

IDENTIFICATION OF GENES INSTRUMENTAL IN PATHOGENESIS OF OLIGODENDROGLIOMA

By

Vijay Govindrao Padul

LIFE09201104007

Tata Memorial Centre,

Navi Mumbai

*A thesis submitted to the
Board of Studies Life Sciences*

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

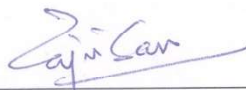
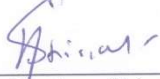
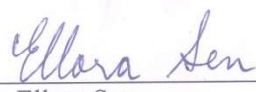
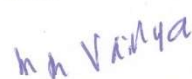
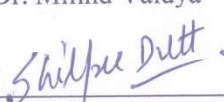

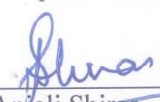


November 2017

Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mr. Vijay Govindrao Padul entitled "Identification of Genes Instrumental in Pathogenesis of Oligodendroglioma." and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

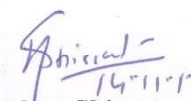
 Chairman – Dr. Rajiv Sarin	14.11.17 Date
 Guide / Convener – Dr. Neelam Shirsat	14.11.17 Date
 Examiner – Dr. Ellora Sen	14/11/2017 Date
 Member 1- Dr. Milind Vaidya	15.11.17 Date
 Member 2- Dr. Shilpee Dutt	14/11/17 Date
 Member 3- Dr. Epari Sridhar	14/11/17 Date
 Invitee - Dr. Anyali Shiras	14.11.17 Date

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 14/11/2017

Place: Navi Mumbai


Dr. Neelam Shirsat
Guide

STATEMENT BY AUTHOR


This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.


19.11.17
Vijay Govindrao Padul

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.


Vijay Govindrao Padul

List of Publications arising from the thesis

Journal

1. "ETV/Pea3 family transcription factor-encoding genes are overexpressed in CIC-mutant oligodendrogliomas.", Padul, V., Epari, S., Moiyadi, A., Shetty, P. and Shirsat, N. V., Genes Chromosom. Cancer, 2015, 54., 725–733.

Chapters in books and lectures notes: nil

Conferences

1. Presented paper (Oral) titled "Genomic Analysis of Oligodendroglial Tumors Indicates Significant Role of RTK/RAS/MAPK Pathway in the Pathogenesis of Oligodendrogliomas" at ISNOCON 2016, 8 th annual conference of Indian Society of Neurooncology, held on 1-3 April 2016 at Hyderabad, India.
2. Presented a poster titled "Integrative Genomic Analysis Indicates Central Role of RTK/RAS/MAPK Pathway Activation as a Major Driver of Oligodendroglioma Pathogenesis", at 2nd Mini Symposium on Cell Biology held on 11th May, 2016 at NCCS- National Centre for Cell Sciences, Pune, India.

Others: nil


15.11.17
Vijay Govindrao Padul

DEDICATIONS

Dedicated to All Well Wishers

ACKNOWLEDGEMENTS

आघ्रातं परि चुम्बितं ननु मुहुर्लीढम् ततः चर्वितं.
त्यक्तम वा भुवि नीरसेन मनसा तत्र व्यथां मा कृथाः.
हे सद्रत्न, तवैतदेव कुशलं यद्वानरेनादरात्.
अन्तःसार विलोकनव्यसनिना चुर्णीकृतं नाश्मना ॥

I thank all people who helped me complete this work. I thank my Doctoral Committee members Dr. Rajiv Sarin, Dr. Milind Vaidya, Dr. Anjali Shiras, Dr. Epari Sridhar, Dr Shilpee Dutt and Dr. Neelam Shirsat for guidance. I also acknowledge gratitude to all those staff at ACTREC, who extended help and support while performing this work.

CONTENTS

Contents.....	x
SYNOPSIS	1
<i>List of Figures</i>	15
<i>List of Tables</i>	17
<i>List of Abbreviations</i>	19
1 INTRODUCTION.....	23
2 REVIEW OF LITERATURE	34
3 MATERIALS AND METHODS.....	60
3.1 LABORATORY METHODS.....	60
3.1.1 Collection of Tumor Tissues and Paired Blood specimens.....	60
3.1.2 Genomic DNA Isolation from Brain tumor tissues.....	60
3.1.3 Genomic DNA Isolation from Blood	63
3.1.4 Total RNA Isolation from Brain Tumor Tissues	66
3.1.5 DNA Agarose Gel Electrophoresis.....	71
3.1.6 DNA Polyacrylamide Gel Electrophoresis [DNA PAGE]	73
3.1.7 DNA Quantification Using Qubit Fluorometer	76
3.1.8 Polymerase Chain Reaction	78
3.1.9 Real Time PCR for Library Quantification	80

3.1.10 Genomic DNA Fragmentation	83
3.1.11 Genomic DNA Libray Preparation.....	85
3.1.12 Exome Enrichment	94
3.1.13 RNAseq Library Preparation	113
3.1.14 Targeted Sequencing on Ion Torrent Platform	131
3.2 BIOINFORMATIC METHODS	133
4 RESULTS	158
4.1 Detection of <i>IDH1</i> and <i>IDH2</i> Mutation status by Sanger Sequencing.	158
4.2 Loss of Heterozygosity analysis for chromosomal arms 1p/19q in oligodendroglial tumours.	162
4.3 Identification of Somatc Mutations and Copy Number Variations by Exome Sequencing of Oligodendroglial Paired Tumour-Normal Samples.....	164
4.4 Comparative Transcriptome Analysis.....	179
4.5 In Silico Analysis of HMG Domain	203
4.6 Mutation Based Stratification of Copy Number Alterations of TCGA Low Grade Glioma Tumours.	210
4.7 Targeted Sequencing of Oligodendroglial Tumour Samples.	213
5 DISCUSSION	218
6 SUMMARY AND CONCLUDING REMARKS	236
BIBLIOGRAPHY	244
APPENDIX – I: PCR Primer List.....	267
Appendix – II: Exome Sequencing Data Quality	269

Appendix - III: RNAseq Data Quality	285
Appendix - IV: Somatic Mutations in 11 Tumours	291
Appendix - V: TCGA Sample Information	317
Appendix - VI: Tables of Differentially Expressed Genes	323
PUBLICATION	333

SYNOPSIS



Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

- 1. Name of the Student: Vijay Govindrao Padul**
- 2. Name of the Constituent Institution: Tata Memorial Centre, ACTREC**
- 3. Enrolment No. : LIFE09201104007 Enrolled on 25/07/2011**
- 4. Title of the Thesis: Identification of Genes Instrumental in Pathogenesis of Oligodendroglioma**
- 5. Board of Studies: Life Sciences**

SYNOPSIS

1. INTRODUCTION

Brain tumours are a major cause of deaths resulting from cancer in children and adults. Gliomas account for almost 80% of primary malignant brain tumours [1]. Diagnosis of gliomas still primarily depends on histopathologic analysis of H & E stained slides of tumour tissue sections [2]. Histopathologic diagnosis is particularly challenging when it comes to reliably distinguishing between oligodendroglial and astrocytic component in low grade to intermediate grade gliomas. Accurate diagnosis of glioma subtypes is not just of academic interest but is necessary for deciding treatment strategy and for prognostication. Oligodendrogliomas generally have slower growth rates and have better prognosis than astrocytomas of similar grade [3]. Oligodendrogliomas especially those having combined loss of chromosome 1p and 19q are known to be particularly sensitive to chemotherapy with much longer progression free survival [4]. Combined loss of chromosome 1p and 19q is known to occur in about 60-90%

oligodendrogliomas [2, 5]. *CIC* on chromosome 19q and *FUBP1* gene on chromosome 1p has been reported to be in about 50% and 25% of 1p/19q codeleted Oligodendrogliomas respectively [6, 7]. In order to profile genetic alterations in oligodendrogliomas, Exome sequencing of histopathologically diagnosed oligodendrogliomas was done. Further transcriptome analysis was done in parallel to understand the role of *CIC* gene in oligodendroglioma pathogenesis

2. OBJECTIVE

Identification of genetic alterations in oligodendroglial tumours.

3. WORK DONE AND RESULTS

As a part of the study till date 11 paired tumour-normal samples have been Exome sequenced. Nine samples have been used for transcriptome sequencing.

Procurement of tumour tissues: The study is approved by the Tata Memorial Centers Institutional Ethics Committee III. The patients were recruited for the study after administering Informed Consent of the patient. Fresh tumour tissues were acquired following surgery and snap frozen in liquid Nitrogen. A part of tissue is acquired fresh for cell culture to establish primary culture. Five ml peripheral blood is also collected and stored at -80°C.

Genomic DNA extraction: Genomic DNA of histopathologically diagnosed 11 Oligodendroglioma tumours (after ensuring 70-80% tumour cell content by Hematoxyline & Eosin staining) and corresponding paired blood was isolated using QIAamp DNA mini kit (Qiagen).

IDH1/2 mutation Analysis: To analyze the status of *IDH1/2* mutation in the tumour samples corresponding exon 4 was sequenced by sanger sequencing method. *IDH1* mutation was identified in 9 out of 11 tumour samples analyzed. One sample harbored *IDH2* mutation while one sample was wild type for both R132 of *IDH1* and R172 of *IDH2*.

Loss of Heterozygosity Analysis by Microsatellite Marker study: Loss of heterozygosity analysis was carried out to assess the status of chromosome 1p and 19q by evaluating five length polymorphic microsatellite markers on chromosome 1p and two on chromosome 19q. Five tumours showed LOH on 1p while seven tumours showed LOH on 19q.

Exome Sequencing: For exome capture Agilent's SureSelect 44Mb Exome capture kit was used at BGI for five paired samples whereas Illumina's TrueSeq 62Mb Exome capture kit was used for six paired samples which were sequenced in-house at ACTREC on Illumina Hiseq 1500. The captured fragments were purified, PCR amplified and quantified by real time PCR (qPCR). The pooled multiplexed libraries were used for cluster generation on Illumina Flow Cell. The flow cell was used for massively parallel sequencing on Illumina Hiseq 1500. Sequence data generated in the form of base call (.bcl) files was demultiplexed to generate raw fastq read files.

Bioinformatic Analysis of the Exome Sequencing Data: The FASTQ Reads of the exome sequence data were aligned to the human reference genome hg19 using the Burrows-Wheeler Aligner version 0.7.9 (www.bio-bwa.sourceforge.net) with default parameters [8]. Duplicate reads were removed using the Picard Tools version 1.80 (<http://broadinstitute.github.io/picard>). The alignment files were refined by local realignment of the reads and base quality recalibration by The Genome Analysis Toolkit (GATK) version 2.1.3 (<https://www.broadinstitute.org/gatk>) [9]. Exome enrichment analysis of the binary reads alignment (BAM) files was done using the

NGSrich Version 0.7.8 (<http://sourceforge.net/projects/ngsrich>). Somatic Single Nucleotide Variants (SNVs) and Insertions and Deletions (Indels) were identified using the VarScan variant detection tool version 2.3.5 (<http://varscan.sourceforge.net>) [10] using the filtering criteria of a minimum coverage 10 and at least 5 somatic variants. Functional annotation of the somatic variant list was done using the ANNOVAR annotation software (www.openbioinformatics.org/annovar) (Wang et al., 2010) [11]. From the ANNOVAR annotated list, variants located in the segmental duplications were excluded. The remaining variants were manually verified in IGV (www.broadinstitute.org/igv). Ambiguous variants (variants represented in reads with low mapping quality, variants present near indels and variants surrounded by mismatched bases) were discarded. The raw fastq files were quality checked by FastQC [17]. The overall base quality of all bases were above 25 and most bases had quality above 35. With no kmer content and no overrepresented sequences the data did not need any QC downstream processing. Post alignment rate of duplication was found to be 2-3% only. Mean exome coverage obtained ranged from 50X to 80X. Total of 319 somatic mutations were identified in 11 tumour samples. Per sample simple somatic nucleotide variations ranged from 12 to 46. All these mutation were manually verified in IGV.

Identification of somatic mutations: Previously reported as well as novel somatic variants have been identified in the exome sequencing data. *CIC*, a gene on chromosome 19q was found to be mutated in four out of nine 1p/19q codeleted tumours. *FUBP1* a gene located on chromosome 1p was identified to be mutated in two tumours, one of which also harbored *CIC* mutation.

Activating mutation (Q61L, G12D) in *KRAS* gene was identified in two tumours with chromosome 1p/19q codeletion which had no alteration in the *CIC* gene. Recurrent mutations were identified in the Notch signaling pathway genes including four tumours with mutation in

NOTCH1 gene and, one tumour with a mutation in *MAML3* gene. Two tumours were found to carry a mutation in the chromatin modifier *ARID1A* gene.

Validation of Selected Mutations: Selected significant genetic alteration were validated by by Sanger sequencing and other were verified in the RNA-seq data.

Copy Number Variation Analysis: The copy number variations in the tumour genome were identified from the paired exome sequence data using the FishingCNV software version 1.5.2 (<http://sourceforge.net/projects/fishingcnv>) [12]. The segmentation means of < -0.3 and > 0.3 were considered as deletion and amplification respectively. The copy number variations in the tumour genome were also analyzed using the Control-FREEC software [<http://bioinfo-out.curie.fr/projects/freec/>]. Coverage based as well as B-allele frequency based methods identified somatic 1p/19q codeletions in 9 out of 11 tumour-normal pairs studied.

Total RNA extraction and Transcriptome Sequencing: Total RNA from nine oligodendroglial tumours was isolated using Qiagen RNeasy Plus mini kit. Libraries for RNA-sequencing were prepared using TruSeq mRNA library preparation kit from Illumina. From 4µg of total RNA, poly adenylated RNA was purified with poly-T RNA purification magnetic beads using two rounds of purification. Purified poly-A RNA was chemically fragmented. The cleaved RNA fragments were primed with random hexamers and reverse transcribed into first strand cDNA. Subsequently RNA template was removed and a replacement strand was synthesized to generate double stranded cDNA. The double stranded cDNA was end repaired and adenylated at 3' ends. Indexed paired end adapters were ligated to the fragments. The ligated DNA fragments were purified, PCR amplified, quantified and validated using qPCR. Sequencing on Hiseq 1500 was performed same as that for exome. Nine tumour transcriptomes were sequenced.

Bioinformatic Analysis of Transcriptome Sequence Data: The reads of the RNA sequencing data containing adapter overlaps were cleaned using the reads trimming tool Trimmomatic version 0.32 (<http://www.usadellab.org>). The cleaned reads were aligned to the reference human genome hg19 using the TopHat version 2.0.13 (<http://ccb.jhu.edu/software/tophat>) with default parameters [13]. Raw counts for the reads aligned to the gene intervals were produced by the python package HTSeq version 0.6.1 (www-huber.embl.de/users/anders/HTSeq) using the default union-counting mode [14]. The read count based gene level differential expression analysis comparing the transcriptome profiles of the *CIC*-mutant vs *CIC*-wild type oligodendrogliomas was carried out using the EdgeR package of R bioconductor (www.bioconductor.org) [15].

Analysis of the TCGA Data on Low Grade Gliomas: A total of 65 *IDH1/IDH2*-mutant, chromosome 1p/19q codeleted oligodendroglioma tumours for which the RNAseq V2 data was available, were used for the differential gene expression analysis comparing the transcriptome profiles of the *CIC*-mutant (39) and the *CIC*-wild type (26) tumour tissues. The gene level RSEM raw counts from the TCGA RNAseq V2 data were rounded to the nearest integer. The data was normalized by variance stabilizing transformation using the DESeq software that takes into account RNA-seq data size of each sample (<http://bioconductor.org/packages/release/bioc/html/DESeq.html>) [16]. The differential gene expression in the *CIC*-mutant vs *CIC*-wild type oligodendrogliomas was analyzed using the Significance Analysis of Microarrays (SAM) tool in the MeV version 4.9.0 (www.TM4.org). The gene set enrichment analysis was carried out using the Web based Gene SeT AnaLysis Toolkit (<http://bioinfo.vanderbilt.edu/webgestalt>). SAM analysis identified 148 genes to be significantly differentially expressed in the 39 *CIC*-mutant oligodendrogliomas as compared to

the 26 *CIC*-wild type oligodendrogliomas from the TCGA cohort at a False Discovery Rate of less than 5%. The differential gene expression comparing the *CIC*-mutant vs. *CIC*-wild type oligodendrogliomas from our cohort as well as the TCGA cohort was also done using EdgeR analysis. The genes identified to be significantly differentially expressed in the TCGA cohort showed differential expression in our cohort as well, although some genes did not reach statistical significance due to the small sample size. *ETV1*, *ETV4* and *ETV5*, the three genes belonging to the *ETS/PEA3* family of transcription factors were found to be upregulated in the *CIC*-mutant tumours. The gene set enrichment analysis identified a number of genes involved in the MAP kinase (MAPK) signaling pathway to be significantly enriched ($P = .0039$) in the *CIC*-mutant oligodendrogliomas. These MAPK pathway genes included *DUSP4*, *DUSP6*, *DUSP19*, the dual specificity phosphatase genes, Sprouty family members *SPRY4*, *SPRED1* and *SPRED2* and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, *EPHB*. The gene set enrichment analysis was carried out using the SeqGSEA package (1.8.0 version) of the R bioconductor (www.bioconductor.org). The gene set enrichment analysis of the TCGA data comparing the expression profiles of the *CIC*-mutant and *CIC*-wild type tumours identified a number of genes involved in the negative regulation of the MAP kinase (MAPK) signaling pathway and those upregulated by the *KRAS* oncogene to be significantly enriched.

In silico analysis of effect of *CIC* gene mutations: An *in silico* analysis of sequence and X-ray crystal structure of conserved HMG domain bound to DNA was carried out using VMD software (<http://www.ks.uiuc.edu/Research/vmd/>). The analysis indicated that the *CIC* gene mutations affect highly conserved amino acids of HMG domain which might impair DNA binding of *CIC* protein

Mutation Based Stratification of Copy Number Alteration: Mutation based stratification of CNA was performed on histologically classified low grade glioma samples available from TCGA dataset. The samples for which both mutation and CNA data was available were used for the analysis. Three major subgroups emerged as a result of stratification analysis ie 1) *IDH1/2&TP53* mutated, 2) *IDH1/2* mutated with chromosome 1p-19q codeleted and 3) *IDH1/2* wild type. The *IDH* wild type samples were found to have molecular alterations similar to that of GBM. This analysis shows that molecular classification of low grade glioma is better than classification based on histology.

Establishment of Oligodendroglioma Primary Cell Cultures: Fresh tumour tissues acquired from ACTREC neurosurgery department were used for establishment of primary cell cultures with intention to develop cell line. Three oligodendroglioma tumour tissue derived primary cultures were established but none of these grew beyond five passages as the cells tended to attain senescence.

Targeted Sequencing Using Ion Torrent Ampliseq: Two oligodendroglioma tumour samples which were already exome sequenced, were used for targeted sequencing of cancer hotspots using Ion AmpliSeqTM Cancer Hotspot Panel v2 on Ion PGM. Ion 314TM Chip v2 was used for sequencing which has 30-50 MB output capacity at 200 bp sequencing length. Ion AmpliSeqTM Cancer Hotspot Panel targets 207 amplicon regions located on 50 oncogenes and tumour suppressor genes with amplicon length in the range of 111–187 bp. Twelve variants were identified for ODG10 of which 2 were found to be hotspot variants. Eighteen variants were identified for ODG11 of which 4 were found to be hotspot variants. All these variants were present in exome data as well.

4. SUMMARY AND CONCLUSIONS

Oligodendroglioma is a histologically defined subtype of glioma along with Astrocytoma and Oligoastrocytoma. Histopathologic diagnosis is particularly challenging and there is intraobserver and interobserver variability in subtype diagnosis. Accurate diagnosis of glioma subtypes is necessary for deciding treatment strategy and for prognostication. Identification of molecular markers specific for glioma subtypes is necessary for resolving the diagnostic dilemma. In order to identify molecular markers specific for Oligodendroglioma, genomic DNA from eleven tumour samples histologically diagnosed as oligodendroglioma and paired normal blood were exome sequenced. Ten of these tumour samples were found to have *IDH* mutations, nine of which had a combined chromosome 1p-19q codeletion. Among the *IDH*-mut, 1p-19q codeletion samples four were found to have *CIC* mutations while two samples without *CIC* mutations had activating *KRAS* mutations. Other genes with recurrent sequence alterations included *FUBP1*, *NOTCH1* and *ARID1A*. One Sample with *IDH1* Mutation was found to have *TP53* and *ATRX* gene mutation. One *IDH* wild type sample was found to have *NF1* mutation, *PDGFRA* amplification and *CDKN2A* deletion. Mutation based stratification of copy number alteration (CNAs) of samples from TCGA dataset segregated low grade gliomas in three major molecular subgroups 1) *IDH* and *TP53* mutated 2) *IDH* mutated with combined chromosome 1p-19q codeletion 3) *IDH* wild type tumours. This indicates that molecular marker based classification of glioma is better than histology based classification. Transcriptome sequencing was performed on nine chr1p-19q codeleted samples. Differential gene expression analysis was carried out for *CIC*-mutant as compared to *CIC*-wild type *IDH1/2*-mutant, chromosome 1p/19q codeleted tumour tissues from study cohort (7 samples) as well as 65 tumours from the TCGA data set. The differential gene expression revealed upregulation of *ETV/Pea3* family transcription factor-

encoding genes *ETV1*, *ETV4* and *ETV5* in *CIC* mutated tumours. Also the negative regulators of tyrosine kinase receptor signaling pathway such as *SPRY4*, *SPRED1*, *SPERD2*, *DUSP4*, and *DUSP6* were found to be upregulated in *CIC* mutated Tumours which is likely due to constitutive activation of inactive *CIC* protein. Higher expression of oncogenic *ETV* transcription factors in the *CIC*-mutant oligodendrogliomas may make these tumours more aggressive than *CIC*-wild type tumours. The study indicates *RTK/RAS/MAPK* pathway activation as a driver of Oligodendroglioma pathogenesis.

References:

1. V. P. Collins, Brain tumours: classification and genes, J Neurol Neurosurg Psychiatry 75 Suppl 2 (2004), ii2-11.
2. M. Gupta, A. Djalilvand, and D. J. Brat, Clarifying the diffuse gliomas: an update on the morphologic features and markers that discriminate oligodendroglioma from astrocytoma, Am J Clin Pathol 124 (2005), 755-68.
3. M. J. van den Bent, Advances in the biology and treatment of oligodendrogliomas, Curr Opin Neurol 17 (2004), 675-80.
4. J. E. Bromberg, and M. J. van den Bent, Oligodendrogliomas: molecular biology and treatment, Oncologist 14 (2009), 155-63.
5. H. H. Engelhard, A. Stelea, and E. J. Cochran, Oligodendroglioma: pathology and molecular biology, Surg Neurol 58 (2002), 111-7; discussion 117.
6. Bettegowda, C., Agrawal, N., Jiao, Y., Sausen, M., Wood, L.D., Hruban, R.H., Rodriguez, F.J., Cahill, D.P., McLendon, R., Riggins, G., Velculescu, V.E., Oba-Shinjo, S.M., Marie, S.K., Vogelstein, B., Bigner, D., Yan, H., Papadopoulos, N., Kinzler, K.W. 2011. Mutations in *CIC* and *FUBP1* contribute to human oligodendroglioma. Science 333: 1453-1455.
7. Sahm, F., Koelsche, C., Meyer, J., Pusch, S., Lindenberg, K., Mueller, W., Herold-Mende, C., von Deimling, A., Hartmann, C. 2012. *CIC* and *FUBP1* mutations in

- oligodendrogliomas, oligoastrocytomas and astrocytomas. *Acta Neuropathol* 123: 853–860.
8. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.
 9. McKenna, A.H. et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303 (2010).
 10. Koboldt, D., Chen, K., Wylie, T. & Larson, D. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25, 2283–2285 (2009).
 11. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164 (2010)
 12. Shi, Y. & Majewski, J. FishingCNV: a graphical software package for detecting rare copy number variations in exome-sequencing data. *Bioinformatics* 29, 1461–1462 (2013).
 13. Trapnell, C., Pachter, L. & Salzberg, S. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111 (2009).
 14. Anders S, Pyl PT, Huber W: HTSeq - A Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015, 31:166.
 15. Robinson M., McCarthy, D., Smyth G. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009, 26:139–140.
 16. Anders S, Huber W (2013) Differential expression of RNA-Seq data at the gene level—The DESeq package.
www.bioconductor.org/packages/devel/bioc/vignettes/DESeq/inst/doc/DESeq.pdf.

17. Andrews S. FastQC A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
18. Padul, V., Epari, S., Moiyadi, A., Shetty, P. and Shirsat, N. V. (2015), ETV/Pea3 family transcription factor-encoding genes are overexpressed in CIC-mutant oligodendrogliomas. Genes Chromosom. Cancer, 54: 725–733. doi: 10.1002/gcc.22283

Publications in Refereed Journal:

a. Published:

Padul, V., Epari, S., Moiyadi, A., Shetty, P. and Shirsat, N. V. (2015), ETV/Pea3 family transcription factor-encoding genes are overexpressed in CIC-mutant oligodendrogliomas. Genes Chromosom. Cancer, 54: 725–733. doi: 10.1002/gcc.22283

b. Accepted: NIL

c. Communicated: NIL

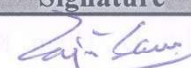
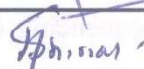
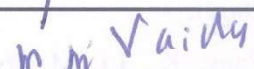
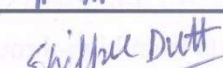
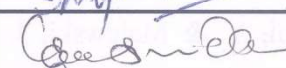
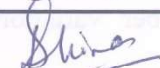
Other Publications: a. Book/Book Chapter: NIL b. Conference/Symposium: NIL

1. **Presented Paper (Oral)** titled “Genomic Analysis of Oligodendroglial Tumors Indicates Significant Role of RTK/RAS/MAPK Pathway in the Pathogenesis of Oligodendrogliomas” at ISNOCON 2016, 8 th annual conference of Indian Society of Neurooncology, held on 1-3 April 2016 at Hyderabad, India.
2. **Presented a Poster** titled “Integrative Genomic Analysis Indicates Central Role of RTK/RAS/MAPK Pathway Activation as a Major Driver of Oligodendroglioma Pathogenesis”, at 2nd Mini Symposium on Cell Biology held on 11th May, 2016 at NCCS- National Centre for Cell Sciences, Pune, India.

Signature of Student: 

Date: 17/02/2017

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. Rajiv Sarin	Chairman		17-2-17
2.	Dr. Neelam Shirsat	Guide/Convener		17-2-17
3.	Dr. Milind Vaidya	Member		17/2/17
4.	Dr. Shilpee Dutt	Member		17/2/17
5.	Dr. Epari Sridhar	Member		17/02/17
6.	Dr. Anjali Shiras	Member		17/2/17

Forwarded through:



To 
Dean-Academic

Dr. K. Sharma,
Director, Academics, T.M.C.


Dr. S.V. Chiplunkar

Director, ACTREC

Chairperson, Academic & Training Programme
ACTREC

Dr. S. V. Chiplunkar

Director

Advanced Centre for Treatment, Research &
Education in Cancer (ACTREC)

Tata Memorial Centre

Kharghar, Navi Mumbai 410210.

List of Figures

Figure 2.1: Molecular alterations in glioma subgroups.....	45
Figure 4.1: <i>IDH1</i> / <i>IDH2</i> mutation status of the studied tumours with oligodendroglioma morphology.....	161
Figure 4.2: Chromosomal Locations of the microsatellite markers used for loss of heterozygosity analysis.....	162
Figure 4.3: Integrated genomic view of the copy number variation and somatic mutations of the 11 oligodendrogliomas as analyzed using the FishingCNV software.....	169
Figure 4.4: Copy number and allelic content profiles from exome data with predicted regions of genomic alteration such as gains, losses and LOH.....	170
Figure 4.5: Bar chart showing different types of somatic mutations identified in the oligodendroglial tumour samples.....	171
Figure 4.6: Electropherograms of the Sanger sequencing of <i>CIC</i> exon 5 from three mutated samples.....	178
Figure 4.7: Coverage of High, Medium and Low expressed transcripts across samples.....	183
Figure 4.8: RNAseq differential gene expression plot of sample relation and MA plots – For present study (ACTREC) cohort by EdgeR.....	185
Figure 4.9: RNAseq differential gene expression plot of sample relations and MA plot for TCGA cohort by EdgeR.....	188
Figure 4.10: RNAseq differential gene expression MA plot, plot of sample relations and p-value histogram for TCGA cohort by DEseq.....	192
Figure 4.11: Heat map showing the top 50 genes most significantly differentially expressed in the 39 <i>CIC</i> -mutant vs. 26 <i>CIC</i> -wild type oligodendrogliomas.....	194

Figure 4.12: The gene sets significantly enriched in the differential expression analysis comparing the <i>CIC</i> -mutant and <i>CIC</i> -wild type tumors in the TCGA.....	202
Figure 4.13: <i>CIC</i> gene mutations from current study and those reported by different groups mapped on <i>CIC</i> gene.....	204
Figure 4.14: BLAST results for the <i>CIC</i> HMG box domain amino acid sequence.....	204
Figure 4.15: FRpred amino acid sequence conservation analysis results.....	205
Figure 4.16: Multiseq Alignment of the <i>CIC</i> HMG box domain amino acid sequence and amino acid sequence from seven HMG box domain crystal structures.....	207
Figure 4.17: Cartoon representation of the structural alignment of seven HMG domains. Left images shows the seven aligned HMG domains.....	208
Figure 4.18: Breakup of 251 TCGA low grade glioma samples according to molecular subgroup and histology as a result of mutation based stratification of copy number alteration.....	211
Figure 4.19: Integrated genomic view of mutation based stratification of copy number alteration of 251 TCGA low grade glioma.....	212
Figure 5.1: A schematic representation of RTK/RAS/MAPK pathway mediated control over downstream <i>CIC</i> repressor protein and inhibitors of components of RTK/RAS/MAPK pathway.....	232

List of Tables

Table 2.1: Frequencies of selected molecular abnormalities among grade II gliomas.....	44
Table 4.1: Histology, WHO grade and <i>IDH1/IDH2</i> mutation status of 11 tumour samples studied.....	159
Table 4.2: Summary of Microsatellite marker loss of heterozygosity on chromosomal arm 1p and 19q.....	163
Table 4.3: Exome Sequencing Data Statistics of 22 exomes.....	166
Table 4.4: The list of the selected somatic alterations in the 11 oligodendroglial tumours identified by the VarScan analysis of the exome sequence data.....	174
Table 4.5: RNA seq Data Statistics.....	181
Table 4.6: RNA seq Data Statistics (Transcript-associated Reads).....	182
Table 4.10 List of the 148 genes significantly differentially expressed (FDR< 0.05) in the <i>CIC</i> -mutant as compared to the <i>CIC</i> -wild type, 1p/19q codeleted, <i>IDH1/IDH2</i> -mutant oligodendrogliomas from the TCGA cohort and the study cohort.....	195
Table 4.11: X-ray crystal structures of HMG domain from different genes and organisms.	206
Table 4.12: The amino acids in SOX17 HMG domain that are closest to the bound DNA.....	209
Table 4.13: Data statistics for Ion torrent run.....	213
Table 4.14: Variants detected in sample ODG10.....	214
Table 4.15: variants detected in sample ODG11.....	215
Table 4.7: List of 106 genes significantly differentially expressed at 5% false discovery rate from the differential expression analysis between <i>CIC</i> mutated and <i>CIC</i> wild type samples from the study cohort using EdgeR.....	323

Table 4.8: List of 158 genes significantly differentially expressed at 5% false discovery rate from the differential expression analysis between *CIC* mutated (n=39) and *CIC* wild type (n=26) samples from TCGA cohort using EdgeR.....326

Table 4.9: List of 76 genes significantly differentially expressed at 10% false discovery rate from the differential expression analysis between *CIC* mutated and *CIC* wild type samples from TCGA cohort using DEseq.....330

List of Abbreviations

AFA - Adaptive focused acoustics

APS - Ammonium persulfate

BLAST - Basic local alignment search tool

CBTRUS - Central brain tumour registry of the United States

CNS - Central nervous system

CNV - Copy number variations

d-2HG - d-2-hydroxyglutarate

DEPC - Diethyl pyrocarbonate

EDTA - Ethylenediaminetetraacetic acid

FDR - False discovery rate

FISH - Fluorescent in situ hybridization

GBMO - GBM with oligodendroglial component

G-CIMP - Glioma CpG island methylator phenotype

H & E - Hematoxylin and eosin

HMG - High mobility group

IDH - IDH1 or IDH2

LMP - Low melting point

LOH - Loss of heterozygosity

MAPK - Mitogen-activated protein kinase

PAGE - Polyacrylamide gel electrophoresis

PCR - Polymerase chain reaction

PCV - Procarbazine, Lomustine (CCNU), Vincristin

RPKM - Reads per kilobase per million mapped reads

RTK - Receptor tyrosine kinase

TAE - Tris acetate EDTA buffer

TCGA - The Cancer Genome Atlas

TEMED - Tetramethylethylenediamine

WHO - World health organization

INTRODUCTION

1 INTRODUCTION

Tumours of the central nervous system are a major cause of deaths resulting from cancer in children and adults. Gliomas account for 29% of all primary brain and other CNS tumours, and almost 80% of primary malignant brain tumours [1].

Gliomas are classified as astrocytic, oligodendroglial, or ependymal depending on the resemblance of tumour cell morphology to the three glial cell types *viz.* astrocytes, oligodendrocytes, and ependymal cells which are presumed to be cell of origin of these tumours [2]. The current WHO classification of primary CNS tumors recognizes four separate tumor grades (I–IV), which can be grouped into low-grade (I and II) or high-grade (III and IV) categories depending on the presence or absence of high-grade features, such as microvascular proliferation and necrosis [3], [4]. Oligodendroglial tumours are classified as Oligodendroglioma (Grade II) and Anaplastic Oligodendroglioma (Grade III). Astrocytic tumours are classified as Pilocytic Astrocytoma (Grade I), Diffuse Astrocytoma (Grade II), Anaplastic Astrocytoma (Grade III) and Astrocytoma/Glioblastoma (Grade IV). Gliomas with histological appearance intermediate to that of oligodendroglioma and astrocytoma are classified as mixed Oligoastrocytoma (Grade II) and Anaplastic Oligoastrocytoma (Grade III) [4]. Pure oligodendrogliomas are composed of single cell type while mixed oligoastrocytomas have morphological characteristics of both pure oligodendrogliomas and astrocytomas. Glioblastoma (WHO grade IV) can be subdivided into primary and secondary tumours. Primary glioblastoma tumours present *de novo* without a preexisting lower grade glioma, and they account for approximately 90% of all glioblastoma tumours. Secondary glioblastoma tumours arise from a preexisting Grade II or III astrocytoma or from a mixed glioma (oligoastrocytoma) [5].

According to 2016 CBTRUS statistical report glioblastomas accounted for 55.4% of primary brain and other CNS gliomas, diffuse astrocytomas for 8.1%, anaplastic astrocytomas for 6.3%, oligodendrogliomas 5.6% while oligoastrocytic tumours accounted for 3.2 % during the period of 2009-2013 [6]. Among the subtypes of gliomas, average annual age-adjusted incidence rates of new cases (per 100,000 individuals) during 2009-2013 have been estimated as: oligodendroglioma (0.25), anaplastic oligodendroglioma (0.1), diffuse astrocytoma (0.51), anaplastic astrocytoma (0.39), oligoastrocytic tumours (0.2) and glioblastoma (3.2) [6].

Diagnosis of gliomas primarily depends on histopathologic analysis of H & E (hematoxylin and eosin) stained slides of tumour tissue sections [7], [8]. Histopathologic diagnosis is particularly challenging when it comes to reliably distinguishing between oligodendroglial and astrocytic component in non-classic low grade to intermediate grade gliomas. Several studies have shown considerable intra-observer variation in the diagnosis of astrocytomas, oligodendrogliomas, and oligoastrocytomas [9], [10]. Accurate pathologic diagnosis needs the ability to distinguish astrocytic from oligodendroglial differentiation in histologic sections. This becomes a challenging feat even for the most experienced neuropathologist. High interobserver variability in the diagnosis of diffuse gliomas arises due to, overlapping morphologic features, and variations in training and practice among pathologists [8], [9], [11], [12]. Further histopathological classification does not perfectly predict clinical outcomes [9], [10].

Specific molecular markers are needed for accurate diagnosis of glioma subgroups. Current FISH (chromosome 1p/19q codeletion) and Immunohistochemistry (TP53) based markers assist in tumour classification but are not sufficient. For example, inactivating point mutations in *TP53* are more common in astrocytomas (50% to 60% in grade II) than in oligodendrogliomas (only 5% to 10%) this suggests that it might be a discriminating marker, but p53 immunoreactivity

does not correlate perfectly with the presence of *TP53* mutations (concordance of 65%-75%) [8]. Chromosome 1p/19q codeletions occur more frequently in oligodendrogliomas (39-70%) than in astrocytomas (0-10%) but are also found frequently in oligoastrocytomas (21-59%) [3] thus precluding its use as specific diagnostic marker.

Primary treatment of diffuse grade II and III gliomas is by surgical resection. Depending upon the anatomic location of the tumor and the clinical condition of the patient biopsy, debulking surgery, or maximal resection is undertaken. Aim of surgery is to maximize cytoreduction without causing neurologic deficit [13]. Complete neurosurgical resection of diffuse gliomas is not possible because of their highly invasive nature. The existence of residual tumour may result in recurrence and malignant progression [8]. Along with surgical resection as primary treatment, chemotherapy and/or radiation therapy may be given as a post-surgical adjuvant therapy. Studies have suggested that astrocytomas show poor response to chemotherapy regimens, whereas oligodendrogliomas are sensitive to PCV chemotherapy that consists of Procarbazine, Lomustine (CCNU), and Vincristine [14]–[18]. Thus, a correct diagnosis of oligodendroglioma is important for deciding treatment strategy and for predicting therapy responsiveness.

The World Health Organization (WHO) definition of oligodendroglioma is “a well-differentiated, diffusely infiltrating tumor of adults, typically located in the cerebral hemispheres and composed predominantly of cells morphologically resembling oligodendroglia” [19]. Oligodendrogliomas generally have slower growth rates and have better prognosis than astrocytomas of similar grade [20]. Oligodendrogliomas especially those having combined loss of chromosome 1p and 19q tend to have slower growth rates and are known to be particularly sensitive to chemotherapy with much longer progression free survival [19].

Patients with an astrocytoma (WHO grade II) have an average survival of approximately 7 years, patients with anaplastic astrocytoma have a median survival of 3.5 years, while glioblastoma patients have an average survival of between 9-11 months [2]. The median survival duration for anaplastic oligodendroglioma and anaplastic oligoastrocytoma was found to be $> 6 - 7$ years in the presence of the chromosome 1p/19q codeletion, and 2–3 years in the absence of the codeletion. For Grade II oligodendroglioma or Grade II oligoastrocytoma, median survival time reported is 12–15 years for chromosome 1p/19q codeleted patients and 5– 8 years for patients without the deletion [19]. As the survival probabilities for glioma subtypes are different, accurate diagnosis is needed for correct prognostication and treatment design. There is a need for a better understanding of these neoplasms and a fresh approach to their treatment [8]. Identification of genetic alterations in gliomas is not just of academic interest but is necessary for accurate diagnosis, prognostication, treatment design and development of more effective treatments with least side effects.

Genetic alterations in astrocytic gliomas particularly grade IV Glioblastoma tumours have been characterized in great detail. The Cancer Genome Atlas project has reported integrated analysis of DNA copy number, gene expression profile and promoter methylation alterations in 206 glioblastoma tumours and nucleotide sequence alterations in 601 cancer related genes from 91 glioblastoma tumours [21]. Deregulation of three pathways *viz.* inactivation of tumour suppressive TP53 and pRB pathways and activation of oncogenic RTK/RAS/PI3K signalling pathway drive pathogenesis of glioblastomas. At least one of the genes belonging to each of the three pathways is mutated/deleted or amplified in glioblastoma tumour tissues.

Mutations in *IDH1* gene encoding Isocitrate Dehydrogenase were first identified in the year 2008 during integrated genomic analysis of glioblastomas. Recurrent mutations were identified in the

Arginine (R132) residue located at the active site of *IDH1* gene in about 12% of glioblastomas many of which were secondary glioblastomas progressed from grade II/III diffuse gliomas. Subsequently R132 mutations in *IDH1* gene were identified in about 70% of grade II/III astrocytomas and oligodendrogliomas. Some of the tumors without *IDH1* mutation were found to carry mutation in analogous R172 residue in *IDH2* gene. *IDH1/2* mutations were found to be common in grade II astrocytomas (90%), anaplastic astrocytomas (73%), oligodendrogliomas (84%), anaplastic oligodendrogliomas (94%) and secondary glioblastomas (85%) than in primary glioblastomas (5%) [22]. *TP53* mutations are more common in diffuse astrocytomas (74%), anaplastic astrocytomas (65%), and secondary glioblastomas (62%) than in oligodendrogliomas (16%) or anaplastic oligodendrogliomas (9%) [22].

Way back in 1994, concurrent loss of chromosome 1p arm and that of chromosome 19q arm was reported to occur in oligodendrogliomas. Further rarity of mutations in *TP53* gene and codeletion of chromosome 1p/19q arm was found to be distinctive of grade II oligodendrogliomas and indicated to be an early event in pathogenesis of oligodendrogliomas. Combined loss of chromosome 1p and 19q was found to be associated with prolonged survival in patients having pure oligodendrogliomas irrespective of the tumor grade. Combined loss of short arm of chromosome 1 (1p) and long arm of chromosome 19 (19q) is the typical lesion of oligodendrogliomas that has been reported to occur in 60-90% of oligodendrogliomas and 10-20% of mixed oligoastrocytomas [8], [23]. Patients with oligodendrogliomas having combined loss of chromosome 1p and 19q are reported to have better response to chemotherapy, more indolent clinical course and longer response to radiation treatment [24]. It is not understood why oligodendrogliomas and oligoastrocytomas having combined loss of 1p/19q have longer survival and better response to chemotherapy. Putative tumour suppressor genes *PTCH2*, *KIF1B*, *RIZ1*,

INK4C, *NOTCH2* and the *TP53* related *TP73* gene located on chromosome 1p were found not to be mutated in oligodendroglial tumours [25]–[27]. Till the initiation of the present thesis project in 2010 tumour suppressor gene or genes present on chromosome 1p and 19q which may have undergone mutational inactivation in oligodendrogliomas had not been identified. Compared with low grade oligodendrogliomas, anaplastic oligodendrogliomas were reported to have additional chromosomal alterations. Most common alterations in anaplastic oligodendrogliomas included loss of heterozygosity of chr 9p, deletion of *CDKN2A* gene, deletion of chr 10 and amplification of *EGFR*/ chr 7p [28]. These alterations commonly occur in high grade astrocytic gliomas as well [21].

Several studies have been carried out to identify glioma subtypes by genome-wide expression profiling [21], [29]. Profiling studies have primarily been done on astrocytic gliomas. Gene expression profile of oligodendrogliomas having chromosome 1p/19q deletion resembles normal brain tissue profile (pro-neural) as compared to expression profile described as proliferative and mesenchymal for most high grade astrocytic tumours [30]. Nonetheless many grade III astrocytic tumours also exhibit pro-neural expression profile. Therefore in these circumstances, combined loss of 1p and 19q remained sole molecular diagnostic marker for oligodendroglial tumours [31].

It is necessary to identify genes responsible for pathogenesis of oligodendroglial tumours which would also help in accurate diagnosis and prognostication of oligodendroglial tumours and may lead to novel treatment strategies. With the advent of next generation sequencing technology it is now possible to sequence genome more efficiently and economically [32]. Technologies have recently been developed to capture entire coding portions/exons of all protein coding genes from genomic DNA which can then be sequenced to get the genome-wide mutational spectrum of a

tumour tissue. The present study therefore proposed to identify exome-wide mutational spectrum of oligodendroglial tumours by sequencing exomes from oligodendroglial tumours.

In order to identify specific molecular markers for each glioma subtypes it is prudent to include tumour samples from all subtypes in the study i.e. Astrocytoma, Oligoastrocytoma and Oligodendroglioma. However this substantially increases the scale of the study and also increases budgetary requirements. Since genetic alterations reported in astrocytomas like mutations in TP53 gene overlapped with those in glioblastomas, it was decided to carry out exome sequencing of oligodendrogliomas. High frequency of chromosome 1p/19q codeletion in oligodendroglioma tumours makes it an attractive choice for molecular marker detection. The putative tumour suppressor gene/s residing on the remaining intact 1p/19q chromosomal arm may undergo mutational inactivation and thereby contribute to the pathogenesis of oligodendrogliomas. Chromosome 1p/19q codeletions were reported rarely in astrocytomas but commonly in oligodendrogliomas. This characteristic molecular alteration presented high likelihood of discovery of genes instrumental in pathogenesis of oligodendrogliomas. Thus tumours with classic oligodendroglioma morphology were selected for exome wide analysis to identify molecular marker specific for oligodendroglial tumours.

In the present study, we performed exome sequencing of 11 tumors with classic oligodendroglial morphology and their paired blood samples in order to identify somatic DNA sequence alterations. The exome sequence analysis identified codeletion of chromosome 1p/19q in nine out of eleven oligodendroglioma tumour samples. Among the nine tumors with combined chromosome 1p/19q codeletion the genes that were frequently found to be mutated in tumours include *IDH1/2*, *CIC*, *FUBP1*, *NOTCH1*, *ARID1A* and *KRAS*. One copy of *CIC* gene, located on chromosome 19q and which encodes a transcriptional repressor, is lost due to 1p/19q codeletion

and the other copy may undergo mutational inactivation. In the present study, four out of 9 tumors with 1p/19q codeletion were found to harbor mutation in the *CIC* gene. Comparative analysis of transcriptome profiles of the *CIC*-mutant with those of the *CIC*-wild type, 1p/19q co-deleted oligodendrogliomas from current study cohort as well as 65 tumours from the TCGA (The Cancer Genome Atlas) cohort was carried out in order to understand role of the *CIC* gene in oligodendroglioma pathogenesis. *ETV1*, *ETV4* and *ETV5*, the three genes belonging to the ETS/PEA3 family of transcription factors were found to be upregulated in the *CIC*-mutant tumours in both cohorts. The gene set enrichment analysis identified a number of genes involved in the negative regulation of MAP kinase (MAPK) signaling pathway and genes upregulated by *KRAS* oncogene to be significantly enriched in the *CIC*-mutant oligodendrogliomas. The KEGG pathway analysis of the gene set significantly differentially expressed between *CIC*-mutant and *CIC*-wild type tumours also identified enrichment of a number of genes in the MAPK signaling pathway ($P = 0.0019$ and $FDR = 0.0199$). Activation of the RTK/RAS/MAPK signaling pathway appears to be a major driver of the oligodendroglioma pathogenesis.

Two tumours with oligodendroglioma morphology showed molecular alterations that are known to be associated primarily with astrocytoma (*IDH1* mutation along with *TP53* and *ATRX* mutation in absence of 1p/19q codeletion) and Glioblastoma (*IDH1/2* wild type, *NF1* mutation, *PDGFRA* amplification and *CDKN2A* deletion). Mutation based stratification of copy number alterations was carried out on 251 low grade glioma (WHO grade II and III) tumour sample data from TCGA, for which both somatic mutation and copy number alteration data was available. The 251 samples segregated in three major molecular subgroups 1) *IDH1/2* mutant tumors with 1p/19q codeletion 2) *IDH1/2* mutant tumors without 1p/19q codeletion and 3) *IDH1/2* wild type tumors. These three tumor types were also found in the present study cohort that included only

histologically pure oligodendroglioma tumors. In the TCGA cohort as well, while most histopathologically diagnosed oligodendrogliomas belonged to *IDH1/2* mutant with 1p/19q codeletion tumors subgroup, and most histopathologically diagnosed astrocytomas belonged to *IDH1/2* mutant tumors without 1p/19q codeletion subgroup and *IDH1/2* wild type subgroup tumors a considerable histological heterogeneity was found in each molecular subtype indicating superiority of molecular classification over histological classification.

REVIEW OF LITERATURE

2 REVIEW OF LITERATURE

2.1 GLIOMA HISTORICAL PERSPECTIVE

Bailey and Cushing established the first diagnostic classification system for primary brain tumours in 1926 [33]. Their classification system was based on understanding of the histogenetic basis of brain development and the microscopically observed morphological resemblance of primary brain tumours to their presumed cells of origin such as oligodendrocytes and astrocytes [34].

Bailey, Cushing and Bucy laid the foundation of current concept of oligodendroglioma. In the article “Oligodendrogliomas of the Brain”, Bailey and Bucy, described oligodendrogliomas as tumors, containing cells with nuclei that are almost all perfectly round, of a fairly constant size and surrounded by a ring of cytoplasm which stains very feebly. They also observed that oligodendrogliomas have a network of fine capillaries and are prone to calcification [7], [8].

Bailey and Cushing published results of a long and laborious research on histopathology and classification of the tumours of the glioma group. They used specimens and records of more than 400 cases drawn from the surgical clinics of the Johns Hopkins and Peter Bent Brigham Hospitals over a period of twenty two years. By dividing brain tumours in fourteen categories Bailey and Cushing attempted to add clarity and order to histopathology of brain tumours. They put forth clinical correlation between their proposed classification and the prognosis and operability of the tumours [33], [35]. This system has been refined periodically subsequently evolving in the current World Health Organization (WHO) scheme [4].

The first edition on the histological typing of tumours of the nervous system was published in 1979 [36]. The second edition by Kleihues et al. incorporated the application of

immunohistochemistry into diagnostic pathology [37]. The third edition incorporated genetic profiles as additional aids to the definition of brain tumours, that was edited by Kleihues and Cavenee and published in 2000 [38]. The third edition included brief sections on epidemiology, clinical signs and symptoms, imaging, prognosis and predictive factors. The fourth edition is the 2007 WHO classification of tumours of the central nervous system [39]. In this edition, several new tumor entities were added if accompanied by different age distribution, location, clinical behavior or genetic profile. WHO series on the classification of CNS tumors was always based on the consensus of international Working Groups such that these criteria for classification of tumor types are accepted and used world wide. In 2016, the WHO released updated guidelines for brain tumor classification that combines biology-driven molecular marker diagnostics with classical histological cancer diagnosis [40].

Since Bailey and Cushing's early attempt at classification of brain tumors in 1926 [33], [35] histological examination has been the primary method for risk class assignment, patient outcome stratification, therapy guidelines, and stratification for clinical trials [10], [39].

2.2 GLIOMA CLINICAL PERSPECTIVE

Tumours of the central nervous system are a major cause of deaths resulting from cancer in children and adults. Gliomas account for 29% of all primary brain and other CNS tumours, and almost 80% of primary malignant brain tumours [1]. Four malignancy grades are recognised by the WHO system, with grade I tumours the biologically least aggressive and grade IV the biologically most aggressive tumours [2]. Based on their infiltrative behavior, gliomas are subdivided into two main subgroups: circumscribed and diffuse.

2.2.1 CIRCUMSCRIBED GLIOMAS

These class of gliomas are low grade, better circumscribed glial and glo-neuronal entities, associated with a more indolent clinical course and usually with favorable prognosis are classified as WHO grade I [4], [41]. The circumscribed gliomas are generally amenable to total surgical resection and patients with these tumors have improved outcomes compared to patients with diffuse gliomas [42]. These tumours typically occur in children and adolescents, and include pilocytic astrocytomas, WHO grade I; pleomorphic xanthoastrocytomas (PXA), WHO grade II; and ganglioglioma, WHO grade I [39].

2.2.2 DIFFUSE GLIOMAS

Diffuse gliomas comprise the more aggressive WHO grades II-IV [39]. Due to infiltration in surrounding normal tissues diffuse gliomas are nearly impossible to resect completely [8]. Diffuse gliomas most frequently arise within the cerebral hemispheres of adults [43]. The aggressive phenotype of diffuse gliomas is because of the tendency of the malignant glioma cells to infiltrate surrounding normal tissue and travel far away from the primary tumor site [8]. The diffuse gliomas tend to progress to higher grade with time which makes them eventually lethal, though with variable survival periods for different subtypes [43]. The primary classes of diffuse gliomas are astrocytomas, oligodendrogliomas and oligoastrocytomas, and these are graded according to World Health Organization (WHO) criteria as grades II–IV [39].

Primary treatment of diffuse grade II and III gliomas is through surgical resection where the primary objective is to provide a histologic diagnosis. There are four objectives when performing surgery in diffuse gliomas: histopathological assessment of the nature of the tumour, improvement of the neurological condition of the patient, reducing the risk of tumor recurrence and progression [44]. Depending on the anatomic location of the tumor and the likely nature of the

tumor based upon the radiological features of the tumor and histological characteristics of tumor biopsy debulking surgery, or maximal resection is undertaken. Extent of resection has impact on progression-free and overall survival, thus maximum possible cytoreduction is attempted without causing neurologic deficit [13]. Because of their highly invasive nature complete neurosurgical resection of diffuse gliomas is impossible. The existence of residual tumour often result in recurrence and malignant progression [8]. Along with surgical resection as primary treatment, chemotherapy and/or radiation therapy may be given as a post-surgical adjuvant therapy. Studies have suggested that astrocytomas show poor response to chemotherapy regimens, whereas oligodendrogliomas are particularly sensitive to PCV chemotherapy that consists of Procarbazine, Lomustine (CCNU), and Vincristine [14]–[18].

2.3 HISTOPATHOLOGICAL CLASSIFICATION OF GLIOMA

Gliomas are classified as astrocytic, oligodendroglial, or ependymal depending on the resemblance of tumour cell morphology to the three glial cell types *viz.* astrocytes, oligodendrocytes, and ependymal cells which are presumed to be the cells of origin of these tumours [2]. The current WHO classification of primary CNS tumors recognizes four separate tumor grades (I, II, III, IV), which can be grouped into low-grade (I and II) or high-grade (III and IV) categories depending on the presence or absence of high-grade features, such as microvascular proliferation and necrosis [3], [4]. Oligodendroglial tumours are classified as Oligodendroglioma (Grade II) and Anaplastic Oligodendroglioma (Grade III). Astrocytic tumours are classified as Pilocytic Astrocytoma (Grade I), Diffuse Astrocytoma (Grade II), Anaplastic Astrocytoma (Grade III) and Astrocytoma/Glioblastoma (Grade IV). Gliomas with histological appearance intermediate to that of oligodendroglioma and astrocytoma are classified as mixed Oligoastrocytoma (Grade II) and Anaplastic Oligoastrocytoma (Grade III) [4]. Pure

oligodendrogliomas are composed of single cell type while mixed oligoastrocytomas have morphological characteristics of both pure oligodendrogliomas and astrocytomas. Glioblastoma (WHO grade IV) can be subdivided into primary and secondary tumours. Primary glioblastoma tumours present *de novo* without a preexisting lower grade glioma, and they account for approximately 90% of all glioblastoma tumours. Secondary glioblastoma tumours arise from a preexisting Grade II or III astrocytoma or from a mixed glioma (oligoastrocytoma) over a period of time [5].

2.4 WHO GRADING

Tumour grade is a significant factor that influences the choice of therapies, particularly determining the use of adjuvant radiation and specific chemotherapy protocols [39]. Grade I applies to lesions with low proliferative potential and the possibility of cure following surgical resection alone. Neoplasms designated as grade II are generally infiltrative in nature. Some type II tumours tend to progress to higher grades of malignancy, for example, low-grade diffuse astrocytomas that transform to anaplastic astrocytoma and glioblastoma. Similar transformation occurs in oligodendroglioma and oligoastrocytomas. The designation WHO grade III is generally used for lesions with histological evidence of malignancy like nuclear atypia and brisk mitotic activity. Most patients with grade III tumours receive adjuvant radiation and/or chemotherapy. The designation WHO grade IV is assigned to cytologically malignant, mitotically active, necrosis-prone neoplasms typically associated with rapid pre- and postoperative disease evolution and a fatal outcome. Glioblastoma is an example of grade IV neoplasm [4].

2.5 HISTOLOGICAL SUBTYPES OF GLIOMAS

OLIGODENDROGLIOMAS

Oligodendrogliomas account for 5.6% of primary brain and other CNS gliomas [6]. Oligodendrogliomas occur in adults most commonly in the cerebral hemispheres.. Oligodendrocytes are presumed to be the cells of origin of oligodendrogliomas. The World Health Organization (WHO) definition of oligodendroglioma is “a well-differentiated, diffusely infiltrating tumor of adults, typically located in the cerebral hemispheres and composed predominantly of cells morphologically resembling oligodendroglia” [19]. Histopathologically oligodendrogliomas consist of moderately cellular, monomorphic tumours with round nuclei, often artefactually swollen cytoplasm on paraffin section, few or no mitoses, no florid microvascular proliferation or necrosis, and are classified as malignancy grade II according to the WHO guidelines. Classically they show a “chicken wire” pattern of capillaries. Grade II oligodendrogliomas are relatively indolent, although they usually recur at the primary site [2]. Histopathologically anaplastic oligodendrogliomas (malignancy grade III) show increase in nuclear pleomorphism and hyperchromatism, as well as pronounced hypercellularity, brisk mitotic activity, prominent microvascular proliferation, and/or spontaneous necrosis [2].

Oligodendrogliomas generally have slower growth rates and have better prognosis than astrocytomas of similar grade [20]. Oligodendrogliomas are associated with longer survival than astrocytic gliomas [45]. In a study done by Cairncross *et al.* about 60%–70% of anaplastic oligodendrogliomas were found to be chemosensitive, particularly to the combination of Procarbazine, Lomustine, and Vincristine (PCV) [46]. Oligodendroglioma patients having combined loss of chromosome 1p and 19q have better response to chemotherapy and radiation treatment, more indolent clinical course and longer survival [19], [24]. It is not known why

oligodendrogliomas and oligoastrocytomas with combined loss of 1p/19q have longer survival and better response to treatment. Putative tumour suppressor genes *PTCH2*, *KIF1B*, *RIZ1*, *INK4C*, *NOTCH2* and the *TP53* related *TP73* gene located on chromosome 1p arm were found not to be mutated in oligodendroglial tumours [25]–[27]. Till the initiation of the current study in 2010 tumour suppressor gene or genes present on chromosome 1p and 19q which may have undergone mutational inactivation in oligodendrogliomas had not been identified. Compared to low grade oligodendrogliomas, anaplastic oligodendrogliomas have additional chromosomal alterations. Most common alterations in anaplastic oligodendrogliomas include loss of heterozygosity of chr 9p, deletion of *CDKN2A* gene, deletion of chr 10 and amplification of *EGFR*/ chr 7p [28]. These alterations commonly occur in high grade astrocytic gliomas including glioblastomas [21].

The median survival duration for anaplastic oligodendroglioma and anaplastic oligoastrocytoma was found to be > 6 – 7 years in the presence of the 1p/19q codeletion, and 2–3 years in the absence of the codeletion. For Grade II oligodendroglioma or Grade II oligoastrocytoma, median survival time reported is 12–15 years for 1p/19q codeleted patients and 5– 8 years for patients without the deletion [19].

While *IDH1/2* mutations are more common in oligodendrogliomas (84%) or anaplastic oligodendrogliomas (94%), *TP53* mutations are rare in oligodendrogliomas (16%) or anaplastic oligodendrogliomas (9%) [22]. Combined loss of short arm of chromosome 1 (1p) and long arm of chromosome 19 (19q) is a characteristic of oligodendrogliomas that has been reported to occur in 60-90% of oligodendrogliomas [8], [23]. Recently *CIC*, *FUBP1*, and *TERT* promoter mutations were found to be frequent in oligodendrogliomas [47]–[50].

ASTROCYTOMAS

Diffuse astrocytomas (grade II) account for 8.1% and anaplastic astrocytomas account for 6.3% of primary brain and other CNS gliomas [6]. Astrocytomas are histopathologically classified as the pilocytic astrocytomas (WHO Grade I) and the diffuse astrocytic tumours including astrocytoma (WHO Grade II) and anaplastic astrocytomas (WHO Grade III). Pilocytic astrocytomas most commonly occur in children in the cerebellar region of brain [2]. Pilocytic astrocytomas are generally biologically non-aggressive and maintain their grade I status over years and even decades and can be cured by surgery alone. Rare cases may progress to more malignant tumours [2], [51]. The adult diffuse astrocytic tumours include astrocytomas (WHO malignancy grade II), anaplastic astrocytomas (WHO malignancy grade III), and glioblastomas (WHO malignancy grade IV) [52]–[54].

The astrocytomas (Grade II) have a peak incidence between 25 and 50 years of age. Median survival for patients having grade II astrocytoma is approximately seven years while that for patients with anaplastic astrocytomas is 3.5 years [55].

The tumour cells of astrocytomas (Grade II) resemble astrocytes, show mild nuclear atypia, and have extensions producing a loosely textured matrix. Anaplastic astrocytomas (Grade III) show increased cellularity but the tumour cells still show histological and immunocytochemical characteristics of astrocytes. The tumour cells of anaplastic astrocytomas are more pleomorphic than those found in astrocytomas, show distinct nuclear atypia, with mitotic activity [2]. *IDH1/2* mutations are common in diffuse astrocytomas (90%), anaplastic astrocytomas (73%), [22]. *TP53* mutations are common in diffuse astrocytomas (74%), anaplastic astrocytomas (65%) [22]. Recently frequent *ATRX* gene mutations were reported to be highly correlated with *IDH1/IDH2* and *TP53* mutations in astrocytomas [52]–[54].

OLIGOASTROCYTOMAS

Oligoastrocytomas account for 3.2% of primary brain and other CNS gliomas [6]. Oligoastrocytomas contain distinct regions of oligodendroglial and astrocytic differentiation. Oligoastrocytomas show overlapping histological characteristics to that of oligodendroglioma and astrocytoma. Oligoastrocytomas are classified as Oligoastrocytoma (WHO Grade II) and Anaplastic Oligoastrocytoma (WHO Grade III) [39].

Oligoastrocytomas consist of tumour cells with astrocytic and oligodendroglial morphological characteristics which can be either diffusely mixed or combined as discrete areas in an individual tumour [2]. The morphological distinctions between astrocytomas, oligoastrocytomas, and oligodendrogliomas are difficult and controversial issues. Oligoastrocytomas are graded using same histopathological criteria as oligodendrogliomas [39]. *IDH1/2* mutations are present at high frequency in oligoastrocytomas (50–100%) and *TP53* mutations are also common (44%) in oligoastrocytomas (Table 2.1). Combined loss of short arm of chromosome 1 (1p) and long arm of chromosome 19 (19q) is reported to occur in 10-20% of mixed oligoastrocytomas [8], [23].

GLIOBLASTOMA

Glioblastomas (WHO grade IV) accounts for 55.4% of primary brain and other CNS gliomas [6]. Glioblastomas can be subdivided into primary and secondary tumours. Primary glioblastoma tumours present *de novo* without a preexisting lower grade glioma, and they account for approximately 90% of all glioblastoma tumours. Secondary glioblastoma tumours arise from a preexisting Grade II or III astrocytoma or from a mixed oligoastrocytoma [5]. The glioblastomas have a peak incidence between 45 and 70 years of age [2]. Patients with glioblastoma have a uniformly poor prognosis, with a median survival of 12-14 months [2], [56]. Histopathologically, glioblastomas are more cellular than the anaplastic astrocytomas. The tumour cells show a wide

spectrum of morphologies, can be very pleomorphic with giant forms, but generally retain some of the phenotypical characteristics of astrocytes. Mitosis, spontaneous tumour necrosis with pseudopalisading of tumour cells, as well as florid endothelial proliferation, are inevitably found in a well sampled tumour [2].

IDH1/2 mutations are more frequent in secondary glioblastomas than in primary glioblastomas. Over 80% of secondary glioblastomas possess an *IDH1* mutation [22]. In contrast, *IDH1* and *IDH2* mutations are rarely detected in primary glioblastoma, with a frequency of 3–7% [22], [57], [58]. The frequency of 1p/19q deletions among glioblastoma is low [59]. TP53 mutations are significantly more frequent in secondary glioblastomas (65%) than in primary glioblastomas (28%) [60]. *TERT* promoter mutations are highly frequent in glioblastomas (83%) [48]. *EGFR* (chromosome 7p12) amplification is a hallmark of glioblastoma. The Cancer Genome Atlas (TCGA) project identified copy number alterations and/or amplification of *EGFR* in 45% of glioblastomas [21]. About 40% of primary glioblastoma and over 70% of secondary glioblastoma display MGMT gene promoter methylation leading to gene silencing [61], [62]. Mutations/ amplifications affecting *EGFR*, *PDGFRA*, *FGFR*, *PIK(3)K*, *MDM2*, *MDM4* and *CDK14* genes while mutations/deletions affecting *PTEN*, *CDKN2A/B*, *TP53*, *RBI* are recurrently found in glioblastomas [63].

GBM WITH OLIGODENDROGLIAL COMPONENT (GBMO)

A subset of glioblastomas shows focal oligodendroglial features, suggesting that some glioblastomas may also have an oligodendroglial origin [64]. The presence of an oligodendroglial component in glioblastoma appears to be an important prognostic factor, outcome being better for GBMO than for classic glioblastoma [65]. An oligodendroglial component is detected in 10% of glioblastoma, and these patients are significantly younger and

survive longer [66]. GBMO patients treated with post-operative chemotherapy and radiotherapy have a better prognosis than those having glioblastoma [66].

Molecular Abnormality	Astrocytoma (%) (Grade II)	Oligoastrocytoma (%) (Grade II)	Oligodendroglioma (%) (Grade II)
TP53 mutations	53	44	13
1p/19q codeletion	0–10	21–59	39–70
MGMT hypermethylation	11	27	62
IDH1 mutations	59–88	50–100	68–82

Table 2.1: Frequencies of selected molecular abnormalities among grade II gliomas. Table source (Bourne, 2010) [3].

2.6 FUNCTIONAL SIGNIFICANCE OF MOLECULAR ALTERATIONS IN GLIOMAS

Molecular and genetic features provide additional information which can be utilized to diagnostically differentiate among glioma subtypes and also to predict clinical outcomes and response to adjuvant therapies [34]. In 2014, a group of expert neuropathologists published the International Society of Neuropathology–Haarlem consensus guidelines, with suggestions as to how molecular information could be incorporated in the routine classification of CNS tumors [67].

MUTATIONS in *IDH1/IDH2* GENE

Isocitrate dehydrogenase catalyzes oxidative decarboxylation of isocitrate, producing alpha-ketoglutarate and CO₂ in citric acid cycle. This enzyme has three isoforms encoded by genes *IDH1*, *IDH2* and *IDH3*. *IDH3* is localized in mitochondrial matrix and catalyzes the third step of

the citric acid cycle within mitochondria and reduces NAD^+ to NADH in the process. IDH1 and IDH2 catalyze this same reaction but reduces NADP^+ to NADPH . IDH1 is the only isoform

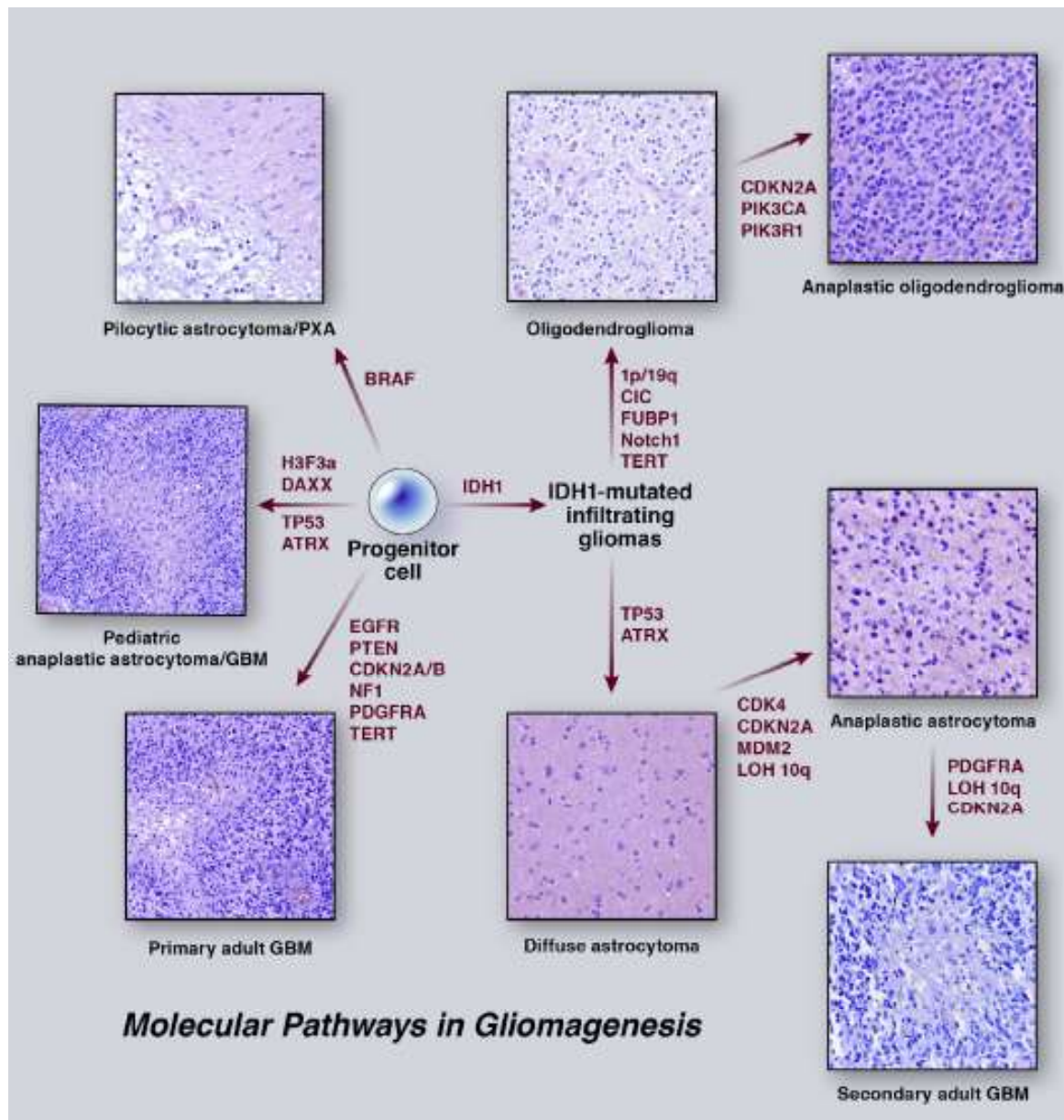


Figure 2.1: Molecular alterations in glioma subgroups. The square images show histological appearance of the given tumor types. The figure describes the different possible molecular alterations a progenitor cell may acquire leading to gliomagenesis and additional molecular alterations acquired by the tumours during progression from lower grades to higher grade. Figure source (Appin, 2015) [43].

localized to the cytoplasm and is also targeted to peroxisomes [68]. IDH1 and IDH2 play role in lipid synthesis, cellular defense against oxidative stress and oxidative respiration.

A genome wide mutational analysis performed by Parsons et al of glioblastomas by sequencing 20,661 protein coding genes revealed somatic mutations at codon R132 located in exon 4 of the *IDH1* gene in about 12% of glioblastomas sequenced. Many of these tumors were known to have evolved from lower-grade gliomas (i. e. secondary glioblastomas) [69]. Subsequently *IDH1* and *IDH2* were found to be mutated in >70% of lower-grade gliomas such as astrocytoma, oligodendroglioma and oligoastrocytoma (WHO grades II and III) [22] (Table 2.1, Figure 2.1). *IDH1* mutations affect the amino acid Arginine at position 132 of the amino acid sequence which belongs to an evolutionary conserved region located at the binding site of isocitrate. Wild type arginine at position 132 is found to be replaced by histidine (R132H) in the vast majority (92-100%) of the cases [57][69]. The mutations reported were always heterozygous [57]. *IDH2* gene mutations affect an analogous position R172 but at a smaller frequency and are mutually exclusive with *IDH1* mutations (i.e *IDH1* and *IDH2* mutation do not occur in the same tumor).

A study by Dang *et al.* demonstrated that *IDH1* mutations result in the production and accumulation of d-2-hydroxyglutarate (d-2HG), which appears to act as an oncogenic metabolite [70]. Mutant *IDH1/IDH2* induces a hypermethylator phenotype, the glioma-CpG island methylator phenotype (G-CIMP) [71], [72]. G-CIMP is characteristic of grade II-III diffuse glioma. G-CIMP is associated with improved prognosis, and with a proneural molecular gene expression profile [73], [74].

Presence of *IDH1/IDH2* mutation is a positive prognostic marker in glioma patients. Patients with *IDH1* or *IDH2* mutated tumors were found to have better outcome than those having wild-type

IDH1/IDH2 genes. The hazard ratio for death among glioblastoma patients with wild-type *IDH1* (n = 79), as compared to those with mutant *IDH1* (n = 11), was 3.7 (95 percent confidence interval, 2.1 to 6.5; P < 0.001) [69]. The median survival was 3.8 years for glioblastoma patients with mutant *IDH1*, as compared to 1.1 years for patients with wild-type *IDH1* [69].

CHROMOSOMES 1p/19q LOSS OF HETEROZYGOSITY

The presence of chromosome 1p/19q co-deletion in oligodendroglioma was first described by Reifenberger et al. in 1994 [75]. Co-deletion of chromosomal arms 1p and 19q is a well known prognostic marker found frequently (>60%) in grade II-III oligodendrogliomas [31], [76], [77]. The combined loss of 1p and 19q is mediated by an unbalanced translocation of 19p arm to 1q arm [76], [78]. presumably a centrosomal or pericentrosomal translocation of chromosomes 1 and 19 results in two derivative chromosomes, der(1,19)(p10;q10) and der(1,19)(q10;p10), after which the derivative chromosome with the short arm (p) of chromosome 1 and the long arm (q) of chromosome 19 is lost [19].

Codeletion of 1p and 19q is highly correlated with a classic oligodendroglioma histological appearance (Figure 2.1), [18], [79], [80]. The 1p/19q codeletion is present in 61%– 89% of anaplastic oligodendroglioma cases, but in only 13%–20% of patients with anaplastic oligoastrocytoma [19]. 1p/19q codeletions are more common in oligodendrogliomas and oligoastrocytomas than in astrocytomas (Table 2.1).

1p/19q codeletion is most commonly assessed by Fluorescence *in situ* Hybridization (FISH) [81] or PCR-based LOH assays [82]. FISH allows pathologists to correlate chromosomal arm copy number findings with tissue morphology without the requirement of normal control samples. Chromosome 1p/19q co-deletion is found only in *IDH1/IDH2* mutant gliomas [83]

indicating an association of 1p/19q codeletion with the G-CIMP and proneural expression phenotypes [73], [74], an association that was demonstrated in grade II-III oligodendrogliomas [73], [84]. Several studies have shown that 1p/19q co-deletion can aid in risk stratification of *IDH1/2* mutant gliomas with *IDH1/2* mutant, 1p/19q co-deleted gliomas having the best prognosis, followed by *IDH1/2* mutant, 1p/19q non-co-deleted gliomas and finally by *IDH1/2* wild-type, 1p/19q non-co-deleted tumors having the worst outcome [73], [85]–[87]. Besides prognostic significance, 1p/19q co-deletion is a marker of chemotherapeutic response [88], [89]. Patients with 1p/19q co-deleted diffuse gliomas responded better to adjuvant chemotherapy PCV or temozolomide [88], [90], [91]. The mechanism for the 1p/19q codeletion associated chemosensitivity remains unknown.

***MGMT* PROMOTER METHYLATION**

Standard therapy for glioblastoma includes radiation and chemotherapy with temozolomide, which acts by crosslinking DNA by alkylating multiple sites including the O6 position of guanine [43], [92]. The DNA repair gene O6-methylguanine DNA methyltransferase (*MGMT*) reverses DNA damage and removes alkyl groups from the O6 position of guanine. Since it is the site of a number of chemotherapy-induced DNA alkylations, *MGMT* interferes with the therapeutic effects of these alkylating chemotherapy agents [93]. Therefore, low levels of *MGMT* lead to enhanced response to alkylating agents. The expression level of *MGMT* is dependant upon the methylation status of the gene's promoter. Methylation of a CpG island is associated with silencing of the gene, thus making it a useful predictor of the responsiveness of glioblastomas to alkylating agents. *MGMT* promoter methylation occurs in up to 40%–50% of glioblastomas [19]. *MGMT* promoter methylation is associated with prolonged progression-free and overall survival in patients with glioblastoma treated with chemotherapy and radiation

therapy [94]. *MGMT* promoter methylation is strongly correlated with 1p/19q codeletion, which was observed in up to 80%–90% of 1p/19q codeleted tumors [95]–[97]. *MGMT* promoter methylations are more common in oligodendrogliomas than in astrocytomas or oligoastrocytomas (Table 2.1).

***TERT* PROMOTER MUTATIONS**

Telomeres are nucleoprotein complexes at the ends of eukaryotic chromosomes that are required for chromosomal integrity. Telomere maintenance is essential for malignant cells as in the case of all actively growing cells. While the ALT (alternative lengthening of telomeres) serves the purpose of maintaining telomere length in astrocytomas, it is rarely seen in oligodendrogliomas. Instead of ALT, activating mutations in the telomerase reverse transcriptase (*TERT*) promoter are present in oligodendrogliomas. *TERT* gene encoded telomerase reverse transcriptase is a part of telomere maintenance mechanism [48]. Activating mutations in the telomerase reverse transcriptase (*TERT*) promoter are present in nearly all oligodendrogliomas (Figure 2.1) [48], [98]. Two *TERT* promoter mutations C228T and C250T corresponding to the positions 124 and 146 bp, respectively, upstream of the *TERT* ATG start site are known to occur [99], [100]. Killela *et al.* evaluated *TERT* promoter mutations in 1,230 specimens of 60 different tumor types and identified a total of 231 mutations (18.8%). C228T and C250T mutations accounted for 77.5% and 20.8% of the alterations, respectively. Point mutations in the promoter of the telomerase reverse transcriptase (*TERT*) gene increases telomerase expression [48]. Among 220 gliomas that were morphologically classified, *TERT* promoter mutations were found to be frequent in glioblastomas (83%) and oligodendrogliomas (78%), but uncommon in grades II and III astrocytomas (10%) and were mutually exclusive with *ATRX* mutations [48]. In

Oligoastrocytomas the prevalence of *TERT* promoter mutations (25% of 24 tumors) was found to be intermediate between oligodendrogliomas and astrocytomas [48].

***TP53* MUTATIONS**

Studies based on morphologic classification have consistently shown that the large majority of grade II and III diffuse astrocytomas and secondary glioblastomas have *TP53* mutations (Figure 2.1), [22], [101], [102]. *TP53* mutations are significantly more frequent in secondary glioblastomas (65%) than in primary glioblastomas (28%)[60]. *TP53* mutations are more frequent in astrocytomas and oligoastrocytomas than in oligodendrogliomas (Table 2.1). When restricted to *IDH1/IDH2*-mutated infiltrating astrocytomas, an even higher percentage have *TP53* mutations, since *IDH1/IDH2* wild-type diffuse gliomas have lower frequencies of *TP53* mutation. In a study of 939 tumors of diverse histologies, 80% of *IDH1/IDH2*-mutated anaplastic astrocytomas and glioblastomas also harbored *TP53* mutations [22]. Conversely, others have shown that 63% of grade II diffuse astrocytomas that contained a *TP53* mutation also harbored an *IDH1/IDH2* mutation [102]. More recently, The Cancer Genome Atlas project demonstrated that 94% of *IDH1/IDH2* mutant grade II and III diffuse gliomas that lacked 1p/19q co-deletion had *TP53* mutations [103].

***ATRX* MUTATIONS**

ATRX (Alpha Thalassemia/Mental Retardation Syndrome X-linked), a member of the SWI/SNF family of chromatin remodelers was found to be mutated in pediatric and adult high-grade gliomas [53], [54]. *ATRX* mutations are strongly associated with the combination of *IDH1/IDH2* and *TP53* mutations (Figure 2.1). In one large analysis of 363 brain tumors, *ATRX* mutations were found to be most frequent in grades II and III astrocytomas and oligoastrocytomas (67–73%) as well as in secondary glioblastomas (57%), but were uncommon in primary

glioblastomas (4%), oligodendroglial tumors (14%) and pediatric glioblastomas (20%) [52]. Nearly all diffuse gliomas with an *ATRX* mutation had an *IDH1/IDH2* mutation as well. In tumors having *IDH1/IDH2* and *ATRX* mutations, 94% had a *TP53* mutation as well. *ATRX* mutations are associated with the Alternative Lengthening of Telomeres (ALT) phenotype [52], an alternative mechanism for maintaining telomere length in tumors that do not have constitutive telomerase activity [48].

***CIC* MUTATIONS**

CIC gene, which encodes a transcriptional repressor, located on chromosome 19q13.2 was found to be mutated in ~60-70% of *IDH1/IDH2* mutated, 1p/19q codeleted oligodendrogliomas (Figure 2.1), [47], [50]. *CIC* mutations are rare in non-1p/19q codeleted gliomas [50]. One copy of *CIC* gene is lost due to chromosome 19q deletion and other copy undergoes either truncating mutations or point mutations affecting DNA interaction domain (encoded by *CIC* exon 5) or protein interaction domain (encoded by *CIC* exon 19-20) [50]. This mutational spectrum suggests complete abrogation of *CIC* repressor protein activity in *CIC* mutant tumours.

CIC is the human homologue of capicua gene of *Drosophila* which has been extensively analyzed for its functional role in *Drosophila*. *CIC* is a HMG box-containing DNA-binding protein that is evolutionarily conserved across species [104], and functions as a transcriptional repressor by preferential binding to TGAATGA/GA sequences in *Drosophila* and mammals [105], [106]. *Cic*, Capicua meaning head-and-tail in Catalan, was identified in developmental studies of *Drosophila* [104]. *Drosophila Cic* plays an essential role downstream of the Torso and the epidermal growth factor receptors, two tyrosine kinases that transmit the signaling via the RAS-RAF-MAP kinase pathway mediating specification of interveing areas in the wing, correct development of the head and tail, and dorsoventral patterning [104], [107], [108]. Apart from

Cic's role in the cell fate determination downstream of RTK pathways in *Drosophila*, Cic is also known to play a role in regulating growth of imaginal discs downstream of the RTK/RAS/MAPK pathway. In human cells as well, CIC appears to play a role downstream of the RTK-MAPK signaling pathway [109]. In mammalian HEK293 and melanoma cells, MAPK signaling results in phosphorylation of CIC and subsequent loss of CIC-mediated transcriptional repression of PEA3 group genes [109].

***FUBP1* MUTATIONS**

FUBP1- the far upstream element (FUSE) binding protein 1- gene located on chromosome 1p31.1 is found to be mutated in 30% of *IDH1/IDH2* mutated chromosome 1p/19q codeleted oligodendrogliomas [47], [49], [50]. The protein encoded by *FUBP1* is a single-stranded DNA binding protein. It was identified as a binding partner for the far upstream element (FUSE) of *MYC* gene [110].

OTHER GENETIC ALTERATIONS

The Cancer Genome Atlas reported recurrent somatic alterations in glioblastoma after comprehensive genomic characterization of more than 500 tumors. Recurrent copy number alterations were found to affect genes involved in the p53 pathway. Amplifications have been found in *MDM2* (7.6%), *MDM4* (7.2%) genes that are known to be involved in degradation or inactivation of TP53 while deletion/mutations in *TP53* gene were found in about 28% of glioblastomas. Alterations are also found in cell cycle regulatory pRb pathway. Amplifications have been found in *CDK4* (14%), *CDK6* (1.6%), while deletions are found in *CDKN2A/B* (61%), and *RBI* (7.6%) genes of the pRB pathway. Further alterations in PI3K signaling pathway genes have also been found in glioblastomas. Amplifications of *PIK3CA*, *EGFR*, *PDGFRA*, gene were

found while deletions in *PTEN* and *NF1* gene were found in 41% and 10% of glioblastomas respectively [63]. At least one Receptor Tyrosine Kinase was found to be altered in 67.3% of glioblastoma overall: *EGFR* (57.4%), *PDGFRA* (13.1%), *MET* (1.6%), and *FGFR2/3* (3.2%). PI3K family gene mutations were found in 25.1% of glioblastoma (Figure 2.1). The TP53 pathway was found to be dysregulated in 85.3% of tumors. 78.9% of tumors had one or more alteration affecting Rb function [63].

2.7 GLIOMA DIAGNOSTIC CONUNDRUM AND NEED FOR SPECIFIC DIAGNOSTIC MARKERS

Diagnosis of gliomas primarily depends on histopathologic analysis of H & E (Hematoxylin and Eosin) stained slides of tumour tissue sections [7], [8]. Morphological evaluation of cancers by light microscopy has been the foundation for diagnosis, prognostication, and therapeutic stratification for well over a century. Histopathologic diagnosis is particularly challenging when it comes to reliably distinguishing between oligodendroglial and astrocytic component in non-classic low grade to intermediate grade gliomas. Several studies have shown considerable intra-observer variation in diagnosis of the glioma subtypes astrocytomas, oligodendrogliomas, and oligoastrocytomas. Accurate pathologic diagnosis needs the ability to distinguish astrocytic from oligodendroglial differentiation in histologic sections. This becomes a challenging feat even for the most experienced neuropathologist. Studies have demonstrated poor intraobserver reproducibility and an unacceptably wide variation in the diagnosis of astrocytoma, oligodendroglioma, and mixed oligoastrocytoma, even among the most experienced neuropathologists [11], [12]. High interobserver variability in the diagnosis of diffuse gliomas arises due to subjective diagnostic criteria, overlapping morphologic features, and variations in training and practice among pathologists [8], [9], [11], [12]. The study by Bruner of diagnostic discrepancies in a series of neuropathology referral cases demonstrated that there are frequent

clinically significant errors that could affect patient management and quality of life [111]. Histopathological classification of diffuse gliomas does not correlate consistently with molecular markers and histopathological classification does not perfectly predict clinical outcomes [9], [10]. Also there is distinct prognosis and treatment strategy for each glioma subtype and these are guided by diagnosis. For reliably distinguishing the main subtypes of gliomas, specific genetic and immunohistochemical markers are highly needed. Current FISH (1p/19q codeletion) and Immunohistochemistry (TP53) based markers assist in tumour classification but are not sufficient. For example, inactivating point mutations in *TP53* are more common in astrocytomas (50% to 60% in grade II) than in oligodendrogliomas (only 5% to 10%) this suggests that it might be a discriminating marker, but p53 immunoreactivity does not correlate perfectly with the presence of TP53 mutations (concordance of 65%-75%) [8]. 1p/19q codeletions occur more frequently in oligodendrogliomas (39-70%) than in astrocytomas (0-10%) but are also found frequently in oligoastrocytomas (21-59%) [3] thus precluding its use as specific diagnostic marker.

Due to prevalent diagnostic conundrum in low grade glioma specific biomarkers for glioma subtypes are of utmost importance for accurate diagnosis, prognostication and for deciding treatment strategies. Also the insights gained from the identification of specific molecular alterations affecting a particular gene or pathway may lead to development of targeted therapies.

2.8 EXPRESSION PROFILING OF GLIOMAS

Morphological evaluation of cancers by light microscopy has been the foundation for diagnosis, prognostication, and therapeutic stratification. However, patients with morphologically identical tumours can have significantly different clinical outcomes [34]. The current WHO classification does not comprehensively reflect diffuse glioma biology and patient outcome. There is extensive

evidence that tumors with indistinguishable morphology under the microscope from different patients do not necessarily share the same biology and do not necessarily reflect similar patient outcomes [29], [34], [112]–[114]. Expression profiling provides an objective method to classify tumors [115], [116]. Defining glioma subtypes based on objective genetic and molecular signatures may allow for a more rational, patient-specific approach to molecularly targeted therapy [114].

Many Gene expression profiling studies have confirmed that significant molecular heterogeneity exists within the various morphologically defined gliomas [34]. French et al. identified gene expression profiles associated with treatment response in oligodendrogliomas using expression profiling [117]. Phillips et al. identified three subtypes when they molecularly profiled several high-grade glioma samples. The subtypes were designated proneural, proliferative and mesenchymal to recognize the dominant feature of the signature genes that characterizes each subclass [118]. Gravendeel et al. identified seven molecular subtypes with distinct prognosis using 5000 genes with highly variable expression in 276 gliomas of all histological subtypes and grades [29]. Using a cohort of 225 glioma tumours of major glioma subtypes Yan et al. identified 3 subtypes with differences in clinical characteristics. The G1 subgroup was characterized by good clinical outcome, young age, low malignant behaviors, and extraordinarily high IDH1 mutation. G3 groups exhibited the opposite effect. The G2 subtype is the middle class of the aforementioned 2 subtypes [114]. The TCGA network described a gene expression based molecular classification of glioblastomas that divided them into proneural, neural, classical, and mesenchymal subtypes [119]. However these molecular subtypes displayed considerable overlap in expression profiles as well as in copy number variations and mutation spectra. Further gene expression profiling could not distinguish astrocytomas from oligodendrogliomas and often

glioblastomas as well. Both oligodendrogliomas and astrocytomas clustered with proneural glioblastomas. The overlapping gene expression profiles of the glioma subtypes is consistent with the overlapping morphological features suggesting possible overlap in cells of origin of the glioma subtypes.

2.9 RECENT CONCURRENT RESEARCH

During the study period (2011-2016) of current research project described in this thesis, many groups reported their findings from exome sequencing studies which overlapped to an extent with the findings in current study project. Exome sequencing of oligodendrogliomas by Bettgowda et al. identified inactivating mutations in two tumor suppressor genes *CIC* in 53% of oligodendrogliomas and *FUBP1* in 15% of oligodendrogliomas [47]. Other groups investigated *CIC* and *FUBP1* mutational status in grade II-III oligodendrogliomas, oligoastrocytomas and astrocytomas [49], [50]. In 2015 The Cancer Genome Atlas Research Network (TCGA) published a comprehensive analysis of 293 diffuse lower grade gliomas (WHO Grade II-III). Grade II-III gliomas were classified in three major classes 1) *IDH1/2* mutated without 1p/19q codeletion with high frequency of *TP53* and *ATRX* mutations 2) *IDH1/2* mutated with 1p/19q codeletion with high frequency of *CIC* mutations and 3) *IDH1/2* wild type gliomas with molecular features similar to *IDH1/2* wild type glioblastomas. These studies have been discussed in detail in relation with the finding of the present thesis work in the Discussion section.

2.10 WHO UPDATED GUIDELINES FOR BRAIN TUMOR CLASSIFICATION 2016

In 2016, the WHO released updated guidelines for brain tumor classification that combines biology-driven molecular marker diagnostics with classical histological cancer diagnosis [40]. The revised classification made *IDH1/IDH2* mutation and 1p/19q codeletion the defining features of oligodendroglioma and anaplastic oligodendroglioma. The WHO grade II diffuse

astrocytomas and WHO grade III anaplastic astrocytomas are now each divided into *IDH*-mutant, *IDH*-wildtype and NOS (i.e., not otherwise specified) categories. Glioblastomas are also classified according to *IDH* mutation status. Diagnosis of oligoastrocytoma is strongly discouraged. The diagnosis of WHO grade II oligoastrocytoma (NOS) and WHO grade III anaplastic oligoastrocytoma (NOS) can only be made in the absence of appropriate diagnostic molecular testing [40].

MATERIALS AND METHODS

3 MATERIALS AND METHODS

3.1 LABORATORY METHODS

3.1.1 Collection of Tumor Tissues and Paired Blood specimens

Ethics approval was obtained from Institutional Review Board (IRB) of ACTREC/TMC. The tumor tissues were obtained from the Neurosurgery department after obtaining informed consent from the patients. The tumor tissues collected in cryotubes were immediately snap frozen in liquid nitrogen. Frozen cryotubes were then transferred to -80⁰C freezer for long term storage. Paired blood sample (5 ml) was also collected before or after surgery.

3.1.2 Genomic DNA Isolation from Brain tumor tissues

Commercially available QIAamp® DNA Mini and Blood Mini kit from Qiagen, USA was used for extraction of total genomic DNA from the brain tumor tissues. Five micrometer cryosections of the brain tumor tissues were taken on a microscope slide. Subsequently, H & E stained 5 µm cryosections of the frozen tumor tissues were microscopically examined to ensure at least 80% tumor content (by Dr. E. Sridhar, Pathologist) before proceeding for DNA isolation.

Materials and Reagents

Qiagen QIAamp® DNA Mini and Blood Mini kit (Catalogue No. 51304) includes Buffer ATL, Proteinase K , RNase , Buffer AL, Buffer AW1, Buffer AW2 , QIAamp DNA Mini spin columns, collection tubes and Buffer AE.

Absolute Ethanol

Autoclaved 1.5 ml microcentrifuge tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Table-top microcentrifuge (Eppendorf)

Method

- All centrifugation steps were carried out at room temperature (22°C–25°C).
 - Samples and reagents were equilibrated to room temperature (22°C–25°C).
 - Heating block was heated to 56°C for use in the step 4.
1. The fresh frozen tissue sample was removed from -80°C storage. ~25 mg tumor tissue was cut out and used for DNA isolation.
 2. 15-20 micrometer cryosections of tumor tissue were taken and the sections were immediately dissolved in 180 µl buffer ATL in a 1.5 ml microcentrifuge tube. The tube was vortexed to disperse the sections completely, and then short spin was given to settle drops from the inside of the lid.
 3. 20 µl proteinase K was added to the tube and mixed completely by vortexing, and incubated at 56°C until the tissue sections were completely dissolved. The tube was vortexed occasionally during incubation to disperse the sample. Alternatively the tube was incubated in a rotator incubator set at 56°C in which the contents of the tube were continuously swirled at low speed during incubation until the complete dissolution of the sections.
 4. The 1.5 ml microcentrifuge tube was briefly centrifuged to remove drops from the inside of the lid.
 5. To obtain RNA-free genomic DNA, 4 µl of an RNase A stock solution (100 mg/ml) was added. The contents were mixed by vortexing briefly and subsequent spun and incubated at room temperature for 10 min.

6. 200 µl Buffer AL was added to the sample and mixed by pulse-vortexing for 15 sec, and incubated at 70°C for 10 min. The 1.5 ml microcentrifuge tube was briefly centrifuged to remove drops from inside the lid. It was ensured that the sample and Buffer AL were mixed thoroughly to yield a homogeneous solution. A white precipitate may form on addition of Buffer AL, which in most cases dissolved during incubation at 70°C.
7. 200 µl ethanol (96–100% ethanol) was added to the sample, and mixed again by pulse-vortexing for 15 s. After mixing, the 1.5 ml microcentrifuge tube was briefly centrifuged to remove droplets from the inside of the lid.
8. The mixture from step 7 was carefully applied to the QIAamp Mini spin column (placed in a 2 ml collection tube) without wetting the rim. The cap was closed, and centrifuged at 6000 x g (8000 rpm) for 1 min. The QIAamp Mini spin column was placed in a clean 2 ml collection tube, and the tube containing the filtrate was discarded.
9. The QIAamp Mini spin column was carefully opened and 500 µl Buffer AW1 was added without wetting the rim (the contents were triturated three times to resuspend all the sediment from column inner rim). The cap was closed and the tube was centrifuged at 6000 x g (8000 rpm) for 1 min. The QIAamp Mini spin column was placed in a clean 2 ml collection tube and the collection tube containing the filtrate was discarded.
10. The QIAamp Mini spin column was carefully opened and 500 µl Buffer AW2 was added without wetting the rim. The contents were triturated three times to resuspend all the sediment from column inner rim. The cap was closed and the tube was centrifuged at full speed (20,000 x g; 14,000 rpm) for 3 min.
11. The QIAamp Mini spin column was placed in a new 1.5 ml microcentrifuge tube (with cap cut off) and the old collection tube was discarded with the filtrate. The spin column in

microcentrifuge tube was centrifuged at full speed (20,000 x g; 14,000 rpm) for 1 min.

This step helps to eliminate the chance of possible Buffer AW2 carryover.

12. The QIAamp Mini spin column was placed in a clean 1.5 ml microcentrifuge tube (with cap cut off), and the collection tube containing the filtrate was discarded. The QIAamp Mini spin column was carefully opened and 100 µl Buffer AE was added. The column was incubated at room temperature (15–25°C) for 1 min, and then centrifuged at 6000 x g (8000 rpm) for 1 min. Incubating the QIAamp Mini spin column loaded with Buffer AE for 5 min at room temperature before centrifugation generally increases DNA yield. A second elution step with a further 200 µl Buffer AE increases yields by up to 15%.

13. The tube containing the isolated DNA was labeled and stored at -20°C

3.1.3 Genomic DNA Isolation from Blood

This method was used for purification of total genomic DNA from whole blood or blood clot using a microcentrifuge.

Materials and Reagents

Qiagen QIAamp® DNA Mini and Blood Mini kit (Catalogue No. 51304) included

Proteinase K, Buffer AL, RNase A, QIAamp DNA Mini spin column, Collection tubes, Buffer AW1, Buffer AW2, Buffer AE

Absolute Ethanol

Autoclaved 1.5 ml microcentrifuge tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Table-top microcentrifuge (Eppendorf)

Method

- All centrifugation steps were carried out at room temperature (22–25°C).

- Samples and reagents were equilibrated to room temperature (22–25°C).
- Heating block was heated to 56°C for use in step 4.

Procedure was started with [A] for whole blood or [B] for blood clot.

A] For EDTA Anticoagulated Blood (Collected in Purple Capped Vacutainer)

1. 20 µl of QIAGEN Protease (or proteinase K) was pipeted into the bottom of a 1.5 ml microcentrifuge tube.
2. 200 µl of blood sample was added to the microcentrifuge tube and mixed by pulse-vortexing for 15 s.
3. 200 µl Buffer AL was added to the sample and mixed by pulse-vortexing for 15 s. The sample and Buffer AL were mixed thoroughly to ensure efficient lysis to yield a homogeneous solution.
4. The tube was incubated at 56°C for 30 min or till the solution became clear and then proceeded to step 5

B] For Blood Clot (Yellow Capped Vacutainer)

1. 200 µl buffer AL was taken in microcentrifuge tube.
2. 15- 20 µm cryosections (of approx. 200 µl volume) of blood clot were taken and the sections were immediately dissolved in buffer AL in a 1.5 ml microcentrifuge tube. The tube was vortexed to disperse the sections completely, and then short spin was given.
3. 20 µl of proteinase K was added to the tube and mixed thoroughly by pulse-vortexing for 15 s.
4. The tube was incubated at 56°C for 2 hr or till the solution became clear. The contents were mixed by short vortex-spin every 30 min. Alternatively the tube was incubated in a rotator incubator in which the contents of the tube were continuously swirled at low speed during incubation. DNA yield reaches a maximum after lysis for 10 min at 56°C.

Longer incubation times have no effect on yield or quality of the purified DNA.

Proceeded to step 5

5. The 1.5 ml microcentrifuge tube was briefly spun to remove drops from the inside of the lid.
6. To obtain RNA-free genomic DNA 4 μ l of an RNase A stock solution (100 mg/ml) was added. The contents were mixed by short vortex-spin and incubated at RT for 10 min.
7. 200 μ l ethanol (96–100%) was added to the sample, and mixed again by pulse-vortexing for 15 s. After mixing, the 1.5 ml microcentrifuge tube was briefly centrifuged to remove drops from the inside of the lid.
8. The mixture from step 7 was carefully applied to the QIAamp Mini spin column (placed in a 2 ml collection tube) without wetting the rim. The cap was closed, and centrifuged at 6000 \times g (8000 rpm) for 1 min. The QIAamp Mini spin column was placed in a clean 2 ml collection tube, and the tube containing the filtrate was discarded.
9. The QIAamp Mini spin column was carefully opened and 500 μ l Buffer AW1 was added without wetting the rim (the contents were triturated three times to resuspend all the sediment from column inner rim). The cap was closed and the tube was centrifuged at 6000 \times g (8000 rpm) for 1 min. The QIAamp Mini spin column was placed in a clean 2 ml collection tube and the collection tube containing the filtrate was discarded.
10. The QIAamp Mini spin column was carefully opened and 500 μ l Buffer AW2 was added without wetting the rim. The contents were triturated three times to resuspend all the sediment from column inner rim. The cap was closed and the tube was centrifuged at full speed (20,000 \times g; 14,000 rpm) for 3 min.

11. The QIAamp Mini spin column was placed in a new 1.5 ml microcentrifuge tube (with cap cut off) and the old collection tube was discarded with the filtrate. The spin column in microcentrifuge tube was centrifuged at full speed for 1 min. This step helps to eliminate the chance of possible Buffer AW2 carryover.
12. The QIAamp Mini spin column was placed in a clean 1.5 ml microcentrifuge tube (with cap cut off), and the collection tube containing the filtrate was discarded. The QIAamp Mini spin column was carefully opened and 100 µl Buffer AE was added. The column was incubated at room temperature (15–25°C) for 1 min, and then centrifuged at 6000 x g (8000 rpm) for 1 min. Incubating the QIAamp Mini spin column loaded with Buffer AE for 5 min at room temperature before centrifugation generally increases DNA yield. A second elution step with a further 200 µl Buffer AE increases yields by up to 15%.
13. The tube containing the isolated DNA was labeled and stored in -20°C.

3.1.4 Total RNA Isolation from Brain Tumor Tissues

This method was used for purification of total RNA from brain tumor tissue by mechanical tissue disruption, followed by lysis in solution D and purification using Qiagen RNeasy mini kit. Five µm cryosections of the brain tumor tissue were taken on a microscope slide. H & E stained sections were microscopically examined to ensure at least 80% tumor content before proceeding for RNA isolation.

Materials and Reagents

Qiagen RNeasy Mini kit (Catalogue No. 74104):

RNeasy spin column

Buffer RW1

Buffer RPE

Absolute Ethanol

DEPC –trated MilliQ water-saturated phenol

Chloroform

RNase-free water

Autoclaved 1.5 ml microcentrifuge tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Table-top microcentrifuge (Eppendorf)

Mechanical tissue homogenizer

DEPC-treated Milli-Q water: DEPC treated Milli-Q water was used for preparation of the reagents required for this method. Water was collected in 50 ml autoclavable tubes. 50 µl DEPC was added to 50 ml Milli-Q water. The DEPC was mixed completely by vigorously shaking the tube. The tubes were left over night at 37°C. The tubes were autoclaved at 121°C for 15 minutes at 15 lb/psi pressure.

4 M GITC (Guanidinium Isothiocyanate): (Prepared in 25 mM Sodium citrate pH 7.0, 0.5 % Sarcosyl and 0.1 M β-mercaptoethanol). 23.6 g of guanidine isothiocyanate was dissolved in 40 ml DEPC-treated water. 1.25 ml of 1 M sodium citrate and 2.5 ml of 10 % sarcosine were added and the final volume was made up to 50 ml with DEPC-treated water. The final solution was neither treated with DEPC nor autoclaved.

10 % N-lauryl-sarcosine: 5 g N-lauryl-sarcosine was dissolved in DEPC-treated water and the final volume was made up to 50 ml. The resulting solution was neither treated with DEPC, nor autoclaved. It was kept at 65 °C for 1 h, and stored at room temperature.

1 M Sodium citrate, pH 7.0: 14.7 g Sodium citrate dihydrate was dissolved in about 35 ml of autoclaved Milli-Q water. The pH was adjusted to 7.0 with a few drops of 1 M citric acid and the volume was made up to 50 ml. (1 M Citric acid was prepared by dissolving 10.5 g powder in 50 ml DEPC-treated water.) 50 µl of DEPC was added to both 1 M citrate and citric acid solution, tubes were mixed vigorously and left at 37 °C overnight. The solutions were autoclaved on the next day, and stored at room temperature.

2 M Sodium acetate, pH 4.0: 13.6 g sodium acetate was dissolved in about 25 ml of Milli-Q water and pH was adjusted to 4.0 with glacial acetic acid. Final volume was made up to 50 ml with Milli-Q water. 50 µl DEPC was added to the solution, mixed vigorously and left at 37°C over night. The solution was autoclaved the following day and stored at room temperature.

Phenol (Saturated with DEPC-treated water): 25 ml DEPC-treated water was added to 25 ml distilled phenol at room temperature in a sterile 50 ml autoclavable tube. The tube was mixed vigorously by inverting several times. The tube was kept at 4 °C until the two phases separated (30-60 min). The upper phase of water was replaced with fresh DEPC-treated water, mixed once again and stored at 4 °C.

Solution D: Solution D was prepared from GITC by adding β -mercaptoethanol at a final concentration of 0.1 M. This solution is stable at room temperature for one month.

Method

1. Three ml of chilled solution D was taken in ice-chilled homogenizer tube. ~50 mg of weighed tumor tissue was placed in solution D in the homogenizer tube and the tissue was completely homogenized at medium intensity. With a micropipette the lysate was transferred to 1.5 ml microcentrifuge tube. Immediately the lysate was passed through a 26 gauge syringe needle. The lysate was passed through syringe needle 10 times or until the lysate lost its viscosity.
2. 50 μ l of 2 M Sodium acetate pH 4.0 was added per 0.5 ml solution D to the microcentrifuge tube and mixed by inverting the tube.
3. 0.5 ml of DEPC-treated MilliQ water-saturated phenol and 0.2 ml chloroform were added successively, and the contents of the tube were mixed thoroughly by vortexing for 1 min. The cap of the tube was loosened to release the developed pressure, and the tube was vortexed again for 30 seconds.
4. The tube was kept on ice for 15-20 min, and then centrifuged at 10,000 rpm at 4° C for 10 min in a table top centrifuge. The upper aqueous phase was transferred to a fresh 1.5 ml microcentrifuge tube, and centrifuged once again to settle any traces of phenol. This aqueous phase was used for total RNA isolation by RNeasy mini kit.
5. Equal volume of 70% ethanol was added to the aqueous phase, and mixed immediately by pipetting. Mix was not centrifuged. Proceeded to next step immediately.

6. Up to 700 μ l of the sample, including any precipitate that may have formed, was transferred to an RNeasy spin column placed in a 2 ml collection tube. The lid was closed gently, and centrifuged for 15 s at $\geq 8000 \times g$ ($\geq 10,000$ rpm). The flow-through was discarded. The collection tube was reused in the next step. If the sample volume exceeded 700 μ l, successive aliquots were centrifuged in the same RNeasy spin column. The flow-through was discarded after each centrifugation.
7. 700 μ l Buffer RW1 was added to the RNeasy spin column. The lid was closed gently, and centrifuged for 15 s at $\geq 8000 \times g$ ($\geq 10,000$ rpm) to wash the spin column membrane. The flow-through was discarded. The collection tube was reused in the next step.
8. 500 μ l Buffer RPE was added to the RNeasy spin column. The lid was closed gently, and centrifuged for 15 s at $\geq 8000 \times g$ ($\geq 10,000$ rpm) to wash the spin column membrane. The flow-through was discarded. The collection tube was reused in the next step.
9. 500 μ l Buffer RPE was added to the RNeasy spin column. The lid was closed gently, and centrifuged for 2 min at $\geq 8000 \times g$ ($\geq 10,000$ rpm) to wash the spin column membrane. The long centrifugation dries the spin column membrane, ensuring that no ethanol is carried over during RNA elution. Residual ethanol may interfere with downstream reactions.
10. To eliminate any possible carryover of Buffer RPE, or any residual flow-through the RNeasy spin column was placed in a new 2 ml collection tube, and the old collection tube with the flow-through was discarded. The lid was closed gently, and centrifuged at full speed for 1 min.

11. The RNeasy spin column was placed in a new 1.5 ml collection tube. 55 µl of RNase-free water was added directly to the spin column membrane. The lid was closed gently, and centrifuged for 1 min at $\geq 8000 \times g$ ($\geq 10,000$ rpm) to elute the RNA.
12. The eluted total RNA was collected in autoclaved 1.5 ml microcentrifuge tube, labeled and stored at -80°C .

3.1.5 DNA Agarose Gel Electrophoresis

This method was used for detecting the quality / integrity of the isolated genomic DNA or visualization of PCR products by electrophoretic separation of DNA on the agarose gel along with standard DNA markers.

Materials and Reagents

Agarose

Glass beaker

Microwave oven

Autoclaved 0.2 ml tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Electrophoresis unit with power pack

50 X Tris-acetate-EDTA (TAE) buffer: 121 g Tris and 18.6 g EDTA was dissolved in 300 ml of Milli-Q water followed by addition of 28.55 ml glacial acetic acid. Volume was made up to 500 ml and was autoclaved.

Ethidium Bromide stock solution (10 mg / ml): Dissolve 10 mg Ethidium Bromide in 1 ml of autoclaved Milli-Q water.

6 X DNA loading dye: Dissolve 0.25 % bromophenol blue, 40 % (w / v) glycerol in Milli-Q water.

Method

1. 50X TAE was diluted to 1X by dilution with Milli-Q water. [1 ml 50X TAE + 49 ml Milli-Q water].
2. The gel casting tray was cleaned and sealed with tape on the both open sides. Gel comb was placed in the tray.
3. According to the required gel concentration Agarose was weighed and dissolved in the 1X TAE in a glass beaker. [For 2% Agarose gel 0.5 gm of Agarose was dissolved in 25 ml of 1X TAE].
4. The glass beaker was heated to boiling point in a microwave oven with intermittent shaking to completely dissolve the Agarose in TAE buffer to form transparent solution.
5. The gel was allowed to cool down to ~50⁰C and Ethidium bromide was added to the gel at a final concentration of 0.5 µg / ml. The Ethidium bromide was mixed completely in the molten Agarose gel.
6. The molten Agarose gel was poured in the gel casting tray carefully without allowing any bubbles to be formed. The gel was allowed to cool undisturbed for about half hour.
7. After the gel is set the comb and tape was removed carefully.
8. The Agarose gel with the tray was placed in the electrophoresis unit and the electrophoresis unit was filled with 1X TAE till it covers the Agarose gel slab completely.
9. In a 0.2 ml tube the DNA sample or DNA marker ladder was added to the 6X gel loading dye so that the final gel dye concentration is above 1X. [2 µl gel loading dye + 8 µl DNA

sample]. The DNA sample and gel loading dye was mixed thoroughly with micropipette and carefully transferred to the bottom of the well on the Agarose gel slab.

10. The electrophoresis unit was run on 50 Volts until the gel dye migrates 3/4th of the gel.

11. After the electrophoresis was complete, the gel was visualized under UV transilluminator and an image was saved.

3.1.6 DNA Polyacrylamide Gel Electrophoresis [DNA PAGE]

This method was used for higher resolution separation of PCR products for the analysis of microsatellite length polymorphism in genomic DNA.

Materials and Reagents:

Autoclaved 0.2 ml tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Glass beaker

PAGE Electrophoresis unit with power pack

Horizontal Orbital Shaker

TEMED: Tetramethylethylenediamine

20% Ammonium persulfate (APS): 200 mg APS was dissolved in 1 ml autoclaved Milli-Q water (always freshly prepared)

50 X Tris-acetate-EDTA (TAE) buffer: 121 g Tris and 18.6 g EDTA was dissolved in 300 ml of Milli-Q water followed by addition of 28.55 ml glacial acetic acid. Volume was made up to 500 ml and was autoclaved.

Ethidium Bromide stock solution (10 mg / ml): Dissolve 10 mg Ethidium Bromide in 1 ml of autoclaved Milli-Q water.

6 X DNA loading dye: Dissolve 0.25 % bromophenol blue, 40 % (w / v) glycerol in Milli-Q water.

30 % (acrylamide + bis-acrylamide) solution: 29.2 g acrylamide, 0.8 g bis-acrylamide were dissolved in approximately 50-60 ml autoclaved Milli-Q water and the final volume was made up to 100 ml. Solution was filtered through ordinary filter paper and stored in an amber colored bottle at 4 °C.

Method

PAGE gel preparation and sample loading

1. The PAGE unit was cleaned and assembled. The unit was sealed on three sides except the upper side with molten 2% agar gel using pipette or syringe.
2. Polyacrylamide gel of 10% concentration was prepared according to the table given below.

Reagent	10% gel
30% Polyacrylamide + bisacrylamide	3.34 ml
50X TAE	0.2 ml
MilliQ water	6.46 ml
TEMED	7 µl
20% APS	100 µl
Total Volume	10 ml

3. TEMED and APS were added just before casting the gel. As soon as these polymerization agents were added, the mixture was swirled thoroughly and then the solution was poured carefully in between the glass plates to completely fill the space between the glass plates without trapping any bubbles. Immediately gel comb was inserted at the top of the glass plates carefully avoiding any bubbles being formed.
4. The solution was allowed to polymerize for half hour. After polymerization the comb was carefully removed avoiding any damage to the wells.
5. The electrophoresis unit was filled completely with 1X TAE buffer.
6. The DNA sample was diluted with 6X gel loading dye so that the final gel dye concentration was 1X in a 0.2 ml tube and mixed thoroughly.
7. The DNA samples were carefully added at the bottom of the gel well.
8. The electrophoresis was carried out at 50 volts till the gel loading dye traveled 3/4th of the gel.

Staining, washing and visualization of PAGE gel

1. 2 µl Ethidium Bromide stock solution was added and dissolved completely in 50 ml 1X TAE buffer in a wide and flat bottomed container.
2. After the electrophoresis was complete, the gel was removed from between the glass plates taking extreme care not to tear the gel. The gel was gently placed in the container with 1X TAE buffer containing Ethidium bromide.
3. The container was placed on a horizontal shaker at low speed for 10 minutes.

4. After 10 minutes the TAE-Ethidium Bromide solution was discarded and replaced with 50 ml of 1X TAE for a wash. The container was placed on horizontal shaker at low speed for 15 minutes. This wash step was performed again to give a total of two washes.
5. After the final wash the gel was visualized on UV transilluminator and an image was saved.

3.1.7 DNA Quantification Using Qubit Fluorometer

This method is for quantification of DNA by using Qubit® 2.0 fluorometer and Qubit® dsDNA BR Assay Kit (catalogue No. Q32850, for 0.01 µg/ml to 5 µg/ml DNA concentration range) or Qubit® dsDNA HS Assay Kit (catalogue No. Q32851, for 1 ng/ml to 500 ng/ml DNA concentration range)

Materials and Reagents

Qubit® 2.0 fluorometer

Qubit® dsDNA BR Assay Kit (catalogue No. Q32850) or

Qubit® dsDNA HS Assay Kit (catalogue No. Q32851)

Qubit® dsDNA BR Reagent

Qubit® dsDNA BR Buffer

Standard #1

Standard #2

1.5 ml or 2 ml microcentrifug tubes

0.5 ml Axygen® PCR tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Method

1. The required number of 0.5 ml tubes to be used for standards and samples were set up. The Qubit® dsDNA BR/HS Assay requires 2 standards (Standard #1, Standard #2 and sample assay tubes).
2. The tube lids were labelled. The tubes were handled wearing plastic gloves.
3. The Qubit® working solution was prepared by diluting the Qubit® dsDNA BR Reagent 1:200 in Qubit® dsDNA BR Buffer. A clean plastic tube was used each time Qubit® working solution was prepared. Working solution is never prepared in a glass container.
4. The final volume in each tube must be 200 µl. Each standard tube requires 190 µl of Qubit® working solution, and each sample tube requires anywhere from 180–199 µl. Sufficient Qubit® working solution was prepared to accommodate all standards and samples.
5. 190 µl of Qubit® working solution was added to each of the tubes used for standards.
6. 10 µl of each Qubit® standard was added to the appropriate labeled tube, and then was mixed by vortexing for 2–3 seconds. Care was taken not to create bubbles.
7. Qubit® working solution was added to individual assay tubes so that the final volume in each tube after adding sample was 200 µl. For example for 1 µl of DNA sample and 199 µl of working solution was added to individual assay tube.
8. All tubes were allowed to incubate at room temperature for 2 minutes.
9. On the Home screen of the Qubit® 2.0 Fluorometer, DNA tab was pressed, then from dsDNA Broad Range (BR) and dsDNA High Sensitivity (HS), appropriate assay was selected as the assay type. The “Read standards” screen is displayed. “Read Standards” was pressed to proceed.

10. The tube containing Standard #1 was inserted into the sample chamber, the chamber lid was closed, and then “Read standard” was pressed. When the reading was complete (~3 seconds), Standard #1 tube was removed.
11. The tube containing Standard #2 was inserted into the sample chamber, the chamber lid was closed, and then “Read standard” was pressed. When the reading was complete (~3 seconds), Standard #2 tube was removed.
12. Run samples menu was pressed.
13. On the assay screen, the sample volume and units were selected:
 - a. The scroll wheel was pressed to select the appropriate sample volume added to the assay tube (from 1–20 µl).
 - b. From the dropdown menu, the units for the output sample concentration were selected.
14. A sample tube was inserted into the sample chamber, the chamber lid was closed, and then “Read tube” button was pressed. When the reading was complete (~3 seconds), the sample tube was removed. This step was repeated until all samples have been read and the displayed concentration for each DNA sample was noted down.

3.1.8 Polymerase Chain Reaction

This method was used for *in vitro* amplification of DNA fragments to be used further for Sanger sequencing or Microsatellite Length Polymorphism based loss of heterozygosity analysis.

Materials and Reagents

10X Standard Taq Reaction Buffer (New England Biolabs (NEB), USA)

10 mM dNTPs (NEB, USA)

10 µM Forward Primer

10 µM Reverse Primer

Taq DNA Polymerase (NEB, USA)

Template DNA

Nuclease-Free Water / Autoclaved Milli-Q water

Autoclaved 0.2 ml and 0.5 ml PCR tubes

Micropipettes (20 μ l, 200 μ l, 1000 μ l) and corresponding autoclaved tips

PCR Thermal Cycler (Veriti, Applied Biosystems, USA)

Method

1. The PCR cycling parameters were standardized on Veriti thermal cycler (Applied Biosystems, US). All reagents were thawed and kept on ice.
2. For each PCR reaction, 10 ng of the template DNA was added in the end to the 0.2 ml PCR tube containing the PCR components.
3. The PCR Reaction mix was prepared as follows:

COMPONENT	10 μ l REACTION	50 μ l REACTION	FINAL CONCENTRATION
10X Standard Taq Reaction Buffer	1 μ l	5 μ l	1X
10 mM dNTPs	0.2 μ l	1 μ l	200 μ M
10 μ M Forward Primer	0.2 μ l	1 μ l	0.2 μ M (0.05–1 μ M)
10 μ M Reverse Primer	0.2 μ l	1 μ l	0.2 μ M (0.05–1 μ M)
Taq DNA Polymerase	0.05 μ l	0.25 μ l	1.25 units/50 μ l PCR
Template DNA	10 ng	50 ng	<1,000 ng
Nuclease-Free Water	Make volume to 10 μ l	Make volume to 50 μ l	

4. All precautions were taken to avoid PCR related contamination. All reagents and PCR products were handled using autoclaved tips or filter tips.

5. The PCR cycling parameters were as follows:

STEP	TEMPERATURE	TIME
Initial Denaturation	95°C	30 seconds
30 Cycles	95°C	15–30 seconds
	45–68°C	15–60 seconds
	68°C	1 minute/kb
Final Extension	68°C	5 minutes
Hold	4–10°C	

6. 10 µl of the PCR product was run on a 1 % agarose gel and visualized using an UV Transilluminator.

7. PCR products were further treated with Exonuclease I and Shrimp Alkaline Phosphatase to remove unused primers, dNTPs and used for Sanger sequencing.

3.1.9 Real Time PCR for Library Quantification

This method was used for quantification of genomic DNA libraries, Exome libraries and transcriptome libraries using SYBR green.

Materials and Reagents

2 X SYBR green Master Mix (Applied Biosystems),

Forward primer (10 pmol / µl or 1 pmol / µl)

Reverse primer (10 pmol / µl or 1 pmol / µl) (Primer sequence provided in Appendix I)

Standard Libraries for quantification (Kapa Biosciences, Catalogue No. KK4903)

Tween 20

384-well optical real time PCR plate

Autoclaved 1.5 ml and 2 ml microcentrifuge tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

2 M Tris-Cl pH 8.0: 121.14 g Tris was dissolved in 400 ml Milli-Q water. pH was adjusted to 8.0 with concentrated HCl, the final volume was made up to 500 ml with Milli-Q water and autoclaved. Solutions were stored at 4 °C.

10 mM Tris-Cl (pH 8.0): 0.25 ml of 2 M Tris-Cl pH 8.0 was diluted to 50 ml to prepare 10 mM Tris-Cl (pH 8.0).

Method

1. The DNA libraries were diluted with 10 mM Tris HCl (containing 0.05% Tween 20) to 10^{-2} , 10^{-3} , 10^{-5} and 10^{-7} dilutions according to following table.

199 μ l (10 mM Tris HCl + 0.05% Tween 20) + 1 μ l Library sample	1:2* 10^2 dilution
90 μ l (10 mM Tris HCl + 0.05% Tween 20) + 10 μ l of 1:2* 10^2 dilution	1:2* 10^3 dilution
198 μ l (10 mM Tris HCl + 0.05% Tween 20) + 2 μ l of 1:2* 10^3 dilution	1:2* 10^5 dilution
198 μ l (10 mM Tris HCl + 0.05% Tween 20) + 2 μ l of 1:2* 10^5 dilution	1:2* 10^7 dilution

2. Standard DNA Libraries (six standard libraries 10-fold dilutions 20 pM, 2 pM, 0.2 pM, 0.02 pM, 0.002 pM, 0.0002 pM) were also used as quantification standard as required.
3. The reagents were added in 384 well Real time PCR plate according to following table.

Reagent	Volume
2 X SYBR green PCR Master Mix	2.5 µl
1 pmol / µl Forward primer	0.25 µl
1 pmol / µl Reverse primer	0.25 µl
Diluted library or Library standards	2.0 µl
Total Volume	5 µl

4. The 5 µl PCR reaction mix in duplicates or triplicates per sample was carefully added in the required number of wells in 384-well optical real time PCR plate. An optical cover sheet was used to cover and seal the PCR plate with the help of a plastic applicator. Proper sealing of the wells was ensured to prevent volume loss due to evaporation. To spin down the reaction mixes to the bottom of the well the sealed plate was centrifuged briefly at 2000 rpm for 2 min. If any air bubbles present, they were removed by gently tapping the plate and the plate was centrifuged briefly at 2000 rpm for 2 min.
5. Real time PCR was carried out in QuantStudio™ 12K Flex Real-Time PCR System (Applied Biosystems, NY, USA) with default cycling parameters.
6. Real time PCR Cycling parameter

Temperature	Time	Cycles
50 °C	2 min	1
95 °C	10 min	1
95 °C	15 sec	40
60 °C	1 min	

7. After completion of the Real time PCR steps the data was analyzed and exported for further analysis using the Quant studio 12Kflex software (Applied Biosystems, NY, USA) on the Real time PCR system.

3.1.10 Genomic DNA Fragmentation

This method describes the fragmentation of genomic DNA for 300 bp fragment size using the Covaris Adaptive Focused Acoustics™ (AFA) process on Covaris M220 focused ultrasonicator.

Materials and Reagents

Covaris M220 focused ultrasonicator

AFA snap cap microtube of 130 µl capacity.

AFA-grade water, Ice, DNA Samples

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Method

Set Up of the Instrument for Sonication

1. The instrument was connected to the laptop with USB port.
2. The acoustic assembly and tube holder was cleaned with kimwipe. The tube holder was placed inside acoustic assembly.
3. AFA-grade water was added into acoustic assembly till water was visible at the top. The instrument and laptop were switched on by turning on the power switch.
4. Sonolab software was initialized on windows operating system. The software automatically connects to the acoustic sonicator. Instrument was left idle for ~30 min till all automated instrument checks on 'Instrument Status' turned green.

Fragmentation of Genomic DNA

1. In ‘Method’, the method to run was selected and required parameters were checked. The standardized parameters for 2.5 µg genomic DNA in 100 µl Qiagen AE buffer are as follows.

Time	125 sec
Peak Power	50
Duty Factor	20%
Cycles per burst	200
Sample volume	2.5 µg in 100 µl
Temperature	6 °C

2. 100 µl of 2.5 µg quality checked genomic DNA was carefully added inside chilled AFA snap cap microtube of 130 µl capacity placed in ice. The tube was kept on ice for two minutes.
3. The safety cover was opened, the microtube to be sonicated was loaded into the centre of the tube holder and cover was closed. When the “Run” button turned green the selected method was started.
4. After completion, a dialogue box is posted onscreen notifying the completion of sonication.
5. Steps 6 to 8 were repeated for each sample.

Steps to turn off the sonicator

1. After completion of sonication of all samples, the tube holder was placed on a clean dry surface. The water from the acoustic assembly was carefully drawn out completely using provided plastic syringe.

2. The surface of the transducer and the metal housing was wiped with lint-free cloth and air dried for next use.
3. SonoLab software by was closed. Laptop was shut down USB cord was disconnected and the instrument was switched off.

3.1.11 Genomic DNA Libray Preparation

This procedure describes genomic DNA library preparation for Illumina platform using KAPA Genomic DNA library preparation reagents.

Materials and Reagents

KAPA library preparation kit (Kapa Biosciences, Catalogue No.:K8200) contains the reagents listed below:

10X End Repair Buffer, End Repair Enzyme Mix, 10X A-tailing Buffer, 5X Ligation Buffer, DNA ligase, 2X KAPA HiFi Reaction Mixture, PCR Primer Cocktail, A-tailing Enzyme, 2X KAPA HiFi HS RM, PCR primer cocktail, Resuspension Buffer

Covaris fragmented genomic DNA

Illumina DNA Barcode Adaptors (30 μ M)

AMPure XP Beads

Magnetic stand

Low Melting Point (LMP) agarose

10X SyBr Gold gel stainig dye

5X gel loading dye

50X TAE buffer

80% Ethanol

Qiaquick gel extraction kit contains the reagents listed below:

Buffer QG, Isopropanol, QIAquick spin column, Bufffer EB, Buffer PE

Autoclaved 0.2 ml and 0.5 ml PCR tubes

Autoclaved 1.5 ml and 2 ml microcentrifuge tubes

Micropipettes (20 µl, 200 µl, 1000 µl) and corresponding autoclaved tips

Table-top microcentrifuge (Eppendorf)

Refrigerated vacuum concentrator

PCR thermal Cycler

Method

A] End Repair of Fragmented DNA

1. The end repair reaction was assembled as following in 0.5 ml PCR tube.

10X End repair buffer	10 µl
End repair enzyme mix	5 µl
2 µg sheared dsDNA	85 µl
<hr/>	
Total	100 µl

2. The components were mixed thoroughly, the cap closed and the tube was short spinned.
Incubate for 30 min at 20 °C in a PCR thermal cycler.
3. After incubation, the reaction was processed immediately to cleanup.
4. It was ensured that the AMPure XP Beads were equilibrated to room temperature, and that they were thoroughly resuspended.
5. AMPure Beads were added to the End Repair reaction:

End repair reaction	100 μ l
AMPure XP Beads	160 μ l
<hr/>	
Total	260 μ l

6. The reaction was mixed thoroughly by pipetting up and down at least ten times.
7. The tube was incubated at room temperature for 15 minutes to allow DNA to bind to the beads.
8. The beads were captured by placing the tube on an appropriate magnetic stand at room temperature for 15 minutes or until the liquid was completely clear.
9. Carefully 255 μ l of the liquid was removed and discarded. Care was taken not to disturb or discard any of the beads. Some liquid may remain visible in the tube.
10. Keeping the tube on the magnetic stand and without disturbing the beads, the beads were washed in 200 μ l of 80% Ethanol for at least 30 seconds.
11. Carefully the ethanol was removed and discarded without disturbing the beads, and the process was repeated for a total of 2 washes in 80% Ethanol.
12. The tube was removed from the magnetic stand, and the beads were allowed to dry at room temperature for 15 minutes.
13. The beads were resuspended thoroughly in 32.5 μ l elution buffer, and incubated at room temperature for 2 minutes to release the DNA from the beads.
14. The beads were captured by placing the tube on a magnetic stand at room temperature for 15 minutes or until the liquid was completely clear.
15. The DNA was recovered in 30 μ l of supernatant and transferred to the tube in which the A-tailing reaction was intended to be performed.

Safe Stopping Point: If A-Tailing is not to be proceeded immediately, the protocol can be safely

stopped here. The end repaired DNA fragments can be stored at -20 °C for up to seven days.

B] A-Tailing of End Repaired DNA Fragments

1. The A-Tailing reaction was assembled in a 0.5 ml PCR tube as given below:

Water	12 µl
10X A-tailing Buffer	5 µl
A-Tailing Enzyme	3 µl
End repaired DNA	30 µl
<hr/>	
Total	50 µl

2. The contents were mixed thoroughly and given a short spin. The tube was incubated in a PCR machine for 30 min at 30 °C
3. Immediately after the incubation, cleanup was initiated.
4. It was ensured that the AMPure XP Beads were equilibrated to room temperature, and that they were thoroughly resuspended.
5. AMPure XP Beads were added to the End Repair reaction as follows:

A-tailing reaction	50 µl
AMPure XP Beads	90 µl
<hr/>	
Total	140 µl

6. Beads were mixed thoroughly by pipetting up and down at least ten times.
7. The tube was incubated at room temperature for 15 minutes to allow DNA to bind to the beads.
8. The beads were captured by placing the tube on an appropriate magnetic stand at room temperature for 15 minutes or until the liquid was completely clear.
9. Carefully 135 µl of the liquid was removed and discarded. Care was taken not to disturb or discard any of the beads. Some liquid may remain visible in the tube.

10. Keeping the tube on the magnetic stand and without disturbing the beads, the beads were washed in 200 μ l of 80% Ethanol for at least 30 seconds.
11. Carefully the ethanol removed and discarded without disturbing the beads, and the process was repeated for a total of 2 washes in 80% Ethanol.
12. The tube was removed from the magnetic stand, and the beads were allowed to dry at room temperature for 15 minutes.
13. The beads were resuspended thoroughly in 32.5 μ l elution buffer, and incubated at room temperature for 2 minutes to release the DNA from the beads.
14. The beads were captured by placing the tube on a magnetic stand at room temperature for 15 minutes or until the liquid is completely clear.
15. The DNA was recovered in 30 μ l of supernatant and transferred to the tube in which the adaptor ligation reaction was intended to be performed.

Safe Stopping Point: If Adaptor Ligation is not to be proceeded immediately, the protocol can be safely stopped here. The A-tailed DNA fragments can be stored at -20 °C for up to seven days.

C] Adaptor Ligation

1. the Adaptor Ligation reaction was assembled in 0.5 ml PCR tube as follows:

5X Ligation Buffer	10 μ l
DNA Ligase	5 μ l
DNA adaptor (30 M)	5 μ l
A-Tailed DNA	30 μ l
<hr/>	
Total	50 μ l

2. The adapter used for each sample was noted down.
3. The reaction was incubated for 15 min at 20 °C in a PCR machine.
4. After incubation the reaction mixture was immediately processed for agarose gel separation and gel extraction.

D] Preparation of 2% LMP Agarose for Library Gel Size Selection

1. The agarose gel electrophoresis unit was thoroughly cleaned and rinsed with Milli-Q water.
2. Three wells of a gel comb were joined with a cello tape to form one well to accommodate 50 µl of sample + 10µl gel loading dye for each sample.
3. Two gram of LMP agarose powder was added in 100 ml of 1X TAE buffer, was heated in a microwave oven until the agarose powder completely dissolved.
4. The melted agarose was cooled on the bench for 5 minutes. Then, 10 µl of SyBr Gold was added, mixed by swirling and then poured into the gel tray.
5. The gel tray was placed in a refrigerator to cool for 30 min.
6. When the agarose gel was set, it was put in the gel electrophoresis unit and the tank was filled with 1X TAE buffer.
7. 5X gel loading dye was added to each adapter ligation reaction, mixed thoroughly and loaded carefully in the combined wells.
8. 9 µl TE buffer, 3 µl 100 bp DNA ladder , 3 µl 5X gel loading dye was mixed thoroughly and loaded into one well.
9. The electrophoresis was carried out at 55 volts until the tracking dye reached upto 3/4th of the gel length.

10. After the run, the gel was visualized on UV transilluminator. With a fresh clean scalpel blade, gel piece containing the sample DNA smear corresponding to size between 400-500 bp was excised out.
11. The gel piece was collected in a pre-weighed 2 ml autoclaved microfuge tube. The weight of gel slab was noted down by subtracting empty tube weight from the weight of the gel collected tube. DNA extraction from the gel piece was carried out using Qiaquick gel extraction immediately.

E] QIAquick Gel Extraction

1. It was ensured that the buffer QG had yellow color which indicates a pH ≤ 7.5 .
2. It was ensured that ethanol had been added to buffer PE before use.
3. Three volumes of Buffer QG was added to 1 volume of gel (100 mg \sim 100 μ l). The maximum amount of gel slice per Qiaquick column is 400 mg; for gel slices weighing more than 400 mg more than one QIAquick columns should be used.
4. The contents were incubated at room temperature for 10 min or until the gel slice had completely dissolved. To help dissolve gel, the tube was mixed by vortexing the tube every 2-3 min during the incubation.
5. After the gel had dissolved completely it was checked that the color of the mixture was yellow (similar to buffer QG without dissolved agarose). If the color of the mixture is orange or violet, 10 μ l of 3M sodium acetate, pH 5.0, should be added and mixed. The color of the mixture will turn yellow. The adsorption of DNA to QIAquick membrane is efficient only at pH ≤ 7.5 .
6. One gel volume of isopropanol was added to the sample and mixed. This step increases the yield of DNA fragments.

7. A QIAquick spin column was placed in a provided 2 ml collection tube.
8. To bind DNA, the sample was applied to the QIAquick column and centrifuged at 13,000 rpm for 1 min. The maximum volume of the column reservoir is 800 μ l. For sample volumes of more than 800 μ l the sample was loaded in multiple aliquots with intermediate centrifugation.
9. The flow through was discarded and QIAquick column was placed back in the same collection tube.
10. 0.5 ml of buffer QG was added to QIAquick column and centrifuged at 13,000 rpm for 1 min. This step ensures removal of all traces of agarose.
11. To wash, 0.75 ml of buffer PE was added to QIAquick column, the column was left to stand 2-5 min then the column was centrifuged at 13,000 rpm for 1 min.
12. The flow through was discarded and the QIAquick column was centrifuged for additional 1 min at 13,000 rpm.
13. To elute DNA, 50 μ l of Buffer EB (10 mM Tris-HCl, pH 8.5) was added to the center of the QIAquick membrane, let the column stand for 1 min, and then centrifuged for 1 min at 13,000 rpm.
14. Using refrigerated vacuum concentrator the \sim 50 μ l library elute was concentrated to 20 μ l. (The volume was readjusted to 20 μ l if concentrated to $<$ 20 μ l, by adding buffer EB)

F] Library Amplification

1. The following reaction conditions were set in a PCR machine. The lid was preheated at 100 $^{\circ}$ C

Denaturation	45 sec at 98 °C
10 cycles	15 sec at 98 °C
	30 sec at 60 °C
	30 sec at 72 °C
Final Extension	1 min at 72 °C
Hold	10 °C

2. Prepare the following reaction for each sample.

2X KAPA HiFi HS RM	25 µl
PCR primer cocktail	5 µl
Library DNA	20 µl

Total 50 µl

- The components were mixed thoroughly by pietting up and down 10 times.
- After completion of PCR cycles, the tube was removed from thermal cyclers.
- AMPure XP Beads were vortexed until they were well dispersed.
- 50 µl of the mixed AMPure XP Beads were added to each well of the PCR plate containing 50 µl of the PCR amplified library. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
- The tube was incubated at room temperature for 15 minutes.
- The PCR tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
- 95 µl of the supernatant was removed and discarded while keeping the tube on magnetic stand.
- With the PCR tube remaining on the magnetic stand, 200 µl of freshly prepared 80% ethanol was added to each well without disturbing the beads.

11. The PCR plate as incubated at room temperature for 30 seconds, then all of the supernatant was removed and discarded from each well.
12. Steps 11 and 12 were repeated once again for a total of two 80% Ethanol washes.
13. While keeping the PCR plate on the magnetic stand, the samples were air dried at room temperature for 15 minutes and then the plate was removed from the magnetic stand.
14. The dried pellet was resuspended in each well with 32.5 µl Resuspension Buffer. Gently pipette the entire volume up and down 10 times to mix thoroughly.
15. The PCR tube was incubated at room temperature for 2 minutes.
16. The PCR tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
17. 30 µl of the clear supernatant was transfered from each tube to new 0.5 ml tube.
18. The tubes were labeled and the genomic DNA libraies were stored at -20 °C.
19. 1 µl of library was used for Qubit BR assay quantification.
20. 50 ng of library was loaded on 2% agarose gel to check size distribution of adapter ligated fragments.

3.1.12 Exome Enrichment

This method describes Exome enrichment from the genomic DNA libraries. Exome enrichment process attempts to selectively enrich only gene coding regions of the genome. In human genome it constitutes less than 2% of the genome.

Materials and Reagents

Truseq Exome enrichment kit (FC-121-1008, Illimina) contains

Capture Target Oligos, Resuspension buffer, Streptavidin magnetic beads, Wash solution 1,

Resuspension buffer, Capture target buffer 1, Wash solution 1, Wash solution 2, Wash solution 3

Elute target buffer 1, Elute Target Buffer 2, PCR master mix, PCR Primer Cocktail

Freshly prepared 2N NaOH

AMPure XP Beads

PCR machine (Eppendorff)

Method

First Hybridization

This process mixes the DNA library with capture probes of targeted regions. The recommended hybridization time makes sure that targeted regions bind to the capture probes thoroughly. It also describes how to combine multiple libraries with different indices into a single pool prior to enrichment.

Preparation:-

- The Capture Target Oligos tube was removed from -15° to -25°C storage and thawed on ice.
- The Capture Target Buffer 1 tube was removed from -15° to -25°C storage and thawed at room Temperature.
- The thermal cycler was pre-programmed as follows for settings as given below in step 5
- A new and sterile 0.5 ml PCR tube was labeled as CT1 (Capture Target tube 1).

The table below was referenced for the amount of DNA libraries to be used for enrichment. Illumina recommends using 500 ng of each DNA library, quantified by the Qubit Fluorometric Quantitation system. If pooling libraries, 500 ng of each DNA library was added to the pool. If

the total volume is greater than 40 µl, a vacuum concentrator was used without heat to reduce the pooled sample volume to 40 µl.

Library Pool Complexity	Total DNA Library Mass (ng)
1-plex	500
2-plex	1000
3-plex	1500
4-plex	2000
5-plex	2500
6-plex	3000

1. The Capture Target Buffer 1 tube was vortexed for 5 seconds. Visually it was made sure that no crystal structures were present. (If crystals and cloudiness are observed, the Capture Target Buffer 1 tube should be vortexed until it appears clear.)
2. The following components were added in the order listed to 0.5 ml PCR tube labeled CT1 (Capture Target Tube 1). Multiply each volume by the number of sample pools being prepared. Prepare 5% extra reagent mix if multiple pooled samples are to be prepared. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.

Reagent	Volume (µl)
Diluted DNA library	40
Capture Target Buffer 1	50
Capture Target Oligos	10
Total Volume per Sample	100

3. The CT1 tube was closed. It was ensured that cap was tightly closed.
4. The CT1 tube was centrifuged to 280 xg for 1 minute.
5. The tightly closed CT1 tube was placed on the pre-programmed thermal cycler. The lid was closed and the tube was incubated using the pre-programmed settings:
 - a. Pre-heat lid and set to 100°C

- b. 95°C for 10 minutes
- c. 18 cycles of 1 minute incubations, starting at 93°C, then decreasing 2°C per cycle
- d. 58°C for 16–20 hours

First Wash

This process uses streptavidin beads to capture probes containing the targeted regions of interest. Three wash steps remove non-specific binding from the beads. The enriched library is then eluted from the beads and prepared for a second hybridization.

Preparation

- The Streptavidin Magnetic Beads, Elute Target Buffer 2, Wash Solution 1, and Wash Solution 3 tubes were removed from 2°C to 8°C storage in a refrigerator and left standing at room temperature.
 - The Elute Target Buffer 1 and Wash Solution 2 tubes were removed from -15° to -25°C storage and thawed at room temperature. 2N NaOH was freshly prepared.
1. The CT1 tube was removed from the thermal cycler.
 2. The CT1 tube was centrifuged to 280x g for 1 minute.
 3. The CT1 tube was placed on a tube stand and the cap was opened. Care was taken when removing the cap to avoid spilling the contents of the tubes. (It is normal to see a small degree of sample loss after overnight hybridization. However, if the sample loss is greater than 15%, Illumina does not recommended proceeding with the sample preparation. This amount of loss can be caused by poor sealing or not heating the lid.)
 4. Entire contents of the CT1 tube were transferred to a new 0.5 ml PCR tube and labeled as WT1 (Wash target tube 2).

5. The Streptavidin Magnetic Bead tube was vortexed until the beads were well dispersed, then 250 μ l of well-mixed Streptavidin Magnetic Beads were added to the WT1 tube. The entire volume was gently pipetted up and down 10 times to mix thoroughly.
6. The total 350 μ l volume of the tube was divided equally i.e. 175 μ l each in two 0.5 ml PCR tubes [WT1A and WT1B] for proper separation on magnetic stand.
7. The WT1A and WT1B tubes were closed.
8. The WT1A and WT1B tubes were left to stand at room temperature for 30 minutes.
9. The WT1A and WT1B tubes were centrifuged to 280 xg for 1 minute.
10. The caps of WT1A and WT1B tubes were opened.
11. The WT1A and WT1B tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
12. All of the supernatant from each tube was removed and discarded.
13. The WT1A and WT1B tubes were removed from the magnetic stand.

Wash 1 WT1 and Wash 2 WT1

The WS1 (Wash solution 1) Clean Up and WS2 (wash solution 2) Clean Up on the WT1A and WT1B tubes was performed as follows:

WS1 Clean Up

1. The Wash Solution 1 tube was vortexed for 5 seconds. Visually it was ensured that no crystal structures were present. (NOTE: If crystals are observed, vortex the Wash Solution 1 tube until no crystal structures are visible.)

2. 100 μ l Wash Solution 1 was added to each of the WT1A and WT1B tube. The entire volume was gently pipetted up and down 10–20 times to make sure the beads were fully resuspended.
3. The WT1A and WT1B tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
4. All of the supernatant from each the WT1A and WT1B tube was removed and discarded.
5. The WT1A and WT1B tubes were removed from the magnetic stand.

WS2 Clean Up

1. The Wash Solution 2 tube was vortexed for 5 seconds. Visually it was ensured that the Wash Solution 2 was mixed thoroughly.
2. 100 μ l Wash Solution 2 was added to each WT1A and WT1B tube. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly. Excessive bubbling or foaming was avoided. It was ensured that the beads were fully resuspended.
3. The WT1A and WT1B tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
4. All of the supernatant from each WT1A and WT1B tube was removed and discarded.
5. The WT1A and WT1B tubes were removed from the magnetic stand.
6. 100 μ l of Wash Solution 2 was added to the WT1A and WT1B tube. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly. Excessive bubbling or foaming was avoided. It was ensured that the beads were fully resuspended.

7. The contents of WT1A and WT1B tubes were further equally divided in four new 0.2 ml PCR tubes. 50 µl was transferred to each tube labeled IW1A, IW1B, IW1C, IW1D (Intermediate Wash tube 1).
8. Close the IW1 tubes caps tightly.
9. The IW1 tubes were incubated on the thermal cycler at 42°C for 30 minutes with a heated lid set to 100°C.

(NOTE: For optimal results, it is important that the thermal cycler lid be heated to 100°C.)

10. The magnetic stand was placed next to the thermal cycler for immediate access.
11. The IW1 tubes were removed from the thermal cycler and *immediately* placed on the magnetic stand for 2 minutes until the liquid appeared clear.
12. The IW1 tubes were opened gently.
13. All of the supernatant from each of the IW1 tubes was *immediately* removed and discarded.
14. The IW1 tubes were removed from the magnetic stand.
15. 50 µl Wash Solution 2 was added to each of the IW1 tubes (IW1A, IW1B, IW1C, IW1D). The entire volume was gently pipetted up and down 10–20 times and mixed thoroughly. Excessive bubbling or foaming was avoided. It was ensured that the beads were fully resuspended.

16. Repeated steps 8–13 once.

Wash 3 IW1 tubes

WS3 Clean Up and Elute Target were performed on the IW1 tubes as follows:

WS3 Clean Up

1. The IW1 tubes were removed from the magnetic stand.
2. 50 μ l Wash Solution 3 was added to each of the IW1 tube. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.
3. The IW1 tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
4. All of the supernatant from each of the IW1 tube was removed and discarded.
5. Steps 1–4 were repeated once.
6. To remove any residual Wash Solution 3, the cap of IW1 tube was closed tightly and the IW1 tube was briefly centrifuged to collect any residual Wash Solution 3.
7. The IW1 tube was placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
8. The IW1 tube was carefully opened to avoid spilling the contents of the tubess.
9. Any residual supernatant from the IW1 tubes was removed and discarded.

Elute Target

1. The following reagents were added, in the order listed, to a new 0.2 ml PCR tube to create the elution pre-mix. Multiplied each volume by the number of sample pools being prepared. Prepared 5% extra reagent mix if preparing multiple sample pools. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.

Reagent	Volume (μ l)
Elute Target Buffer 1	28.5
2N NaOH	1.5
Total Volume per Sample	30

2. Removed the IW1 tubes from the magnetic stand.
3. 30 μ l of the same elution pre-mix was added successively to each of the IW1A, IW1B, IW1C, IW1D tubes to collect all distributed pellet volume in one IW1 tube. The entire volume of each tube was gently pipetted up and down 10–20 times to mix thoroughly. It was ensured that the beads were fully resuspended.
4. The cap of IW1 tube was tightly closed.
5. The IW1 tubes were left to stand at room temperature for 5 minutes.
6. The IW1 tube was centrifuged to 280 x g for 1 minute.
7. The IW1 tube was placed on the magnetic stand for 2 minutes until the liquid appeared clear.
8. Carefully the cap of IW1 tube was opened to avoid spilling the contents of the tubes.
9. 29 μ l of supernatant was transferred from IW1 tube to the corresponding new 0.5 ml PCR tube labeled TT1 (Temporary target tube 1).
10. 5 μ l Elute Target Buffer 2 was added to the TT1 tube containing samples to neutralize the elution. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.
11. The TT1 tube was tightly closed.
12. The remaining reagents were stored as follows:
 - a. The Streptavidin Magnetic Beads, Elute Target Buffer 2, Wash Solution 1, and Wash Solution 3 tubes were kept in 2°C to 8°C storage.
 - b. The Elute Target Buffer 1, 2N NaOH, and Wash Solution 2 tubes were kept in -15°C to -25°C storage.
 - c. Any remaining elution pre-mix was discarded.

Safe Stopping Point

If proceeding to *Second Hybridization* is not immediately required, the protocol can be safely stopped here. If stopped, the TT1 tube should be tightly closed and stored at -15°C to -25°C for up to seven days.

Second Hybridization

This process mixes the first elution of the DNA library with the capture probes of target regions. The second hybridization makes sure that the targeted regions are further enriched.

Preparation:-

- The Capture Target Oligos tube was removed from -15°C to -25°C storage and thawed on ice.
- The Capture Target Buffer 1 tube was removed from -15°C to -25°C storage and thawed at room Temperature.
- The thermal cycler was pre-programmed as follows:
 - a. The pre-heat lid option was chosen and set to 100°C
 - b. 95°C for 10 minutes
 - c. 18 cycles of 1 minute incubations, starting at 93°C, then decreasing 2°C per cycle
 - d. 58°C for forever
- A new and sterile 0.5 ml PCR tube was labeled as CT2 (Capture Target tube 2).

6. The Capture Target Buffer 1 tube was vortexed for 5 seconds. Visually it was made sure that no crystals were present. (If crystals and cloudiness are observed, the Capture Target Buffer 1 tube should be vortexed until it appears clear.)
7. The following components were added in the order listed to 0.5 ml PCR tube labeled CT2 (Capture Target Tube 2). Multiplied each volume by the number of sample pools being prepared. Prepared 5% extra reagent mix if multiple pooled samples were to be prepared. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.

Reagent	Volume (µl)
Capture Target Buffer 1	50
Capture Target Oligos	10
PCR grade water (Autoclaved MilliQ)	10
First Elution from TT1 Tube	30
Total Volume per Sample	100

8. The CT2 tube was closed. It was ensured that cap was tightly closed.
9. The CT2 tube was centrifuged to 280 x g for 1 minute.
10. The tightly closed CT2 tube was placed on the pre-programmed thermal cycler. The lid was closed and tube was incubated using the pre-programmed settings:
 - a. pre-heat lid and set to 100°C
 - b. 95°C for 10 minutes
 - c. 18 cycles of 1 minute incubations, starting at 93°C, then decreasing 2°C per cycle
 - d. 58°C for 16–20 hours

Second Wash

This process uses streptavidin beads to capture probes containing the targeted regions of interest. Three wash steps remove non-specific binding from the beads. The enriched library is then eluted from the beads and prepared for sequencing.

Preparation

- The Streptavidin Magnetic Beads, Elute Target Buffer 2, Wash Solution 1, and Wash Solution 3 tubes were removed from 2°C to 8°C storage and left standing at room temperature.
 - The Elute Target Buffer 1 and Wash Solution 2 tubes were removed from -15°C to -25°C storage and thawed at room temperature. 2N NaOH was freshly prepared.
1. The CT2 tube was removed from the thermal cycler.
 2. The CT2 tube was centrifuged to 280 x g for 1 minute.
 3. The CT2 tube was placed on a tube stand and the cap was opened. Care was taken when removing the cap to avoid spilling the contents of the tubes. (It is normal to see a small degree of sample loss after overnight hybridization. However, if the sample loss is greater than 15%, Illumina does not recommend proceeding with the sample preparation. This amount of loss can be caused by poor sealing or not heating the lid.)
 4. Entire contents of the CT2 tube were transferred to a new 0.5 ml PCR tube and labeled as WT2 (Wash Target tube 2).

5. The Streptavidin Magnetic Bead tube was vortexed until the beads were well dispersed, then 250 μ l of well-mixed Streptavidin Magnetic Beads were added to the WT2 tube. The entire volume was gently pipetted up and down 10 times to mix thoroughly.
6. The total 350 μ l volume of the tube was divided equally i.e. 175 μ l each in two 0.5 ml PCR tubes [WT2A and WT2B] for proper separation on magnetic stand.
7. The WT2A and WT2B tubes were closed.
8. The WT2A and WT2B tubes were left standing at room temperature for 30 minutes.
9. The WT2A and WT2B tubes were centrifuged to 280x g for 1 minute.
10. The caps of WT1A and WT1B tubes were opened.
11. The WT2A and WT2B tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
12. All of the supernatant from each tube was removed and discarded.
13. The WT2A and WT2B tubes were removed from the magnetic stand.

Wash 1 WT2 and Wash 2 WT2

The WS1 (Wash solution 1) Clean Up and WS2 (wash solution 2) Clean Up on the WT2A and WT2B tubes was performed as follows:

WS1 Clean Up

1. The Wash Solution 1 tube was vortexed for 5 seconds. Visually it was ensured that no insoluble material was present. (NOTE: If crystals were observed, the Wash Solution 1 tube was vortexed until no insoluble matter was visible.)

2. 100 μ l Wash Solution 1 was added to each of the WT2A and WT2B tube. The entire volume was gently pipetted up and down 10–20 times to make sure the beads were fully resuspended.
3. The WT2A and WT2B tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
4. All of the supernatant from each the WT2A and WT2B tube was removed and discarded.
5. The WT2A and WT2B tubes were removed from the magnetic stand.

WS2 Clean Up

1. The Wash Solution 2 tube was vortexed for 5 seconds. Visually it was ensured that the Wash Solution 2 was mixed thoroughly.
2. 100 μ l Wash Solution 2 was added to each WT2A and WT2B tube. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly. Excessive bubbling or foaming was avoided. It was ensured that the beads were fully resuspended.
3. The WT2A and WT2B tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
4. All of the supernatant from each WT2A and WT2B tube was removed and discarded.
5. The WT2A and WT2B tubes were removed from the magnetic stand.
6. 100 μ l of Wash Solution 2 was added to the WT2A and WT2B tube. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly. Excessive bubbling or foaming was avoided. It was ensured that the beads were fully resuspended.

7. The contents of WT2A and WT2B tubes were further equally divided in four new 0.2 ml PCR tubes. 50 µl was transferred to each tube labeled IW2A, IW2B, IW2C, IW2D (Intermediate Wash tube 2).
8. IW2 tubes caps were closed tightly.
9. The IW2 tubes were incubated on the thermal cycler at 42°C for 30 minutes with a heated lid set to 100°C.

(NOTE: For optimal results, it is important that the thermal cycler lid be heated to 100°C.)

10. The magnetic stand was placed next to the thermal cycler for immediate access.
11. The IW2 tubes were removed from the thermal cycler and *immediately* placed on the magnetic stand for 2 minutes until the liquid appeared clear.
12. The IW2 tubes were opened gently.
13. All of the supernatant from each of the IW2 tubes was immediately removed and discarded.
14. The IW2 tubes were removed from the magnetic stand.
15. 50 µl Wash Solution 2 was added to each of the IW2 tubes (IW2A, IW2B, IW2C, IW2D). The entire volume was gently pipetted up and down 10–20 times and mixed thoroughly. Excessive bubbling or foaming was avoided. It was ensured that the beads were fully resuspended.
16. The steps 8–13 were repeated once.

Wash 3 IW2 tubes

WS3 Clean Up and Elute Target were performed on the IW1 tubes as follows:

WS3 Clean Up

1. The IW2 tubes were removed from the magnetic stand.
2. 50 μ l Wash Solution 3 was added to each of the IW2 tube. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.
3. The IW2 tubes were placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
4. All of the supernatant from each of the IW2 tube was removed and discarded.
5. The steps 1–4 were repeated once.
6. To remove any residual Wash Solution 3, the cap of IW2 tube was closed tightly and the IW2 tube was briefly centrifuged to collect any residual Wash Solution 3.
7. The IW2 tube was placed on the magnetic stand for 2 minutes at room temperature until the liquid appeared clear.
8. The IW2 tube was carefully opened to avoid spilling the contents of the tubes.
9. Any residual supernatant from the IW2 tubes was removed and discarded.

Elute Target

1. The following reagents were added, in the order listed, to a new 0.2 ml PCR tube to create the elution pre-mix. Each volume was multiplied by the number of sample pools being prepared. 5% extra reagent mix was prepared if preparing multiple sample pools. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.

Reagent	Volume (μ l)
Elute Target Buffer 1	28.5
2N NaOH	1.5
Total Volume per Sample	30

2. The IW2 tubes were removed from the magnetic stand.
3. 30 μ l of the same elution pre-mix was added successively to each of the IW2A, IW2B, IW2C, IW2D tubes to collect all distributed pellet volume in one IW1 tube. The entire volume of each tube was gently pipetted up and down 10–20 times to mix thoroughly. It was ensured that the beads were fully resuspended.
4. The cap of IW2 tube was tightly closed.
5. The IW2 tubes were left to stand at room temperature for 5 minutes.
6. The IW2 tube was centrifuged to 280x g for 1 minute.
7. The IW2 tube was placed on the magnetic stand for 2 minutes until the liquid appeared clear.
8. Carefully the cap of IW2 tube was opened to avoid spilling the contents of the tubes.
9. 29 μ l of supernatant was transferred from IW2 tube to the corresponding new 0.5 ml PCR tube labeled TT2 (Temporary target tube 2).
10. 5 μ l Elute Target Buffer 2 was added to the TT2 tube containing samples to neutralize the elution. The entire volume was gently pipetted up and down 10–20 times to mix thoroughly.
11. The TT1 tube was tightly closed.
12. The remaining reagents were stored as follows:
 - d. The Streptavidin Magnetic Beads, Elute Target Buffer 2, Wash Solution 1, and Wash Solution 3 tubes were kept in 2° to 8°C storage.
 - e. The Elute Target Buffer 1, 2N NaOH, and Wash Solution 2 tubes were kept in -15° to -25°C storage.
 - f. Any remaining elution pre-mix was discarded.

PCR Amplification

This process uses PCR to amplify the enriched DNA library for sequencing. PCR is performed with the same PCR primer cocktail used in TruSeq DNA library sample preparation.

Preparation:

- PCR Master Mix tube and PCR Primer Cocktail tube was removed from -15°C to -25°C storage, thawed and then was placed on ice.
- The thawed PCR Primer Cocktail and PCR Master Mix was briefly centrifuged for 5 seconds.
- AMPure XP beads were removed from storage and left to stand for at least 30 minutes to bring them to room temperature.
- The thermal cycler was pre-programmed as follows as required for the step 1 given below

PCR reaction was set up as follows

1. The following reagents were added to a new 0.5 ml PCR tube labelled TA1 (Target Amplification tube 1). The entire volume was gently pipette up and down 10 times to mix thoroughly.
2. The lid of the TA1 tube was closed
3. The tube was centrifuged at 280x g for 1 minute.

Amp PCR

1. The closed TA1 tube was placed on the preprogrammed thermal cycler and following PCR cycles were performed.
 - a. Pre heat lid option was chosen and set to 100⁰C

- b. 98°C for 30 seconds
- c. 10 cycles of
 - 98°C for 10 seconds
 - 60°C for 30 seconds
 - 72°C for 30 seconds
- d. 72°C for 5 minutes
- e. Hold at 10°C

Cleanup of Amplification Reaction was carried out as follows

1. The cap of TA1 tube was opened.
2. The AMPure XP beads were vortexed until the beads were well dispersed, then 90 µl of the mixed AMPure XP beads were added to TA1 tube containing 50 µl of the PCR amplified library. The entire volume was gently pipetted up and down 10 times to mix thoroughly.
3. The TA1 tube was incubated at room temperature for 5 minutes until the liquid appeared clear.
4. Using a 200 µl pipette, 140 µl of the supernatant was removed and discarded from the TA1 tube.
5. The TA1 tube was left on the magnetic stand while performing the following 80% ethanol steps.
6. With the TA1 tube remaining on the magnetic stand, 200 µl of freshly prepared 80% ethanol was added to the tube without disturbing the beads.

7. The TA1 tube was incubated for at least 30 seconds at room temperature, then the supernatant from each tube was removed and discarded.
8. The steps 6-7 were repeated for a total of two 80% ethanol washes.
9. The TA1 tube was kept on the magnetic stand and allowed to stand at room temperature for 15 minutes to dry, and then the tube was removed from the magnetic stand.
10. The dried pellet in the TA1 tube was resuspended with 30 µl of Resuspension Buffer. The entire volume was gently pipetted up and down 10 times to mix thoroughly.
11. The TA1 tube was incubated at room temperature for 5 minutes until the liquid appeared clear.
12. 30 µl of the clear supernatant (Exome enriched library) from the TA1 tube was transferred to new 0.5 ml PCR tube. The tube was labeled and stored at -20°C.

3.1.13 RNAseq Library Preparation

This process purifies the poly-A containing mRNA molecules using poly-dT oligo-attached magnetic beads using two rounds of purification. During the second elution of the poly-A RNA, the RNA is also fragmented and primed for cDNA synthesis

Materials and Reagents

Truseq RNA sample Prep Kit V2 (Catalogue No.:) contains Bead binding buffer, Elution buffer, Bead washing Buffer, Elute, Prime, Fragment mix, First Strand Master Mix, Resuspension Buffer, End Repair Mix, A-Tailing Mix, Ligation Mix, Stop Ligation Buffer

AMPure XP Beads

Magnetic stand

80% Ethanol

5X HF buffer

10 mM dNTPs

NEB Primer Universal

NEB Primer Indexes

NEB Universal Adapter

Autoclaved MilliQ Water

Autoclaved 0.2 ml and 0.5 ml PCR tubes

Autoclaved 1.5 ml and 2 ml microcentrifuge tubes

Micropipettes (20 μ l, 200 μ l, 1000 μ l) and corresponding autoclaved tips

Table-top microcentrifuge (Eppendorf)

PCR thermal Cycler

Method

A) Purify and Fragment mRNA

Preparation:

- The following reagents were removed from -15°C to -25°C storage and thawed at room temperature:

Bead Binding Buffer, Bead Washing Buffer, Elution Buffer, Elute, Prime, Fragment Mix, and Resuspension Buffer

- The RNA Purification Beads tube was removed from 2°C to 8°C storage and left to stand to bring to room temperature.
- The thermal cycler was Pre-programmed with the following three programs. For each program the pre-heat lid option was chosen and set to 100°C:
 - 65°C for 5 minutes, 4°C hold – saved as **mRNAD1**

- 80°C for 2 minutes, 25°C hold – saved as **mRNA1**
- 94°C for 8 minutes, 4°C hold – saved as **mRNAELU2**

Procedure:

1. The total RNA was diluted with nuclease-free ultra pure water to a final volume of 50 µl in a 0.5 ml PCR tube.
2. The room temperature RNA Purification Beads tube was vigorously vortexed to completely resuspend the oligo-dT beads.
3. 50 µl of RNA Purification Beads were added to each 0.5 ml PCR tube to bind the poly-A RNA to the oligo dT magnetic beads. The entire volume was pipetted up and down gently 6 times to mix thoroughly.
4. The tube was tightly closed and was placed on the pre-programmed thermal cycler. The lid was closed and mRNAD1 program was selected to denature the RNA and facilitate binding of the poly-A RNA to the beads.
5. The tube was removed from the thermal cycler when it reached 4°C at the end of the program.
6. The tube was placed on the bench and incubated at room temperature for 5 minutes to allow the RNA to bind to the beads.
7. The tube was placed on the magnetic stand at room temperature for 5 minutes to separate the poly-A RNA bound beads from the solution.
8. All of the supernatant from each tube was removed and discarded.
9. The tube was removed from the magnetic stand.

10. The beads were washed by adding 200 µl of Bead Washing Buffer in each tube to remove unbound RNA. Gently the entire volume was pipetted up and down 6 times to mix thoroughly.
11. The tube was placed on the magnetic stand at room temperature for 5 minutes.
12. The thawed Elution Buffer was spun for 5 seconds.
13. All of the supernatant from each tube was removed and discarded. The supernatant contains the majority of the ribosomal and other non-messenger RNA.
14. The tubes were removed from the magnetic stand.
15. 50 µl of Elution Buffer was added to each tube. Gently the entire volume was pipetted up and down 6 times to mix thoroughly.
16. The cap was closed tightly
17. The Elution Buffer tube was stored at 4°C.
18. The closed tube was placed on the pre-programmed thermal cycler. The lid was closed and mRNA1 was selected to elute the mRNA from the beads. This releases both the mRNA and any contaminant rRNA that has bound the beads non-specifically.
19. The tube was removed from the thermal cycler when it reached 25°C at the end of the program.
20. The tube was placed on the bench at room temperature and the cap was opened.
21. The thawed Bead Binding Buffer was spun for 5 seconds.
22. 50 µl of Bead Binding Buffer was added to each tube. This allows mRNA to specifically rebound the beads, while reducing non-specific binding of rRNA. Gently the entire volume was pipetted up and down 6 times to mix thoroughly.

23. The tube was incubated at room temperature for 5 minutes and the Bead Binding Buffer tube was stored at 2°C to 8°C.
24. The tube was placed on the magnetic stand at room temperature for 5 minutes.
25. All of the supernatant was removed and discarded from each of the tube.
26. The tube was removed from the magnetic stand.
27. The beads were washed by adding 200 µl of Bead Washing Buffer in each tube. Gently the entire volume was pipetted up and down 6 times to mix thoroughly.
28. The Bead Washing Buffer tube was stored at 2°C to 8°C.
29. The tube was placed on the magnetic stand at room temperature for 5 minutes.
30. All of the supernatant was removed and discarded from each tube. The supernatant contains residual rRNA and other contaminants that were released in the first elution and did not rebind the beads.
31. The tube was removed from the magnetic stand.
32. 19.5 µl of 'Elute, Prime, Fragment Mix' was added to each of the tube. Gently the entire volume was pipetted up and down 6 times to mix thoroughly. The Elute, Prime, Fragment Mix contains random hexamers for RT priming and serves as the 1st strand cDNA synthesis reaction buffer.
33. The cap was closed tightly.
34. The Elute, Prime, Fragment Mix tube was stored at -15°C to -25°C.
35. The sealed tube was placed on the pre-programmed thermal cycler. The lid was closed and MRNAELU2 program was selected to elute, fragment, and prime the RNA.
36. The tube was removed from the thermal cycler when it reached 4°C and centrifuged briefly.

37. Steps were immediately proceeded to First Strand cDNA Synthesis

B] First Strand cDNA Synthesis

This process reverse transcribes the cleaved RNA fragments primed with random hexamers into first strand cDNA using reverse transcriptase and random primers.

Preparation:

- One tube of First Strand Master Mix was removed from -15° to -25°C storage and thawed it at room temperature.
- The thermal cycler was Pre-programed with the following program and saved as RNA1STR:
 - Choose the pre-heat lid option and set to 100°C
 - 25°C for 10 minutes
 - 42°C for 50 minutes
 - 70°C for 15 minutes
 - Hold at 4°C

Procedure:

1. The tube was placed on the magnetic stand at room temperature for 5 minutes. without removing the tube from the magnetic stand the cap was opened.
2. 17 µl of the supernatant (fragmented and primed mRNA) was transferred from each tube to the corresponding new 0.5 ml PCR tube.
3. The thawed First Strand Master Mix tube was spun for 5 seconds.

4. First Strand Master Mix and SuperScript II mix was prepared at a ratio of 1 μ l SuperScript II for each 9 μ l First Strand Master Mix. The components were mixed gently, but thoroughly, and centrifuged briefly.
5. 8 μ l of prepared First Strand Master Mix and SuperScript II mix was added to each tube. Gently the entire volume was pipetted up and down 6 times to mix thoroughly.
6. The cap was closed and the tube was spun briefly.
7. The First Strand Master Mix tube was returned to -15° to -25°C storage immediately after use.
8. The tightly closed tube was placed on the pre-programmed thermal cycler. The lid was closed and the RNA1str program was selected.
9. When the thermal cycler reached 4°C, the tube was removed from the thermal cycler and steps were proceeded immediately to Second Strand cDNA Synthesis

C] Second Strand cDNA Synthesis

This process removes the RNA template and synthesizes a replacement strand to generate double stranded (ds) cDNA. AMPure XP beads are used to separate the ds cDNA from the second strand reaction mix.

Preparation:

- The Second Strand Master Mix was removed from -15°C to -25°C storage and thawed at room temperature.
- The Resuspension Buffer was removed from 2°C to 8°C storage and was brought to room temperature.

- The AMPure XP beads were removed from storage and left to stand at room temperature for at least 30 minutes to bring them to room temperature.
- The thermal cycler was pre-heated to 16°C.

Procedure:

1. The thawed Second Strand Master Mix was spun for 5 seconds.
2. 25 µl of thawed Second Strand Master Mix was added to each tube. Gently the entire volume as pipetted up and down 6 times to mix thoroughly.
3. The cap was closed tightly.
4. The sealed tube was placed on the pre-heated thermal cycler. The lid was closed and the tube was incubated at 16°C for 1 hour.
5. The tube was removed from the thermal cycler, the cap was opened, and left to stand at room temperature to bring the plate to room temperature.
6. The AMPure XP beads were vortexed until they were well dispersed, then 90 µl of well mixed AMPure XP beads were added to each tube containing 50 µl of ds cDNA. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
7. The tube was incubated at room temperature for 15 minutes.
8. The tube was placed on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads were bound to the side of the wells.
9. 135 µl of the supernatant was removed and discarded from each tube.
10. With the tube remaining on the magnetic stand, 200 µl of freshly prepared 80% ethanol was added to each well without disturbing the beads.
11. The tube was incubated at room temperature for 30 seconds, and then all of the supernatant was removed and discarded from each tube.

12. Steps 10 and 11 were repeated once for a total of two 80% ethanol washes.
13. The tube was left to stand at room temperature for 15 minutes to dry and then the tube was removed from the magnetic stand.
14. The Resuspension Buffer equilibrated to room temperature was spun for 5 seconds.
15. 52.5 µl Resuspension Buffer was added to each tube. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
16. The tube was incubated at room temperature for 2 minutes.
17. The tube was placed on the magnetic stand at room temperature for 5 minutes.
18. 50 µl of the supernatant (ds cDNA) was transferred from the tube to the new 0.5 ml PCR tube.

Safe Stopping Point: If it is not planned to proceed to Perform End Repair immediately, the protocol can be safely stopped. Tightly closed and sealed with parafilm, the tube can be stored at -15° to -25°C for up to seven days.

D] End Repair

This process converts the overhangs resulting from fragmentation into blunt ends using an End Repair Mix. The 3' to 5' exonuclease activity of this mix removes the 3' overhangs and the polymerase activity fills in the 5' overhangs.

Preparation:

- The End Repair Mix was removed from -15°C to -25°C storage and thawed at room temperature.
- The use of the End Repair Control is optional and it can be replaced with the same volume of Resuspension Buffer.

- The Resuspension Buffer was removed from 2°C to 8°C storage and it was brought to room temperature.
- The AMPure XP beads were removed from storage and left to stand for at least 30 minutes to bring them to room temperature.
- The ds cDNA tube was removed from -15°C to -25°C storage (if it was stored), and left to stand to thaw at room temperature then spun for 1 minute.
- The thermal cycler was pre-heated to 30°C with the pre-heat lid option set to 100°C

Procedure:

1. As the in-line End Repair control reagent was not used, 10 µl of Resuspension Buffer was added to each tube that contains 50 µl of ds cDNA.
2. 40 µl of End Repair Mix was added to each tube containing the ds cDNA. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
3. The cap was closed tightly.
4. The tube was placed on the pre-heated thermal cycler. The lid was closed and incubated at 30°C for 30 minutes.
5. The tube was removed from the thermal cycler.
6. The AMPure XP Beads were vortexed until they were well dispersed, then 160 µl well mixed AMPure XP Beads were added to each tube containing 100 µl of End Repair Mix. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
7. The tube was incubated at room temperature for 15 minutes.
8. The tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.

9. Using a 200 µl single channel pipette set to 127.5 µl, 127.5 µl of the supernatant was removed and discarded from tube.
10. Step 9 was repeated once.
11. With the tube on the magnetic stand, 200 µl of freshly prepared 80% ethanol was added to each tube without disturbing the beads.
12. The tube was incubated at room temperature for 30 seconds, then all of the supernatant was removed and discarded from each tube. Care was taken not to disturb the beads.
13. Steps 11 and 12 were repeated once for a total of two 80% ethanol washes.
14. The tube was left to stand at room temperature for 15 minutes to dry, then the tube was removed from the magnetic stand.
15. The dried pellet in each tube was resuspended with 17.5 µl Resuspension Buffer. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
16. The tube was incubated at room temperature for 2 minutes.
17. The tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
18. 15 µl of the clear supernatant was transferred from each tube to the new 0.5 ml PCR tube.

Safe Stopping Point: If it is not planned to proceed to Adenylate 3' Ends immediately, the protocol can be safely stopped. Tightly closed and sealed with parafilm, the tube can be stored at -15°C to -25°C for up to seven days.

E] Adenylation of 3' Ends

A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single 'T' nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment. This strategy ensures a low rate of chimera (concatenated template) formation.

Preparation:

- The A-Tailing Mix was removed from -15°C to -25°C storage and thawed at room temperature:
- The resuspension Buffer was removed from 2°C to 8°C storage and brought to room temperature.
- End Repaired tube was removed from -15°C to -25°C storage (if it was stored) and left stand to thaw at room temperature. The tube was spun for 1 minute.
- The thermal cycler was pre-programmed with the following program and saved as ADYNYLT:

The pre-heat lid option was chosen and set to 100°C

- 37°C for 30 minutes
- 70°C for 5 minutes
- Hold at 4°C

Procedure:

1. As the in-line A-Tailing Control reagent was not used, 2.5 µl of Resuspension Buffer was added to each tube.

2. 12.5 µl of thawed A-Tailing Mix was added to each tube. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
3. The cap was closed tightly.
4. The tightly closed tube was placed on the pre-programmed thermal cycler. The lid was closed and program ADYNYLT was started.
5. When the thermal cycler temperature reached 4°C, the tube was removed from the thermal cycler, then steps were proceeded immediately to Ligate Adapters.

F] Adapter Ligation

This process ligates multiple indexing adapters to the ends of the ds cDNA, preparing them for hybridization onto a flow cell.

Preparation:

- The required NEB Universal Adapter tubes were removed from -15°C to -25°C storage and thawed at room temperature.
- The Resuspension Buffer was removed from 2°C to 8°C storage and brought to room temperature.
- The AMPure XP beads were removed from storage and left to stand for at least 30 minutes to bring them to room temperature.
- The thermal cycler was Pre-heated to 30°C for ligation step.
- The thermal cycler was pre programmed with heated lid at 100°C and 37°C for 15 minutes for USER enzyme treatment step.

Procedure:

1. The thawed NEB Universal Adapter was spun for 5 seconds.
2. The Ligation Mix tube was removed from -15°C to -25°C storage just before use.
3. As the in-line Ligation Control reagent was not used, 2.5 µl of Resuspension Buffer was added to each tube.
4. Reagents were added according to following table:

A Tailed Reaction	30 µl
Resuspension Buffer/ Ligation Control(1:100 diluted)	2.5 µl
Ligation Mix	2.5 µl
NEB Universal Adapter	1 µl
Resuspension Buffer	1.5 µl
<hr/>	
Total	37.5 µl

5. The tube was placed on the pre-heated thermal cycler. The lid was closed and the tube was incubated at 30°C for 10 minutes.
6. The Ligation Mix tube was returned back to -15°C to -25°C storage immediately after use.
7. The tube was removed form thermal cycler.
8. 3 µl of USER enzyme was added to the reaction tube and the tube was incubated at 37°C for 15 Minutes.
9. The AMPure XP Beads were vortexed until they were well dispersed, then 40 µl of mixed AMPure XP Beads were added to each tube. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.

10. The tube was incubated at room temperature for 15 minutes.
11. The tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
12. The supernatant from each tube was removed and discarded.
13. With the tube remaining on the magnetic stand, 200 μ l of freshly prepared 80% ethanol was added to each tube without disturbing the beads.
14. The tube was incubated at room temperature for 30 seconds, then all of the supernatant from the tube was removed and discarded.
15. Steps 13 and 14 were repeated once for a total of two 80% Ethanol washes.
16. While keeping the tube on the magnetic stand, the samples were left to air dry at room temperature for 15 minutes and then the tube was removed from the magnetic stand.
17. The dried pellet in each tube was resuspended with 22.5 μ l Resuspension Buffer. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
18. The tube was incubated at room temperature for 2 minutes.
19. The tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
20. 20 μ l of the clear supernatant from each tube was transferred to a fresh 0.5 ml PCR tube.

Safe stopping point: If it is not planned to proceed to Enrich DNA Fragments immediately, the protocol can be safely stopped. Tightly closed and sealed with parafilm, the tube can be stored at -15°C to -25°C for up to seven days.

G] Enrichment of adapter ligated DNA Fragments/Amplification of the library

This process uses PCR to selectively enrich those DNA fragments that have adapter molecules on both ends and to amplify the amount of DNA in the library. The PCR is performed with a PCR primer cocktail that anneals to the ends of the adapters. The number of PCR cycles should be minimized to avoid skewing the representation of the library.

Preparation:

- The Resuspension Buffer was removed from 2°C to 8°C storage and brought to room temperature.
- The AMPure XP beads were removed from storage and left to stand for at least 30 minutes to bring them to room temperature.
- The following reagents were thawed to room temperature and spun for 5 sec and placed on ice.

5X HF buffer

10 mM dNTPs

NEB Primer Universal

NEB Primer Index

- The adapter ligated sample tube was removed from -15° to -25°C storage, if it was stored at the conclusion of Adapter Ligation and left to stand to thaw at room temperature.
- The thermal cycler was pre-programmed with the following program and saved as NEBPCR:

- The pre-heat lid option was chosen and set to 100°C
- 98°C for 10 seconds
- 15 cycles of:

- 98°C for 10 seconds
- 65°C for 30 seconds
- 72°C for 30 seconds
- 72°C for 5 minutes
- Hold at 4°C

Procedure:

1. The following reaction was set:

Purified Adapter Ligated Fragments	20 µl
5X HF buffer	10 µl
10 mM dNTPs	1 µl
NEB Primer Universal	1 µl
NEB Primer Index(Sample Specific)	1 µl
Phusion Polymerase	0.5 µl
Autoclaved MilliQ Water	16.5 µl
<hr/>	
Total	50 µl

2. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
3. The tightly closed tube was placed on the pre-programmed thermal cycler. The lid was closed and NEBPCR program was started.
4. The tubes were removed from thermal cycler after completion of the PCR.
5. The AMPure XP Beads were vortexed until they were well dispersed, then 50 µl of the mixed AMPure XP Beads were added to each tube containing 50 µl of the PCR amplified library. The entire volume was gently pipetted up and down 10 times to mix thoroughly.

6. The tube was incubated at room temperature for 15 minutes.
7. The PCR tubes were placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
8. 95 μ l of the supernatant was removed and discarded from each tube.
9. With the tube remaining on the magnetic stand, 200 μ l of freshly prepared 80% ethanol was added to each tube without disturbing the beads.
10. The PCR tube was incubated at room temperature for 30 seconds, then all of the supernatant was removed and discarded from each tube.
11. Repeated steps 8 and 9 for a total of two 80% ethanol washes.
12. While keeping the tube on the magnetic stand, the samples were left to air dry at room temperature for 15 minutes and then the tube was removed from the magnetic stand.
13. The dried pellet in each tube was resuspended with 32.5 μ l Resuspension Buffer. Gently the entire volume was pipetted up and down 10 times to mix thoroughly.
14. The tube was incubated at room temperature for 2 minutes.
15. The tube was placed on the magnetic stand at room temperature for 5 minutes or until the liquid appeared clear.
16. 30 μ l of the clear supernatant i.e. transcriptome library was transferred from each tube to a new 0.5 ml PCR tube.
17. The tube was tightly closed, labeled and sealed with a parafilm and then stored at -15°C to -25°C.

3.1.14 Targeted Sequencing on Ion Torrent Platform

Libraries for targeted sequencing on Ion Torrent platform were prepared using Ion Ampliseq Library Kit 2.0 (Thermo Fisher Scientific Inc, Catalogue number: 4475345) following standard protocol Ion Ampliseq DNA library preparation user guide (Thermo Fisher Scientific Inc, Publication part number: MAN0006735, Revision B.0) from Ion Torrent (www.thermofisher.com). Ion Ampliseq Cancer Hotspot Panel v2 (Thermo Fisher Scientific Inc, Catalogue number: 4477685) which amplifies cancer hotspot 207 regions from 50 genes using 207 primer pairs. To amplify DNA targets 10 ng of input tumour DNA was used. The libraries were sequenced on Ion PGM (Thermo Fisher Scientific Inc, Personal Genome Machine, Life technologies) following standard protocols: Ion PGM Template OT2 200 Kit (Thermo Fisher Scientific Inc, Life technologies, Catalogue Number: 4480974), user guide (Thermo Fisher Scientific Inc, Publication number: MAN0007220 , Revision B.0, www.thermofisher.com), and Ion PGM Sequencing 200 Kit v2 (Thermo Fisher Scientific Inc, Life technologies, Catalogue Number: 4482006), user guide (Thermo Fisher Scientific Inc, Publication number: MAN0007273, Revision 3, www.thermofisher.com).

3.2 BIOINFORMATIC METHODS

3.2.1 Quality Analysis of fastq Reads

FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) a quality control tool for high throughput sequence data was used for quality analysis of fastq reads. FastQC is a GUI based software. The fastq reads data was imported in to the software in .fastq or .bam format. Results were displayed after completion of the analysis by the software. The results were saved.

3.2.2 Alignment of fastq Reads to hg19 by BWA

The fastq files were aligned to the human reference genome hg19 with (BWA) Burrows-Wheeler Aligner version 0.7.9 (www.bio-bwa.sourceforge.net). Following commands were used.

```
$ /bwa-0.7.9a/bwa aln -f sample_R1.sai hg19_GATK.fa sample_R1.fastq
```

```
$ /bwa-0.7.9a/bwa aln -f sample_R2.sai hg19_GATK.fa sample_R2.fastq
```

```
$ /bwa-0.7.9a/bwa sampe -r '@RG\tID:sample-GATK\tPL:illumina\tLB:Illumina-TrueSeq\tSM:sample\tCN:ACTREC' hg19_GATK.fa sample_R1.sai sample_R2.sai sample_R1.fastq sample_R2.fastq > sample_GATK.sam
```

3.2.3 Removal of Adapter Overlaps

The reads of the RNA sequencing data containing adapter overlaps were cleaned using the reads trimming tool Trimmomatic version 0.32 (<http://www.usadellab.org>).

```
$ java -jar trimmomatic-0.32.jar PE -threads 2 -trimlog sample-trimmomatic-trimlog  
sample_R1.fastq sample_R2.fastq sample_R1_paired.fastq sample_R1_unpaired.fastq  
sample_R2_paired.fastq sample_R2_unpaired.fastq ILLUMINACLIP:/Neb-Next-Primers-Index-  
PE.fa:2:30:10:1:true SLIDINGWINDOW:4:20 MINLEN:36
```

3.2.4 Alignment of fastq Reads to hg19 by Tophat

The the adapter overlap cleaned reads were aligned to the reference human genome hg19 using TopHat version 2.0.13 (<http://ccb.jhu.edu/software/tophat>) with default parameters. Following commands were used.

```
$ /Tophat-2.0.13/bin/tophat --num-threads 3 /hg19 sample_R1.fastq sample_R2.fastq
```

```
$ java -Xmx4g -Djava.io.tmpdir=/tmp -jar /picard-tools-1.80/SortSam.jar SO=coordinate  
INPUT=sample-accepted_hits.bam OUTPUT=sample-accepted_hits_picard.bam  
VALIDATION_STRINGENCY=LENIENT CREATE_INDEX=true
```

```
$ java -Xmx4g -Djava.io.tmpdir=/tmp -jar /picard-tools-1.80/MarkDuplicates.jar  
INPUT=sample-accepted_hits_picard.bam OUTPUT=sample-  
accepted_hits_picard_MarkRmDup.bam METRICS_FILE=sample-  
accepted_hits_picardRMDUP-metrics.txt VALIDATION_STRINGENCY=LENIENT  
CREATE_INDEX=true REMOVE_DUPLICATES=true
```

3.2.5 Removal of Duplicate Reads

Duplicate reads were removed using the Picard Tools version 1.80 (<http://broadinstitute.github.io/picard>). Following commands were used.

```
$ java -Xmx4g -Djava.io.tmpdir=/tmp -jar picard-tools-1.80/SortSam.jar SO=coordinate  
INPUT=sample_GATK.sam OUTPUT=sample_GATK_picard.bam  
VALIDATION_STRINGENCY=LENIENT CREATE_INDEX=true
```

```
$ java -Xmx4g -Djava.io.tmpdir=/tmp -jar picard-tools-1.80/MarkDuplicates.jar  
INPUT=sample_GATK_picard.bam OUTPUT=sample_GATK_picard_MarkRmDup.bam  
METRICS_FILE=sample-RMDUP-metrics.txt VALIDATION_STRINGENCY=LENIENT  
CREATE_INDEX=true REMOVE_DUPLICATES=true
```

3.2.6 Refinement of Alignment by Genome Analysis Toolkit

The alignment files were refined by local realignment of the reads and base quality recalibration by the Genome Analysis Toolkit (GATK) version 2.1.3 (<https://www.broadinstitute.org/gatk>). Following commands were used.

```
$ java -Xmx4g -jar /GenomeAnalysisTK-2.1.3/GenomeAnalysisTK.jar -T  
RealignerTargetCreator -R hg19.fa -o sample_GATK_picard_MarkRmDup.bam.list -I  
sample_GATK_picard_MarkRmDup.bam
```

```
$ java -Xmx4g -Djava.io.tmpdir=/tmp -jar /GenomeAnalysisTK-2.1.3/GenomeAnalysisTK.jar -  
T IndelRealigner -R hg19.fa -I sample_GATK_picard_MarkRmDup.bam -targetIntervals  
sample_GATK_picard_MarkRmDup.bam.list -o  
sample_GATK_picard_MarkRmDup_realigned.bam
```

```
$ java -Djava.io.tmpdir=/tmp -jar /picard-tools-1.80/FixMateInformation.jar  
INPUT=sample_GATK_picard_MarkRmDup_realigned.bam  
OUTPUT=sample_GATK_picard_MarkRmDup_realigned_fixed.bam SO=coordinate  
VALIDATION_STRINGENCY=LENIENT CREATE_INDEX=true
```

```
$ java -Xmx4g -jar /GenomeAnalysisTK.jar -I INFO -R hg19.fa -knownSites:dbsnp,VCF  
hg19_dbsnp_132_sorted.vcf -I sample_GATK_picard_MarkRmDup_realigned_fixed.bam -T  
CountCovariates -cov ReadGroupCovariate -cov QualityScoreCovariate -cov CycleCovariate -  
cov DinucCovariate -recalFile sample_GATK_picard_MarkRmDup_realigned_fixed_recal-  
data.csv
```

```
$ java -Xmx4g -jar /GenomeAnalysisTK.jar -I INFO -R hg19.fa -I  
sample_GATK_picard_MarkRmDup_realigned_fixed.bam -T TableRecalibration --out  
sample_GATK_picard_MarkRmDup_realigned_fixed_recal.bam -recalFile  
sample_GATK_picard_MarkRmDup_realigned_fixed_recal-data.csv
```

3.2.7 Pileup of aligned reads

Read pileup file was created using following command. Samtools (samtools.sourceforge.net/)
version samtools-0.1.18 was used.

```
$ /samtools-0.1.18/samtools mpileup -f hg19.fa  
Sample_GATK_picard_MarkRmDup_realigned_fixed_recal.bam >  
Sample_GATK_picard_MarkRmDup_realigned_fixed_recal.mpileup
```

3.2.8 Exome Enrichment Analysis

The quality of exome data was analyzed using NGSrich and ngsCAT. Following commands were used.

NGSrich

```
$ java NGSrich evaluate -r Sample_GATK_picard_realigned_fixed_recal.bam -u hg19 -t  
TruSeq_exome_targeted_regions.hg19.bed -a refGene.txt -T tmp/ -o output
```

ngsCAT

```
$ ngscat.v0.1/ngscat.py --bams=/filelist --bed=TruSeq_exome_targeted_regions.hg19.bed --  
out=out/ coveragethrs=10,15,20,30,50 --reference/hg19.fa --saturation=y  
coveragethrs=10,15,20,30,50 --tmp=/tmp/ --threads=4
```

3.2.9 RNAseq quality analysis

RNAseq data quality was estimated using RNA-seQC [<http://archive.broadinstitute.org/cancer/cga/rna-seqc/>] software (version v1.1.8.1). As an annotation file gencode.v7 was used.

```
$ java -jar RNA-SeQC v1.1.8.jar -n 1000 -s SampleInfo.txt -t / annotation.gtf -r hg19.fa -o  
output -strat gc -gc gencode.v7.gc.txt --transcriptDetails
```

3.2.10 Identification of Somatic Simple Nucleotide Variations

Simple nucleotide variants (SNVs) and insertions and deletions (indels) were identified using the VarScan variant detection tool version 2.3.5 (<http://varscan.sourceforge.net>) using the filtering criteria of a minimum coverage 10 and at least 5 somatic variants.

1. Reads pileup files were input to VarScan to detect somatic simple nucleotide variations with following parameters.

```
$ java -jar /home/Dell/Vijay/VarScan.v2.3.5.jar somatic  
SampleB_GATK_picard_MarkRmDup_realigned_fixed_recal.mpileup  
SampleT_GATK_picard_MarkRmDup_realigned_fixed_recal.mpileup  
Sample_Varscan_somatic_defaultstrn --tumor-purity 0.80 --output-vcf 1
```

2. The variant calls in the somatic output file were separated by somatic status (Germline, Somatic, LOH) and were classified as high-confidence (.hc) or low-confidence (.lc) using command processSomatic.


```
$ java -jar /home/Dell/Vijay/VarScan.v2.3.5.jar processSomatic  
PDT2_Varscan_somatic_defaultstrn.indel.vcf --min-tumor-freq .20 --max-normal-freq .05  
  
$ java -jar /home/Dell/Vijay/VarScan.v2.3.5.jar processSomatic  
PDT2_Varscan_somatic_defaultstrn.snp.vcf --min-tumor-freq .20 --max-normal-freq .05
```

3. To further refine the somatic mutation list to identify true positive mutation calls
somaticFilter Command was used with default minimum read depth of 10 and minimum
supporting 2 reads for a variant

```
$ java -jar VarScan.v2.3.5.jar somaticFilter  
Sample_Varscan_somatic_defaultstrn.indel.vcf.Somatic.hc --output-file  
Sample_Varscan_somatic_defaultstrn.indel.vcf.Somatic.hc.filterMinCov10MinReads2-4  
  
$ java -jar VarScan.v2.3.5.jar somaticFilter  
Sample_Varscan_somatic_defaultstrn.snp.vcf.Somatic.hc --output-file  
Sample_Varscan_somatic_defaultstrn.snp.vcf.Somatic.hc.filterMinCov10MinReads2-4
```

3.2.11 Annotation of Somatic Variants

Annotation of the somatic variants identified in the analysis of Exome data was done using the ANNOVAR software (www.openbioinformatics.org/annovar) (Wang et al., 2010).

1. The vcf4 file output generated by VarScan software was converted into the annovar input format using following command.

```
$ /annovar/convert2annovar.pl -format vcf4 -includeinfo  
Sample_Varscan_somatic_defaultstrn.indel.vcf.Somatic.hc.filterMinCov10MinReads2-4 >  
Sample_Varscan_somatic_defaultstrn.indel.vcf.Somatic.hc.filterMinCov10MinReads2-4.avinput
```

2. Variant annotation was carried out using summarize annovar script.

```
$ /annovar/summarize annovar.pl --verbose --outfile  
Sample_Varscan_somatic_defaultstrn.indel.vcf.Somatic.hc.filterMinCov10MinReads2-4.avinput  
--buildver hg19 --remove --verdb SNP 137 --ver1000g 1000g2012apr --veresp 6500si --genotype  
refgene --checkfile --alltranscript  
Sample_Varscan_somatic_defaultstrn.indel.vcf.Somatic.hc.filterMinCov10MinReads2-4.avinput  
/annovar/humandb
```

3. From the ANNOVAR-annotated list, variants located in segmental duplications were excluded. The remaining variants were manually verified in IGV (www.broadinstitute.org/igv). Ambiguous variants (variants represented in reads with low mapping quality, variants present near indels, and variants surrounded by mismatched bases) were discarded.

3.2.12 Analysis of Somatic Copy Number Variations (CNVs) in Exome Data

The copy number variations in the tumor genome were identified from the paired exome sequence data using the FishingCNV software version 1.5.2 (<http://sourceforge.net/projects/fishingcnv>). Segmentation means of less than -3 and more than 0.3 were considered as deletion and amplification, respectively. The copy number variations in the tumor genome were also analyzed using the Control-FREEC software (<http://bioinfoout.curie.fr/projects/freec/>), which uses input aligned reads in samtools mpileup format to construct and normalize the copy number profile and the B-allele frequency (BAF) profile. By performing segmentation of both profiles, it ascribes the genotype status and annotates genomic alterations using both copy number and allelic frequency information.

3.2.12.1 Analysis of Somatic Copy Number Variations using FishingCNV

1. The software settings were set as follows.

```
Rscript=/usr/bin/R
PairedCNV= /FishingCNV_1.5.2/pairedCNV.R
UnpairedCNVwPCA= /FishingCNV_1.5.2/unpairedCNVwPCA_2012_11_26.R
ExtractRPKM=/FishingCNV_1.5.2/extractRPKM_FishingCNV.r
UnpairedCNV= /FishingCNV_1.5.2/unpairedCNV_v1.0.R
CorrectP=/FishingCNV_1.5.2/correctp_2013_01_14.R
Fasta= /hg19.fa
CCDS= 44M_029368_D_BED_20101013.bed or
CCDS= TruSeq exome targeted regions.hg19.bed
GATK=/home/Dell/Vijay/FishingCNV_1.5.2/GenomeAnalysisTK-2.3-9/GenomeAnalysisTK.jar
```

2. On “RPKM file” tab “Produce RPKM files from BAM files” was selected. Exome bam files were loaded and program was run to produce RPKM files.
3. RPKM files were generated for all Blood and Normal Exome sample data.

4. In the Paired CNV tab in the program, from the scroll down menu, “CNV from coverage files” option was chosen.
5. Blood (Control) RPKM coverage file and paired Tumour (Test) RPKM coverage file was loaded.
6. The software was run for analysis. The segmentation files produced as a result of the analysis were visualized in IGV (Integrative genomics Viewer).

3.2.12.2 Analysis of Somatic Copy Number Variations using Control-FREEC

1. The configuration file config_hg19.txt was set as follows.

```
[general]

chrLenFile = /FREEEC/Databases/hg19.len

window = 3000

step = 1000

ploidy = 2

chrFiles= hg19/

intercept=1

minMappabilityPerWindow = 0.7

printNA=FALSE

outputDir = SAMPLE-FREEC/

sex=XX

#sex=XY

breakPointType=4

[sample]
```

```
mateFile = Tumour.mpileup  
  
inputFormat = pileup  
  
mateOrientation = FR  
  
[control]  
  
mateFile = Blood.mpileup  
  
inputFormat = pileup  
  
mateOrientation = FR  
  
[BAF]  
  
SNPfile = /FREEC/Databases/hg19_snp137.SingleDiNucl.1based.txt  
  
minimalCoveragePerPosition = 10  
  
[target]  
  
captureRegions = 44M_029368_D_BED_20101013.bed  
  
#captureRegions = TruSeq_exome_targeted_regions.hg19.bed
```

2. The software was run with following command.

```
$ FREEC/freec -conf config_hg19.txt
```

3. The resultant Tumour.mpileup_ratio.txt Tumour.mpileup_BAF.txt were used to plot the CNV status of the sample.

```
$ cat /home/Dell/Vijay/FREEC/makeGraph.R | R --slave --args 2 Tumour.mpileup_ratio.txt
Tumour.mpileup_BAF.txt
```

3.2.13 Differential Gene Expression Analysis Using EdgeR

The read count based gene level differential expression analysis comparing the transcriptome profiles of the *CIC*-mutant and *CIC*-wild type oligodendrogliomas was carried out using the EdgeR package of R bioconductor (www.bioconductor.org).

1. Sample information was loaded with following command

```
> samples<-read.table(file="Sample_LGG-CIC-mut-Vs-wt.txt", sep='\t', header=TRUE)
```

2. The edgeR package as loaded and the utility function, readDGE, was used to read in the COUNT files created from htseq-count:

```
> library("edgeR")
> counts<-readDGE(samples$countf)$counts
```

3. Weakly expressed and noninformative (e.g., non-aligned) features were removed.

```
> noint<-rownames(counts) %in% c("no_feature", "ambiguous", "too_low_aQual",  
"not_aligned", "alignment_not_unique", "__alignment_not_unique", "__ambiguous", "__no_feature",  
"__too_low_aQual", "__not_aligned")  
> cpms<-cpm(counts)  
> keep<-rowSums(cpms>1) >=3 & !noint  
> counts<-counts[keep,]
```

4. The count table was visualized and inspected as follows.

```
> colnames(counts)<-samples$shortname  
> head(counts[,order(samples$condition)],5)
```

5. A DGEList object was created.

```
> d<-DGEList(counts=counts,group=samples$condition)
```

6. Normalization factors were estimated using:

```
> d=calcNormFactors(d)
```

7. The relationships between samples was inspected using a multidimensional scaling (MDS) plot.

```
> plotMDS(d, labels="o",col=c("darkgreen","blue")[factor(samples$condition)])
```

8. Tagwise dispersion was estimated.

```
> d=estimateCommonDisp(d)
> d=estimateTagwiseDisp(d)
```

9. A visual representation of the mean-variance relationship was created using the plotMeanVar and plotBCV functions.

```
> plotMeanVar(d, show.tagwise.vars=TRUE, NLine=TRUE)
> plotBCV(d)
```

10. A differential expression test ('classic' edgeR) was carried out.

```
> de<-exactTest(d, pair=c("1p19qDel-IDH1/2mut-CICmut","1p19qDel-IDH1/2Mut-CICwt"))
```

11. To present a tabular summary of the differential expression statistics topTags function was used.

```
> tt<-topTags(de, n=nrow(d))
> head(tt$table)
```

12. The depth-adjusted reads per million was inspected for some of the top differentially expressed genes.


```
> nc<-cpm(d, normalized.lib.sizes=TRUE)
> rn=rownames(tt$table)
> head(nc[rn,order(samples$condition)],5)
```

13. A graphical summary, such as an M (log-fold change) versus A (log-average expression) was plotted.

```
> deg<-rn[tt$table$FDR<.05]
> plotSmea(d, de.tags=deg)
```

14. The result table was exported as a CSV file.

```
write.csv(tt$table, file="CIC-mut-Vs-wt-toptags_edgeR.csv")
```

3.2.14 Gene Expression Read Count

Raw counts for the reads aligned to the gene intervals were produced by the python package HTSeq version 0.6.1 (www.huber.embl.de/users/anders/HTSeq) using the default union-counting mode. gencode.v19.annotation.gtf file was used for gene annotation.

```
$ htseq-count -s no -a 10 sample-Trimomatic-Tophat-accepted hits-nameSorted.sam gtf >
sample-htseq.v19.count
```

3.2.15 Differential Gene Expression Analysis Using SAM

The variance stabilized transformed gene expression count data was obtained following steps in section 3.2.16 (Gene Expression Data Normalization Using DESeq). The normalized transformed

data was used as input in Significance Analysis of Microarrays software, MeV version 4.9.0 (<http://tm4.org>) [120]. Subsets of samples were assigned groups i.e. *CIC*-mutant and *CIC*-wild type accordingly using graphical user interface. The differential gene expression analysis was performed between the two subgroups.

3.2.16 Gene Expression Data Normalization Using DESeq

A total of 65 *IDH1/IDH2*-mutant, 1p/19q codeleted oligodendroglioma tumors for which the RNAseq V2 data were available were used for differential gene expression analysis comparing the transcriptome profiles of the *CIC*-mutant and *CIC*-wild type tumor tissues. The gene level RSEM (<http://deweylab.biostat.wisc.edu/rsem/>) raw counts from the TCGA RNAseq V2 data were rounded to the nearest integer for each gene in each sample. The data were normalized by variance stabilizing transformation using the R-Bioconductor (R-3.1.2) package DESeq that takes into account the RNA-seq data size of each sample (<http://bioconductor.org/packages/release/bioc/html/DESeq.html>).

1. The sample information was loaded from sample information file.

```
> samples<-read.csv(file="samples.csv", sep=';', header=TRUE)
```

2. A data frame was created with the required metadata. The read counts from the read count files were already rounded to the nearest integer.

```
> samplesDESeq=with(samples,data.frame(shortname=l(shortname),  
countf=l(countf),condition=condition,LibraryLayout=LibraryLayout))
```

3. The DESeq package was loaded and a CountDataSet object was created.

```
> library("DESeq")  
> cds=newCountDataSetFromHTSeqCount(samplesDESeq)
```

4. Normalization factors were estimated.

```
> cds=estimateSizeFactors(cds)
```

5. The size factors were inspected using.

```
> sizeFactors(cds)
```

6. A variance-stabilizing transformation was applied.

```
> cdsB=estimateDispersions(cds, method="blind")  
> vsd=varianceStabilizingTransformation(cdsB)
```

7. The variance stabilized data was exported to a text file. This data file was used as an input for the differential gene expression analysis in the *CIC*-mutant vs *CIC*-wild type oligodendrogliomas using the significance analysis of microarrays (SAM) tool in the MeV version 4.9.0 (www.TM4.org).

```
> write.table(as.data.frame(exprs(vsd)),file="TCGA-LGG-OligodendrogliomaOnly-  
VarianceStabilizedData.txt", sep="\t",col.names=TRUE,row.names=TRUE)
```

8. To inspect sample relationships a principal component analysis (PCA) plot was generated.

```
> p=plotPCA(vsd, intgroup=c("condition","LibraryLayout"))
```

9. Function `estimateDispersions` was used to calculate dispersion values:

```
> cds=estimateDispersions(cds)
```

10. The estimated dispersions were inspected using the `plotDispEsts` function.

```
> plotDispEsts(cds)
```

11. The test for differential expression was performed by using function `nbinomTest`.

```
> res=nbinomTest(cds,"IDH1-2-1p-19q-CIC-Mut","IDH1-2-1p-19q-CIC-Wt")
```

12. To display differential expression (log-fold changes) versus expression strength (log-average read count), function `plotMA` was used.

```
> plotMA(res)
```

13. The result table was exported to a comma separated value text file.

```
> write.csv(res, file="TCGA-LGG-CIC-Mut-Vs-Wt-res_DESeq.csv")
```

3.2.17 Gene Set Enrichment Analysis

The gene set enrichment analysis was carried out using the SeqGSEA package (1.8.0 version) of the R bioconductor (www.bioconductor.org) (R-3.1.2). Gene set enrichment analysis was performed to find out gene sets significantly enriched in the differentially expressed genes between *CIC*-mutant and *CIC*-wild type samples.

1. Gene count table was prepared using CountTable function of DESeq package of R-Bioconductor.

```
> samples<-read.table(file="LGG_TCGA_All_CNV2_RNAseqV2_Sample_LGG-CIC-mut-Vs-wt.txt", sep='\t', header=TRUE)
> samplesDESeq=with(samples,data.frame(shortname=I(shortname), countf=I(countf), condition=condition, gender=gender,grade=tumor_grade))
> library("DESeq")
> cds=newCountDataSetFromHTSeqCount(samplesDESeq)
> CountTable<-(counts(cds))
> write.table(CountTable, file="TCGA-OligoDendroglioma65sample1to39CICmut40to65CICwtReadCountsForGSEA.txt", sep="\t", row.names = TRUE, col.names = TRUE)
```

2. Libraries SeqGSEA and doParallel were loaded.

```
> library(SeqGSEA)
> library(doParallel)
```

3. Multithreading option was specified that 4 cores to be used in computing by assigning four clusters.

```
> cl <- makeCluster(4)
> registerDoParallel(cl)
```

4. Permutation number of times were set

```
> perm.times <- 1000
```

5. Output prefix for files was set

```
> output.prefix<-"TCGA-LGG-OligoD-CIC-mutVswt"
```

6. Data was imported as data frame with rownames and column names. The first 39 samples were *CIC*-mutant while rest 26 were *CIC*-wild type, thus they were labeled as such. Differential gene expression analysis was carried out to find out differentially expressed genes.

```
> geneCounts <- read.csv(file="TCGA-
OligoDendroglioma65sample1to39CICmut40to65CICwtReadCountsForGSEA.txt", sep = "\t",
header=TRUE, row.names=1)
> label <- as.factor(c(rep("CIC-mut",39), rep("CIC-wt",26)))
> DEG <-runDESeq(geneCounts, label)
> DEGres <- DENBStat4GSEA(DEG)
```

7. DE NB statistics was calculated on the permutation data sets.

```
> permuteMat <- genpermuteMat(label, times=perm.times)
> DEpermNBstat <- DENBStatPermut4GSEA(DEG, permuteMat) # permutation
```

8. Differential expression score was calculated

```
> DEscore.normFac <- normFactor(DEpermNBstat)
> DEscore <- scoreNormalization(DEGres$NBstat, DEscore.normFac)
> DEscore.perm <- scoreNormalization(DEpermNBstat, DEscore.normFac)
```

9. Differential Expression and Differential Splicing score integration. As Differential Splicing score was not calculated, it was not included. (DSscore can be null).

```
> gene.score <- geneScore(DEscore, DEweight=1)
> gene.score.perm <- genePermuteScore(DEscore.perm, DEweight=1)
> plotGeneScore(DEscore, DEscore.perm, pdf="GSEA.DEScore.pdf", main="Expression")
```

10. Gene set data was loaded. msigdb.v5.0.symbols.gmt (c5: gene ontology (GO) gene sets) and c6.all.v5.2.symbols.gmt (c6: oncogenic signatures gene sets) data bases were used in two separate analysis run.

```
> geneIDs <- rownames(geneCounts)
> gene.set <- loadGenesets("msigdb.v5.0.symbols.gmt", geneIDs, geneID.type="gene.symbol",
genesize.min = 5, genesize.max = 1000)
```

11. Enrichment analysis was performed.

```
> gene.set <- GSEnrichAnalyze(gene.set, gene.score, gene.score.perm, weighted.type=1)
```

12. The results were exported to a text file.

```
> write.table(GSEAres, paste(output.prefix, ".GSEA.result.txt", sep=""), quote=FALSE, sep="\t",
row.names=FALSE)
```

3.2.18 Mutation Based Stratification of Copy number Alterations

The somatic mutation and copy number alteration level 3 data was downloaded from TCGA data portal [<https://tcga-data.nci.nih.gov>]. The copy number alteration data was stratified according to the mutation frequencies of the samples. The stratified data was visualized in integrative genomics viewer (IGV) [software.broadinstitute.org/software/igv/].

3.2.19 HMG box Domain Analysis

In silico analysis was carried out to predict functional effects of the nonsynonymous mutations in the *CIC* exon 5 which codes for HMG (high mobility group) box domain, which is a DNA binding domain.

CIC HMG box Domain Sequence (highlighted bold and underlined):

```
>NP_055940.3 protein capicua homolog isoform CIC-S [Homo sapiens]
MYSÄHRPLMPASSAASRGLGMFVWTNVEPRSVAVFPWHSLVPFLAPSQPDPSVQPSEAQQPASHPVASNQ
SKEPAESA AVAHERPPGGTGSADPERPPGATCPESP GPGPPHPLGVVESGKGPPPTTEEEASGPPGEPRL
DSETESDHDDAFLSIMSPEIQLPLPPGKRRTQSLSALPKERDSSSEKDGRSPNKRE KDHIRRP MNAFMIF
SKRHRALVHQ RHPNQDNRTVSKILGEWWYALGPKEKQKYHDLAFQVKEAHFKAHPDWKWCNKDRKKSSE
AKPTSLGLAGGHKETRERSMSETGTAAAPGVSSSELLSVAAQTLLSSDTKAPGSSSCGAERLHTVGGPGSA
RPRAFSHSGVHSLDGGGEVDSQALQELTQMVSGPASYS GPKPSTQYGAPGPFAAPGEGGALAATGRPPLLP
TRASRSQRAASEDMTSDEERMVICEEGDDDDVIADDGFGTTDIDLKCKERVTDSESGDSSGEDPEGNKG F
GRKVFSPVIRSSFTHCRPPLDPEPPGPPDPVAFGKGYSAPSSSASSPASSSASAATSFSLSGSGTFKAQ
ESGQGSTAGPLRPPPPGAGGPATPSKATRFLPMDPATFRKRPE SVGGLEPPGPSVIAAPPSGGGNILQT
LVLPPNKEEQEGGGARVPSAPAPSLAYGAPAAPLSRPAATMVTNVVRPVSSTPVPIASKPFPTSGRAEAS
```



```
PNDTAGARTEMGTGSRVPGGSPLGVSLVYSDKKSAAATSPAPHLVAGPLLGTGKAPATVTNLLVGTPGY
GAPAPPAVQFIAQGAPGGGTTAGSGAGAGSGPNGPVPLGILQPGALGKAGGITQVQYILPTLPQQQLQVAP
APAPAPGTKAAAPSGPAPTTTSIRFTLPPGTSTNGKVLAAATPTPGIPILQSVPSAPPPKAQSVSPVQAPP
PGGSAQQLLPKGKVLVPLAAPSMSVRGGGAGQPLPLVSPFVSVPVQNGAQPPSKI IQLTPVPVSTPSGLVPP
LSPATLPGPTSQPQKVLLPSSTRITYVQSAGGHALPLGTSPASSQAGTVTSYGPTSSVALGFTSLGSPGP
AFVQPLLSAGQAPLLAPGQVGVSVPVSPQLPPACAAPGGPVITAFYSGSPAPTSSAPLAQPSQAPPSLVY
TVATSTTPPAATILPKGPPAPATATPAPTSPPFPSATAGSMTYSLVAPKAQRPSPKAPQKVKAASIASIPVG
SFEAGASGRPGPAPRQPLEPGPVREPTAPESELEGQPTPPAPPPLPETWTPTARSSPPLPPPAEERTSAK
GPETMASKFPSSSSDWRVPGQGLENRGEPTPPSPAPAPAVAPGGSSESSSGRAAGDTPERKEAAGTGKK
VKVRPPPLKKTDFSVDNRVLSEVDFEERFAELPEFRPEEVLPSPTLQSLATSPRAILGSYRKKRKNSTD
DSAPEDPTSPKRKMRRSSSCSEPNTPKSAKCEGDI FTFDRTGTEAEDVLGELEYDKVPYSSLRRTLDQR
RALVMQLFQDHGFFPSAQATAAFQARYADIFPSKVCLQLKIREVRQKIMQAATPTEQPPGAEAPLPVPPP
TGTAAPAPPTSPAGGPDPTSPSSDSGTAAQAPPLPPPPESGPGQPGWEGAPQSPPPPPGPSTAATGR
```

HMG domain amino acids 197-276 [KDHIRRPMNAFMIFSKRHRALVHQRHPNQDNRTVSKI
LGEWWYALGPKEKQKYHDLAFQVKEAHFKAHPDWKWCNKDRKK]

3.2.19.1 BLAST

CIC HMG domain amino acid sequence was queried with Standard Protein BLAST [<https://blast.ncbi.nlm.nih.gov/Blast.cgi>] with default parameters.

3.2.19.2 FRpred conservation Analysis

FRpred Analysis was performed to identify conserved amino acids among the amino acid sequence of the CIC –HMG box domain. The CIC HMG box domain amino acid sequence [a.a. 197-276] [KDHIRRPMNAFMIFSKRHRALVHQRHPNQDNRTVSKILGEWWYALGPKEKQKYHDLAFQVKEAHFKAHPDWKWCNKDRKK] was analyzed in FRpred [<https://toolkit.tuebingen.mpg.de/frpred>] with default PSI-BLAST parameters (Max. number of iterations 10, E-value threshold 1E-3, Min. coverage 80%, Min. identity 25%).

3.2.19.3 Analysis in VMD

VMD (Visual Molecular Dynamics, Version 1.9.2, <http://www.ks.uiuc.edu/Research/vmd/>) software suit was used to perform *in silico* structure analysis. Crystographically derived HMG box domain structures were downloaded from pdb database [<http://www.rcsb.org/pdb/home/home.do>].

The amino acid sequences of the HMG box domains from the pdb structure files were aligned using Clustalw in Multiseq [www.ks.uiuc.edu/Research/vmd/plugins/multiseq/]. HMG box domain structures from pdb structure files were aligned using STAMP structural alignment tool from Multiseq.

RESULTS

4 RESULTS

Based on the histopathological diagnosis, oligodendroglioma tumor tissues were recruited for the whole exome sequencing study. Before proceeding for the exome sequencing study, the tumor DNA was analyzed for known alterations reported to occur frequently in oligodendroglioma tissues. Mutation in the *IDH1* gene encoding Isocitrate dehydrogenase, at the amino acid residue R132 or that in *IDH2* gene at the position R172 was reported to occur in 70-80% low grade gliomas [57], [86], [114]. Therefore, genomic DNA from 11 pure oligodendroglioma tissues was analysed for the presence of *IDH1* or *IDH2* mutation. Further, loss of one copy of chromosome 1p arm and that of chromosome 19q arm was also reported to occur frequently in oligodendrogliomas [31], [75], [76]. Genomic DNA from the tumor tissues and their paired blood DNA was also therefore, analyzed for the Loss of Heterozygosity (LOH) of chromosome 1p/19q arms using microsatellite marker analysis.

Before DNA extraction, it was ensured that the tumor tissue used contained at least 80% tumor cells by microscopic examination (done by Dr. E. Sridhar, Pathologist, TMH) of Hematoxylin & Eosin stained cryosections of the tissues. Genomic DNA was extracted from the tumor tissues and paired whole blood using QIAamp DNA mini Kit as per the manufacturer's protocol (Qiagen, GmbH, Hilden, Germany).

4.1 Detection of *IDH1* and *IDH2* Mutation status by Sanger Sequencing.

IDH1 mutations affecting amino acid R132 were detected by Sanger sequencing of the exon 4 of *IDH1* gene. The samples which were found to be wild type for *IDH1* R132 were further tested for *IDH2* R172 mutational status by sequencing exon 4 of *IDH2* gene. The sequences of the

primers used for PCR are given in Appendix I Electropherograms of the exon 4 sequences of the *IDH1/IDH2* gene of the 11 tumour tissues sequenced are shown in the figure 4.1.

Table 4.1: Histology, WHO grade and *IDH1/IDH2* mutation status of 11 tumour samples studied.

Sample	Histology	WHO Grade	IDH1 / IDH2 Status
ODG1	Oligodendroglioma	Grade III	IDH1 R132H
ODG2	Oligodendroglioma	Grade III	IDH1 R132H
ODG3	Oligodendroglioma	Grade II	IDH1 R132H
ODG4	Oligodendroglioma	Grade III	IDH1 R132H
ODG5	Oligodendroglioma	Grade III	IDH1 R132H
ODG6	Oligodendroglioma	Grade III	IDH1 R132H
ODG7	Oligodendroglioma	Grade II	IDH1 R132H
ODG8	Oligodendroglioma	Grade II	IDH1 R132H
ODG9	Oligodendroglioma	Grade III	IDH2 R172K
ODG10	Oligodendroglioma	Grade III	IDH1 R132C
ODG11	Oligodendroglioma	Grade III	IDH1 R132 wild type and IDH2 R172 wild type

Of the 11 tumour samples tested for *IDH1/IDH2* mutation status, 10 tumour tissues were found to be *IDH1/IDH2* mutated (Figure 4.1). One tumor was found to be wild type for both *IDH1* R132 and *IDH2* R172. Among the 10 mutated tumour tissues, nine tumour samples were found to be *IDH1* mutated (90%) and one tumor was found to be *IDH2* (R172K) mutated. Among the nine *IDH1* mutated tumour tissues, eight (88.9%) were found to carry R132H mutation while one was found to have R132C mutation.

Table 4.1 shows the summary of the *IDH1/IDH2* mutational status of the 11 oligodendrogliomas studied as well as their histopathological grading done by pathologist (Dr. E. Sridhar, Tata Memorial Hospital).

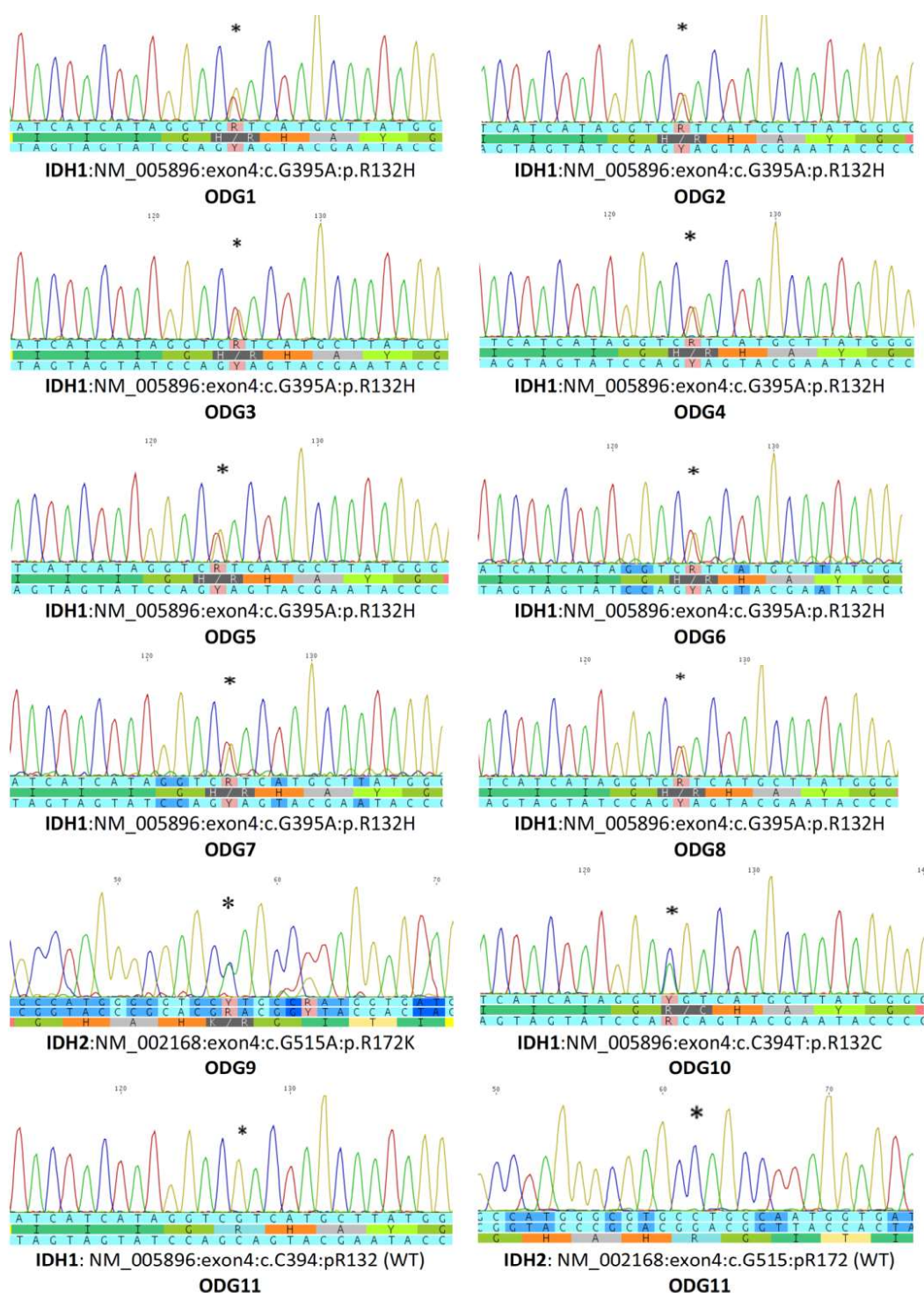


Figure 4.1: *IDH1* / *IDH2* mutation status of the studied tumours with oligodendroglioma morphology. The Mutated nucleotide is indicated with *.

4.2 Loss of Heterozygosity analysis for chromosomal arms 1p/19q in oligodendroglial tumours.

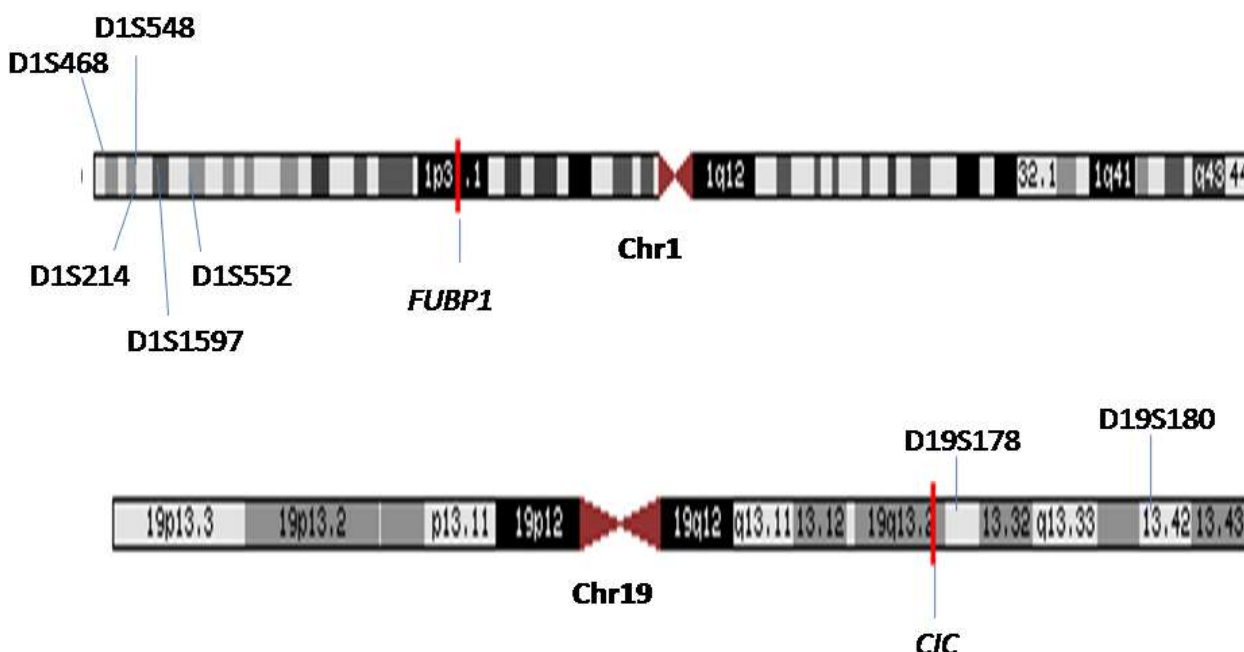
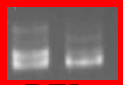


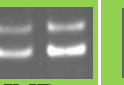
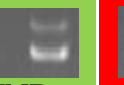
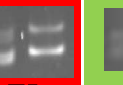

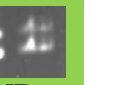
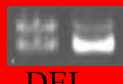

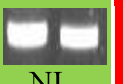
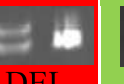
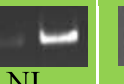
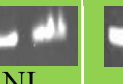
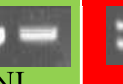



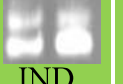
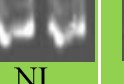




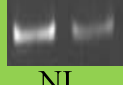

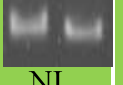

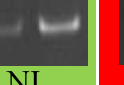



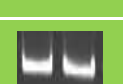






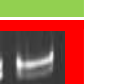




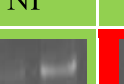













Figure 4.2: Chromosomal Locations of the microsatellite markers used for loss of heterozygosity analysis.

Microsatellites or Short Tandem Repeats are short DNA sequences of 2-5 bp in length that are repeated multiple times (typically 5 to 50 times) in genome. The microsatellites exhibit length polymorphisms and hence serve as markers for loss of heterozygosity analysis. [121], [122]. Loss of heterozygosity (LOH) analysis using microsatellite markers present on chromosome 1p and 19q was carried to identify of 1p/19q chromosomal arm deletion. Five microsatellite markers located on chromosomal arm 1p and two microsatellite markers located on 19q were used for the loss of heterozygosity analysis. Figure 4.2 shows the microsatellite markers used; their location on chromosomal arms.

Table 4.2: Summary of Microsatellite marker loss of heterozygosity on chromosomal arm 1p and 19q. Five microsatellite markers used for assessing the loss of heterozygosity status of chromosome arm 1p (name begins with D1) and two microsatellite markers used for assessing the loss of heterozygosity status of chromosome arm 19q (name begins with D19) in eight oligodendroglial tumour samples and deletion status of each marker in each sample. The left lane in Polyacrylamide gel image is for blood DNA and the right lane is for tumour DNA used to amplify that particular microsatellite. DEL-deleted, NI-non informative, IND-informative not deleted, NA-not available.

Microsatellite	ODG1	ODG 2	ODG3	ODG4	ODG 5	ODG6	ODG 7	ODG12
D1S468	 DEL	 IND	 IND	 IND	 IND	 DEL	 IND	 IND
D1S548	 DEL	 DEL	 NI	 DEL	 NI	 NI	 NI	 DEL
D1S214	 IND	 NI	 IND	 NI	 NI	 NI	 NI	 IND
D1S1597	 NI	 DEL	 NI	 NI	 NI	 DEL	 NI	 NI
D1S552	 NI	 NI	 NI	 NI	 NI	 NI	 NI	 DEL
D19S178	 DEL	 DEL	 DEL	 DEL	 IND	 DEL	 DEL	 IND
D19S180	 IND	 IND	 IND	 IND	 IND	 IND	 IND	 IND

The primer sequences used for amplification of these microsatellite markers are listed in Appendix I. The genomic DNA of the oligodendroglial tumour tissues and that of the corresponding paired blood (as normal tissue control) was used for PCR amplification of the

individual microsatellite markers and subsequent resolution and visualization on 10 % Polyacrylamide gel. The results are summarized in table 4.2.

Informative cases are those which showed two or more bands on gel after PCR amplification. If amplified microsatellite marker from DNA extracted from blood showed two bands whereas the same microsatellite marker amplified from DNA extracted from tumour shows single band the tumour sample was interpreted as loss of heterozygosity with respect to that microsatellite on the particular chromosome arm. If the amplified microsatellite region from DNA extracted from blood showed single band the case was termed as non-informative due to absence of polymorphism.

High frequency of microsatellite loss of heterozygosity was observed in the oligodendroglial samples studied. Five out of eight samples studied showed microsatellite loss of heterozygosity on chromosomal arm 1p while six out of eight samples studied showed microsatellite loss of heterozygosity on chromosomal arm 19q.

4.3 Identification of Somatic Mutations and Copy Number Variations by Exome

Sequencing of Oligodendroglial Paired Tumour-Normal Samples.

The genomic DNA from tumor tissues and paired blood was quantified in a Qubit 2.0 fluorimeter (Life Technologies, Carlsbad, CA) and the quality was checked by agarose gel electrophoresis before proceeding for library preparation for exome sequencing

DNA libraries for the paired end multiplex deep sequencing were prepared using the DNA library preparation kit for Illumina from Kapa Biosystems (Wilmington, MA) as per the manufacturer's protocol. Two microgram genomic DNA was sheared using a Covaris M220 focused ultrasonicator (Covaris, Woburn, MA), end repaired, and ligated to the single indexed

DNA adapters, followed by the separation of the fragmented DNA by agarose gel electrophoresis and purification of ~300 bp DNA fragments using QIAquick gel extraction kit (Qiagen). Exome capture was performed using the Truseq Exome enrichment kit (Catalogue No.-FC-121-1008, Illumina, San Diego, CA) that captures 62 Mb exomic region corresponding to the 20,794 genes. The multiplexed exome libraries were subjected to 100 bp paired end deep sequencing using the HiSeq 1500 ultra-high-throughput sequencing system (Illumina). The exome sequencing was done to get at least 50 X average depth of coverage.

4.3.1 Exome Sequence Data Quality.

Quality of the raw sequence data was analyzed using FastQC [<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>]. Mean and median base Phred quality score for almost all bases was above 30. FastQC analysis of the aligned reads for each sample is provided in the Appendix II - Exome Data Quality.

All exomes were sequenced at > 50 X average depth of coverage except one sample ODG11-Blood which was sequenced at 42.59 X average depth of coverage. The average depth of coverage ranged from 42.59 X to 79.64 X. More than 95% of the targeted Exome region was covered at least 1 X in each sample. The percentage of exome covered at ≥ 30 X ranged between 50.45% in ODG11-B to 85.25% in ODG9-B. The percentage of exome covered at ≥ 10 X ranged between 88.09% in ODG11-B to 95.56% in ODG9-B. The detailed statistics of the total number of reads obtained, coverage of the targeted exome for each of tumor and its paired blood DNA sequenced is given in table 4.3.

Table 4.3 : Exome Sequencing Data Statistics of 22 exomes. (B-Blood, T-Tumour)

Exome Sample	# Unique Reads	# On Target ± 100 bp	Target Size (bp)	# Target Regions	Coverage Mean	Coverage Std Deviation	Covered 1x	Covered 5x	Covered 10x	Covered 20x	Covered 30x
ODG1-T	61060488	36126909	46401242	194060	56.1	33.42	96.15%	92.66%	88.66%	79.31%	69.13%
ODG1-B	57721820	36693473	46401242	193825	56.9	31.62	96.2%	93.11%	89.66%	81.03%	71.21%
ODG2-T	68530920	42957664	46401242	193999	67.46	40.05	96.26%	93.26%	90.07%	82.35%	73.75%
ODG2-B	54799566	34078840	46401242	193768	53.26	29.68	96.17%	92.91%	89.11%	79.63%	69%
ODG3-T	57625664	35069131	46401242	194476	54.35	31.38	96.48%	93.11%	89.34%	79.72%	68.88%
ODG3-B	50269122	33638194	46401723	193612	52.1	28.66	96.15%	93.1%	89.46%	79.98%	69%
ODG6-T	56699726	34130172	46401242	194065	53.15	32.51	96.18%	92.52%	88.16%	77.63%	66.28%
ODG6-B	67808548	36901519	46401242	195053	57.4	32.51	96.81%	93.12%	89.44%	80.73%	71.09%
ODG7-T	61095960	38433460	46401242	194119	59.7	34.7	96.17%	92.9%	89.2%	80.58%	71.25%
ODG7-B	59852610	36811311	46401242	194292	57.19	32.15	96.34%	93.24%	89.78%	81.04%	71.09%

ODG4-T	66332809	33075991	62286597	200834	69.24	51.4		97.57%	95.44%	93.57%	88.94%	80.38%
ODG4-B	66712423	32775927	62286597	201025	68.7	49.69		97.92%	95.87%	94.02%	89.78%	81.55%
ODG5-T	53140831	26466597	62286597	201021	55.56	45.82		97.38%	94.94%	92.12%	82.5%	67.19%
ODG5-B	57048202	28579487	62286597	201028	59.77	45.6		97.54%	95.21%	93.05%	86.45%	73.62%
ODG8-T	72718512	43667659	62286597	200866	57.44	44.81		96.01%	93.43%	91%	82.85%	69.85%
ODG8-B	73176614	46130311	62286597	200895	60.63	48.16		95.36%	92.85%	90.52%	83.09%	71.54%
ODG9-T	63049395	30626755	62286597	200847	64.09	47.75		97.73%	95.47%	93.45%	88.02%	77.41%
ODG9-B	85017967	37328017	62286597	200838	79.64	58.78		98.79%	97.15%	95.56%	91.89%	85.25%
ODG10-T	69397156	41355244	62286597	201025	54.49	44.56		96.01%	93.32%	90.47%	80.63%	65.94%
ODG10-B	80973115	50178742	62286597	201027	65.83	56.32		95.72%	93.26%	90.96%	83.65%	72.47%
ODG11-T	77896918	48105792	62286597	200853	63.23	56.17		95.8%	93.29%	90.9%	83.07%	70.93%
ODG11-B	54270056	32370160	62286597	200834	42.59	36.32		95.91%	92.64%	88.09%	71.47%	50.45%

4.3.2 Copy Number Variations in the Oligodendroglial Tumours.

The copy number variation analysis was performed by comparing tumor exome data in with that of its paired blood exome data. Copy number variation analysis of the exome sequence data of 11 tumors diagnosed histopathologically as oligodendrogliomas was done using two algorithms, FishingCNV software version 1.5.2 (<http://sourceforge.net/projects/fishingcnv>) [123] and Control-FREEC [124], taking into account both the coverage and B allele frequency. FishingCNV software identifies copy number variations in paired Exome data by comparing the coverage depth of the tumour exome reads aligned to reference genome hg19 to the coverage depth of aligned exome reads from blood (control). Segmentation means of less than 0.3 and more than 0.3 were considered as deletion and amplification, respectively.

The copy number variations in the tumor genome were also analyzed using the Control-FREEC software (<http://bioinfo-out.curie.fr/projects/freec/>), which uses input aligned reads in samtools mpileup format to construct and normalize the copy number profile and the B-allele frequency (BAF) profile. By performing segmentation of both profiles, it ascribes the genotype status and annotates genomic alterations using both copy number and allelic frequency information.

Both FishingCNV and Control FREEC analysis identified concurrent loss of a copy of 1p and 19q, a known characteristic of oligodendrogliomas, in 9 out of the 11 tumor tissues (Figure 4.3 and 4.4). Other than the 1p/19q codeletion, recurrent chromosome 14 deletions were found in three 1p/19q codeleted tumors (ODG2, ODG5, ODG8).

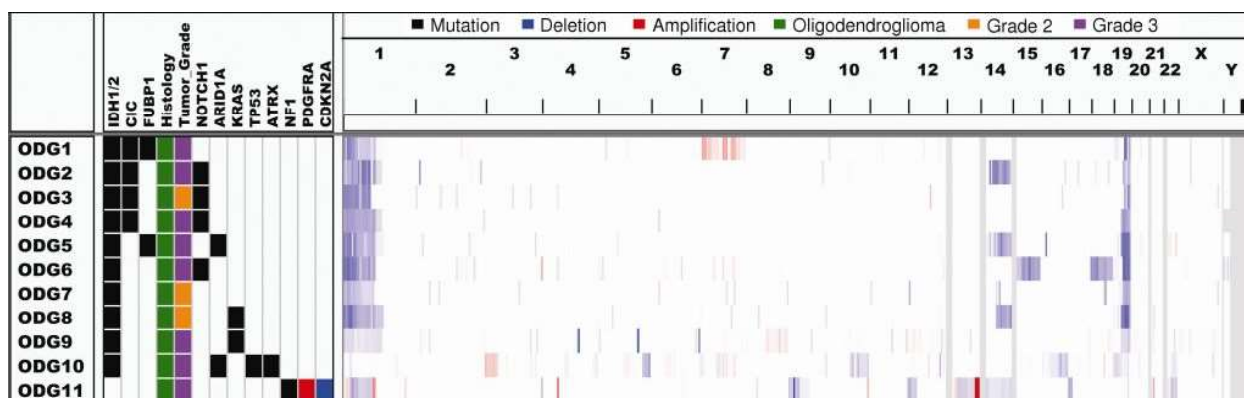


Figure 4.3: Integrated genomic view of the copy number variation and somatic mutations of the 11 oligodendrogliomas as analyzed using the FishingCNV software. The tumors are numbered sequentially (ODG1 to ODG11).

Other copy number alterations that were detected were chromosome 7 amplification in one sample ODG1. Chromosome 15 and chromosome 18 deletions were found in sample ODG6. A focal deletion on chromosome 9p containing *CDKN2A* gene and a focal amplification of region containing gene *PDGFRA* was found in *IDH1/2* wild type sample ODG11.

Using Control FREEC similar pattern of copy number variations was detected (Figure 4.4). The brown colored points represent each diploid heterozygous nucleotide (frequency approaching 0.5) and blue colored points represent nucleotides approaching homozygous frequency (0.0 or 1.0) due to loss of heterozygosity as compared to heterozygous status of the nucleotide in the normal sample. The black horizontal lines highlight predicted deletion and violet horizontal lines highlight predicted amplification.

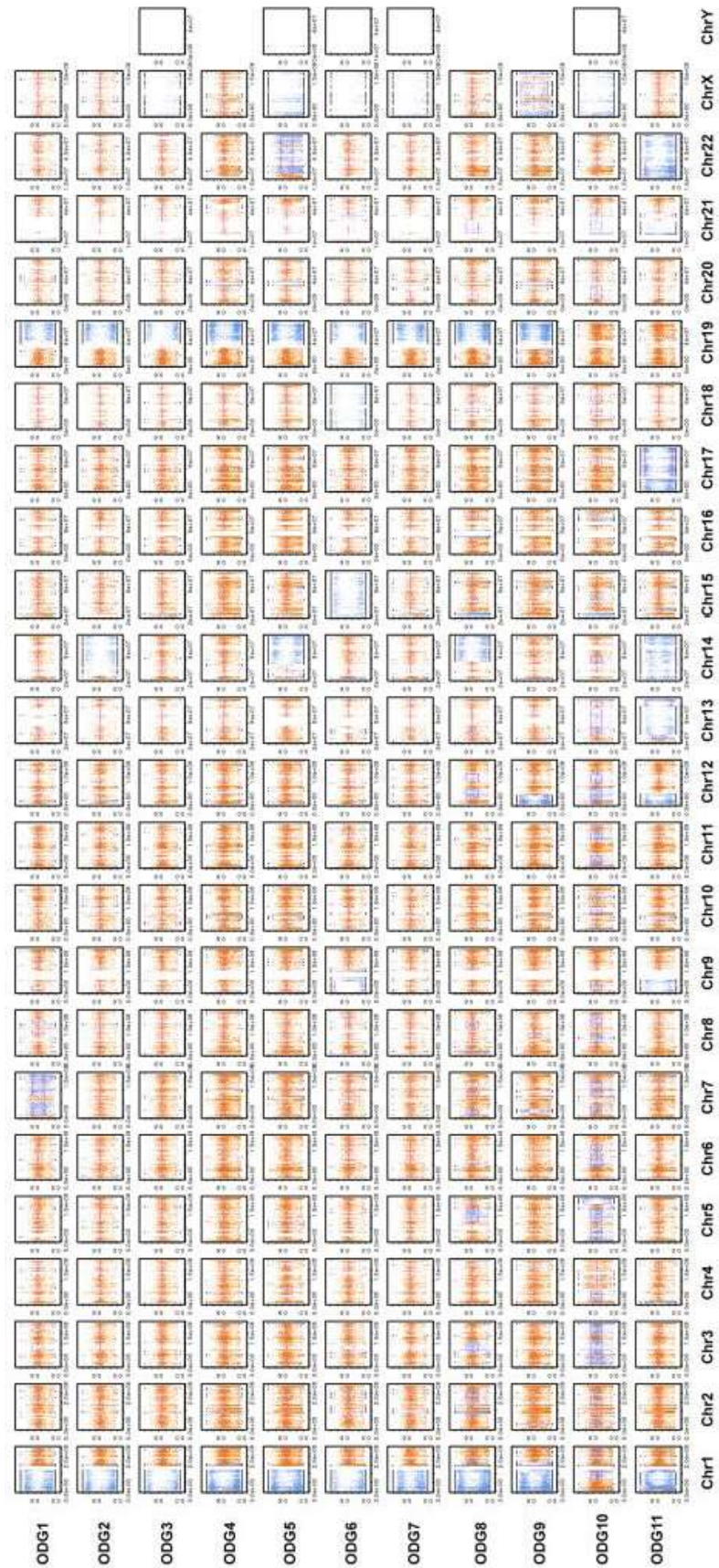


Figure 4.4: Copy number and allelic content profiles from exome data with predicted regions of genomic alteration such as gains, losses and LOH.

4.3.3 Somatic Mutations in Oligodendroglial Tumours.

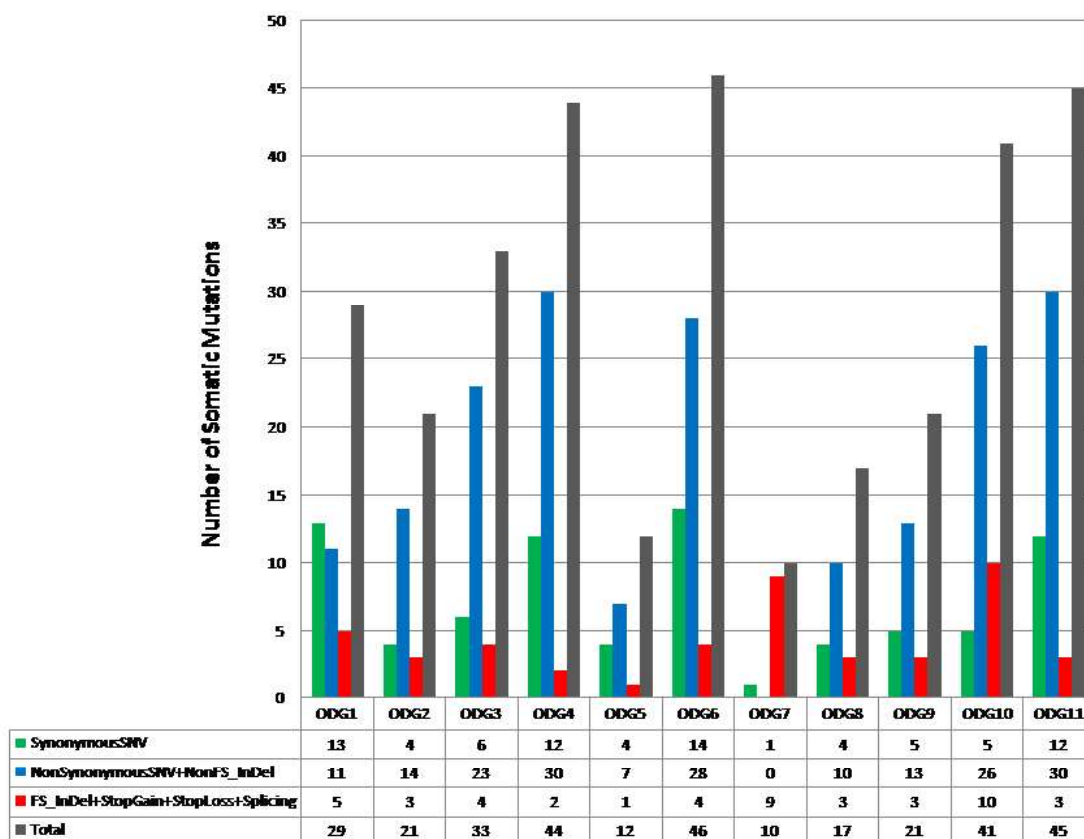


Figure 4.5: Bar chart showing different types of somatic mutations identified in the oligodendroglial tumour samples (ODG1 to ODG11) from the paired Exome sequencing data.

The Exome raw reads for each sample and paired blood were aligned to human reference genome hg19 using BWA aligner [125]. Somatic single nucleotide variants (SNVs) and insertions and deletions (indels) were identified using the VarScan [126] variant detection tool version 2.3.5 (<http://varscan.sourceforge.net>) using the filtering criteria of a minimum coverage 10 and at least 5 somatic variants. Functional annotation of the somatic variant list was done

using the ANNOVAR software (www.openbioinformatics.org/annovar)[127]. From the ANNOVAR-annotated list, variants located in segmental duplications were excluded. The remaining variants were manually verified in IGV (www.broadinstitute.org/igv). Ambiguous variants (variants represented in reads with low mapping quality, variants present near indels, and variants surrounded by mismatched bases) were discarded.

The number of non-synonymous somatic mutations per tumor exome ranged from 10 to 46 (Figure 4.5). R132H/R132C or R172K mutations in the *IDH1* or *IDH2* gene, respectively, were identified in 10 tumors while one tumor lacked mutation in the *IDH1* as well as *IDH2* gene (Figure 4.1). Four of the 9 tumors with chromosome 1p/19q codeletion were found to carry a missense or a frame-shift deletion mutation in the *CIC* gene, located on 19q while two tumors carried mutations in *FUBP1* gene located on 1p. Two tumors with chromosome 1p/19q codeletion, but no somatic alteration in the *CIC* gene were found to carry an activating mutation (Q61L, G12D) in the *KRAS* gene. Recurrent mutations were identified in the Notch signaling pathway genes, including four tumors with mutation in *NOTCH1* and one tumor with a mutation in *MAML3*. Two tumors were found to carry a mutation in the chromatin modifier *ARID1A* gene (Table 4.4). The tumor ODG10 lacking chromosome 1p/19q codeletion carried mutations in the *ATRX*, *TP53*, and *IDH1* genes. ODG11 lacked chromosome 1p/19q codeletion as well as a mutation in *IDH1/IDH2*. This tumor was found to carry a frame-shift deletion in the *NF1* gene, amplification of the *PDGFRA* gene, and deletion of chromosome 9p arm including the *CDKN2A* gene locus. The complete list of all 319 somatic nucleotide alterations in the 11 oligodendroglial tumour-normal paired exomes studied is given in Appendix IV.

Table 4.4: The list of the selected somatic alterations in the 11 oligodendroglial tumours identified by the VarScan analysis of the exome sequence data. (Reference Assembly hg19)

Sample	Gene	Exonic Function	Transcript:Exon: NucleotideChange: AminoAcid Change	Chr	Start	End	Ref	Obs	R1	R2	R3	R4
ODG10	<i>ARID1A</i>	frameshift deletion	NM_006015:exon20: c.5684delA:p.E1895fs	chr1	27106073	27106073	A	-	74	0	39	11
ODG5	<i>ARID1A</i>	frameshift deletion	NM_006015:exon20: c.5577_5578del:p.1859 _1860del	chr1	27105966	27105967	CG	-	44	1	13	5
ODG10	<i>ATRX</i>	frameshift deletion	NM_138270:exon25: c.5953_5957del:p.1985 _1986del	chrX	76849205	76849209	AAG AG	-	13	0	3	7
ODG11	<i>CDKN2C</i>	stopgain SNV	NM_078626:exon2: c.A163T:p.R55X	chr1	51439598	51439598	A	T	34	0	12	19
ODG1	<i>CIC</i>	nonsynonymous SNV	NM_015125:exon5: c.G692A:p.S231N	chr19	42791806	42791806	G	A	78	0	16	23
ODG2	<i>CIC</i>	frameshift deletion	NM_015125:exon3: c.342delG:p.L114fs	chr19	42791282	42791282	G	-	33	0	18	6
ODG2	<i>CIC</i>	frameshift deletion	NM_015125:exon2: c.150_151del:p.50_51del	chr19	42791005	42791006	CC	-	43	0	22	8
ODG3	<i>CIC</i>	nonsynonymous SNV	NM_015125:exon5: c.C643T:p.R215W	chr19	42791757	42791757	C	T	77	0	1	36
ODG4	<i>CIC</i>	nonsynonymous SNV	NM_015125:exon5: c.A614T:p.N205I	chr19	42791728	42791728	A	T	269	1	62	78
ODG1	<i>FUBP1</i>	stopgain SNV	NM_003902:exon16:	chr1	78425915	78425915	C	T	35	0	7	19

R1- Reference Genome-supporting reads in normal/ blood; R2- Variant-supporting reads in normal/blood; R3- Reference Genome-supporting reads in tumour; R4- Variant-supporting reads in tumour.

Sample	Gene	Exonic Function	Transcript:Exon: NucleotideChange: AminoAcid Change	Chr	Start	End	Ref	Obs	R1	R2	R3	R4
			c.G1530A:p.W510X									
ODG5	<i>FUBP1</i>	nonframeshift deletion	NM_003902:exon3: c.248_250del:p.83_84del	chr1	78433848	78433851	CAGT	-	46	0	17	9
ODG1	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	7	0	3	6
ODG10	<i>IDH1</i>	nonsynonymous SNV	INM_005896:exon4: c.C394T:p.R132C	chr2	209113113	209113113	G	A	11	0	5	5
ODG2	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	9	0	6	5
ODG3	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	12	0	3	10
ODG4	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	13	0	15	5
ODG5	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	19	0	11	10
ODG6	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	9	0	4	8
ODG7	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	7	0	8	1
ODG8	<i>IDH1</i>	nonsynonymous SNV	NM_005896:exon4: c.G395A:p.R132H	chr2	209113112	209113112	C	T	8	0	9	5
ODG9	<i>IDH2</i>	nonsynonymous SNV	NM_002168:exon4: c.G515A:p.R172K	chr15	90631838	90631838	C	T	55	0	9	19

R1- Reference Genome-supporting reads in normal/ blood; R2- Variant-supporting reads in normal/blood; R3- Reference Genome-supporting reads in tumour; R4- Variant-supporting reads in tumour.

Sample	Gene	Exonic Function	Transcript:Exon: NucleotideChange: AminoAcid Change	Chr	Start	End	Ref	Obs	R1	R2	R3	R4
ODG8	<i>KRAS</i>	nonsynonymous SNV	NM_004985:exon3: c.A182T:p.Q61L	chr12	25380276	25380276	T	A	87	1	58	30
ODG9	<i>KRAS</i>	nonsynonymous SNV	NM_004985:exon2 :c.G35A:p.G12D	chr12	25398284	25398284	C	T	45	0	20	44
ODG8	<i>MAML3</i>	frameshift deletion	NM_018717:exon3: c.1513_1514del: p.505_505del	chr4	140811064	140811075	TGCT GCTG CTGC	-	68	0	40	57
ODG11	<i>NF1</i>	frameshift deletion	NM_000267:exon8: c.797delT:p.V266fs	chr17	29509592	29509592	T	-	21	0	5	43
ODG2	<i>NOTCH1</i>	nonframeshift deletion	NM_017617:exon6: c.1070_1072del: p.357_358del	chr9	139413070	139413072	AGA	-	32	0	28	15
ODG3	<i>NOTCH1</i>	nonframeshift deletion	NM_017617:exon6: c.1070_1072del: p.357_358del	chr9	139413070	139413072	AGA	-	31	0	29	8
ODG4	<i>NOTCH1</i>	nonsynonymous SNV	NM_017617:exon27: c.G5110A:p.G1704R	chr9	139397691	139397691	C	T	100	0	71	38
ODG6	<i>NOTCH1</i>	nonsynonymous SNV	NM_017617:exon7: c.A1163G:p.D388G	chr9	139412681	139412681	T	C	22	0	7	5
ODG10	<i>TP53</i>	frameshift deletion	:NM_001126118:exon3: c.248_249del:p.83_83del	chr17	7579321	7579322	CA	-	16	0	8	5
ODG10	<i>TP53</i>	stopgain SNV	NM_001126115:exon5: c.A562T:p.K188X	chr17	7576888	7576888	T	A	56	0	19	13

R1- Reference Genome-supporting reads in normal/ blood; R2- Variant-supporting reads in normal/blood; R3- Reference Genome-supporting reads in tumour; R4- Variant-supporting reads in tumour.

R1 - Reference Genome-supporting reads in normal/ blood; R2- Variant-supporting reads in normal/blood; R3- Reference Genome-supporting reads in tumour; R4- Variant-supporting reads in tumour.

4.3.4 Sanger Sequencing of *CIC* Exon 5 from Mutated Samples.

Tumour samples ODG1, ODG3 and ODG4 had nonsynonymous mutation in exon 5 of the *CIC* gene. The mutations were validated by sanger sequencing the exon 5 of the *CIC* gene (Figure 4.6)

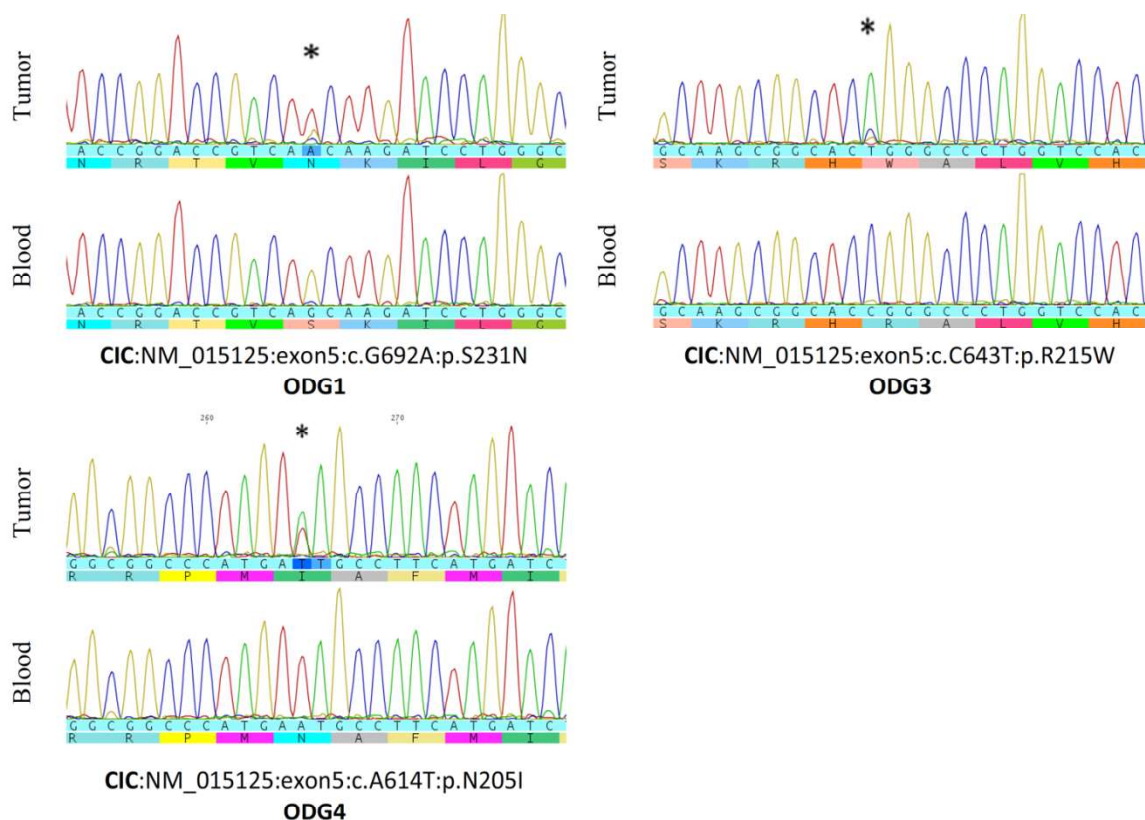


Figure 4.6: Electropherograms of the Sanger sequencing of *CIC* exon 5 from three mutated samples. Mutated nucleotides are indicated with *.

The tumour ODG1 had amino acid change S231N due to G692A mutation. The tumour ODG3 had R215W amino acid change due to C643T mutation and the tumour ODG4 had amino acid change N205I due to mutation A614T.

4.4 Comparative Transcriptome Analysis.

4.4. 1 Transcriptome Sequencing of Oligodendroglial Tumour Samples and Comparative Transcriptome Analysis.

From the Exome sequencing of eleven paired tumour-blood samples of oligodendroglioma morphology, nine tumours were found to have *IDH1/2* mutation and chromosome 1p/19q codeletion. Of these nine tumour tissues (ODG1 to ODG9) four tumors were found to carry *CIC* mutation (ODG1, ODG2, ODG3 and ODG4). Two out of the remaining five tumors though were wild type for *CIC*, carried activating *KRAS* mutations (ODG8 (Q61L) and ODG9 (G12D)). *CIC* gene located on chromosome arm 19q encodes a transcriptional repressor protein. To identify genes transcriptionally regulated by *CIC*, genes significantly differentially expressed in the *CIC*-mutant oligodendrogliomas as compared to that in the *CIC*-wild type oligodendrogliomas were identified by comparing gene expression profiles of the *CIC* mutant tumors with that of the *CIC* wild type

Total RNA was extracted from the tumor tissues after ensuring at least 80% tumor cell content. RNA was extracted using RNeasy plus mini kit as per the manufacturer's protocol (Qiagen). The RNA was quantified in a Qubit 2.0 fluorimeter (Life Technologies, Carlsbad, CA).

Single indexed RNA libraries were prepared using the Truseq RNA sample prep kit V2 (Catalogue No. RS-122-2001, Illumina) using 4 µg of total RNA as per the manufacturer's protocol. The kit specifically purifies poly-A mRNAs from total RNA using oligo-dT probes. The multiplexed RNA libraries were subjected to 150 bp paired end deep sequencing using the HiSeq 1500 ultra-high-throughput sequencing system (Illumina). The RNA sequencing was done to obtain a minimum of 20 million reads per sample.

4.4.2 Transcriptome Sequencing Data Quality.

Transcriptomes of nine oligodendroglial tumour samples carrying chromosome 1p/19q codeletion were sequenced. The mean and median Phred base quality of most of the bases was found to be above 30. Mapped reads are those that were aligned. Total mapped reads per sample ranged from 14.8 million to 22.9 million. Unique mapped are both aligned as well as non-duplicate reads. Unique read mapping rate ranged from 41.8% to 88.1%. rRNA reads are non-duplicate and duplicate reads aligning to rRNA regions. The rRNA content was minimal ranging from 0.011 to 0.096. The sequence data quality is given Appendix-III. The sequence data statistics has been described in detail in Table 4.5 and Table 4.6.

Transcript associated read statistics: The fraction of reads that mapped within genes (within introns or exons) called as intragenic read mapping rate ranged from 96.1% to 97.9%. The reads mapping within exonic region ranged from 86.4% to 91.3% while the intronic read mapping rate was found to be 6.6% to 10.2%. Overall gene expression profiling efficiency ranged from 86.4% to 91.3%. Detailed transcript associated read statistics is given in Table 4.6.

Mean coverage for high expressed transcripts ranged from 125 to >1750, mean coverage for medium expressed transcripts ranged from 2 to 70 and for Low expressed transcripts ranged from 1 to 22. The coverage of the transcripts across the length of the transcript was found to be uniform except towards 3' end of the medium and low expressed transcripts (Figure 4.7).

Table 4.5: RNA seq Data Statistics

Sample	Total Purity Filtered Reads Sequenced	Alternative Alignments	Read Length	Estimated Library Size	Mapped	Unique Mapped Reads	Unique Rate of Mapped	rRNA Reads	rRNA rate
ODG1	20,350,593	1,546,599	151	48,220,099	20,350,593	17,897,333	0.879	316,324	0.016
ODG2	19,600,652	1,808,362	151	6,821,626	19,600,652	10,435,395	0.532	1,448,311	0.074
ODG3	20,806,600	1,890,462	151	10,347,221	20,806,600	12,967,628	0.623	1,333,003	0.064
ODG4	22,924,168	1,732,945	151	26,880,022	22,924,168	18,291,115	0.798	501,653	0.022
ODG5	19,343,844	1,779,592	151	4,541,868	19,343,844	8,080,209	0.418	1,849,294	0.096
ODG6	14,836,944	1,188,961	151	6,456,798	14,836,944	8,767,719	0.591	1,083,058	0.073
ODG7	21,824,823	1,800,832	151	7,620,562	21,824,823	11,569,980	0.530	1,716,215	0.079
ODG8	16,656,707	1,162,959	101	35,597,102	16,656,707	14,637,575	0.879	181,030	0.011
ODG9	17,001,751	1,397,519	101	37,556,873	17,001,751	14,971,209	0.881	218,893	0.013

Table 4.6: RNA seq Data Statistics (Transcript-associated Reads)

Sample	Intragenic Rate	Exonic Rate	Intronic Rate	Intergenic Rate	Split Reads	Expression Profiling Efficiency	Transcripts Detected	Genes Detected
ODG1	0.961	0.874	0.086	0.037	5,688,120	0.874	110,350	20,194
ODG2	0.979	0.913	0.066	0.020	2,704,573	0.913	94,139	17,537
ODG3	0.965	0.878	0.087	0.035	2,300,190	0.878	87,811	18,419
ODG4	0.969	0.891	0.078	0.030	5,377,150	0.891	107,800	19,662
ODG5	0.977	0.887	0.090	0.023	1,058,580	0.887	83,154	17,291
ODG6	0.971	0.903	0.069	0.028	1,582,694	0.903	88,328	17,687
ODG7	0.974	0.898	0.076	0.025	2,082,771	0.898	93,784	18,606
ODG8	0.962	0.864	0.099	0.036	3,545,867	0.864	107,025	19,415
ODG9	0.966	0.864	0.102	0.033	3,684,009	0.864	109,032	19,851

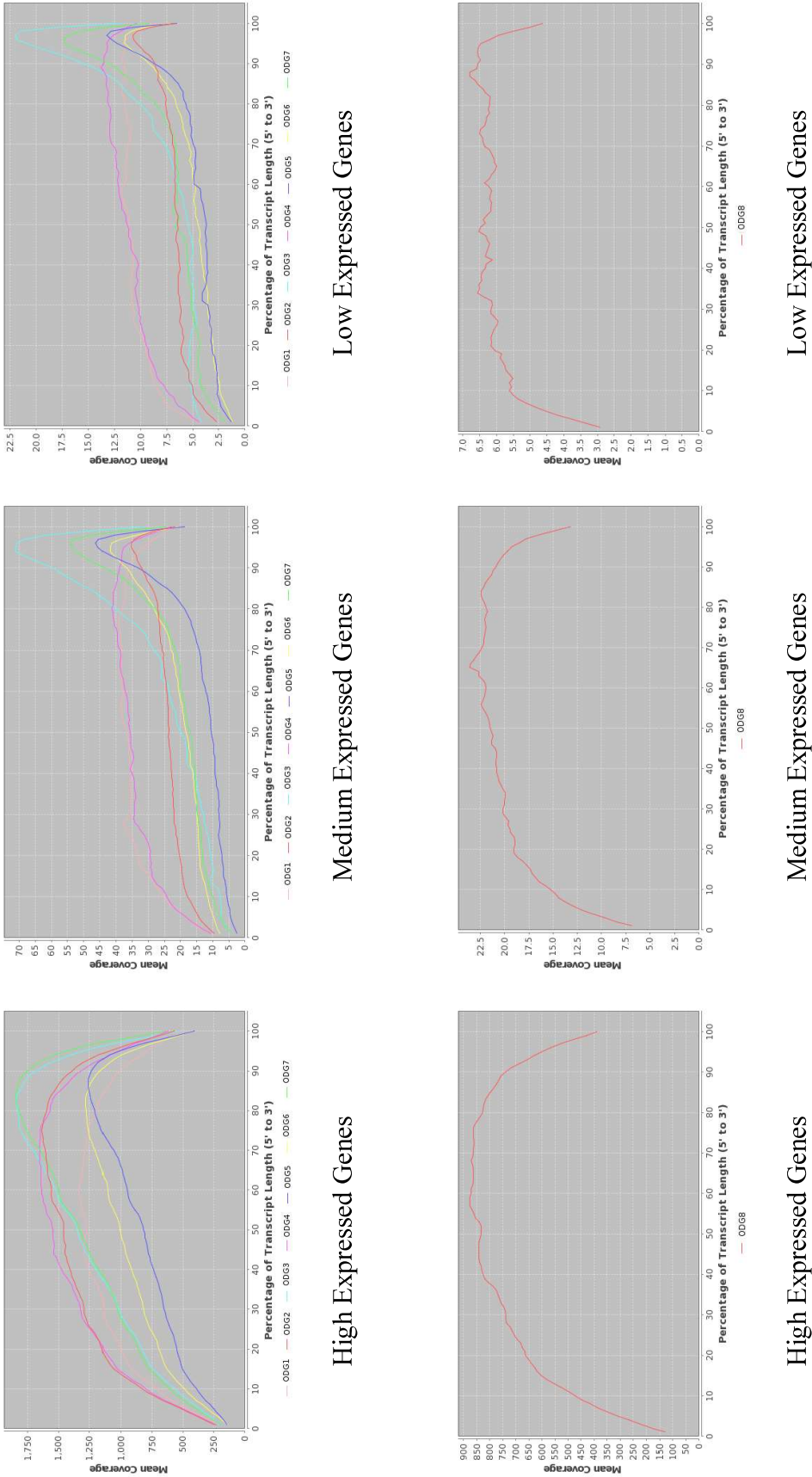


Figure 4.7: Coverage of High, Medium and Low expressed transcripts across samples. The X axis describes percentage of transcript length (5' to 3') and the Y axis shows the mean coverage.

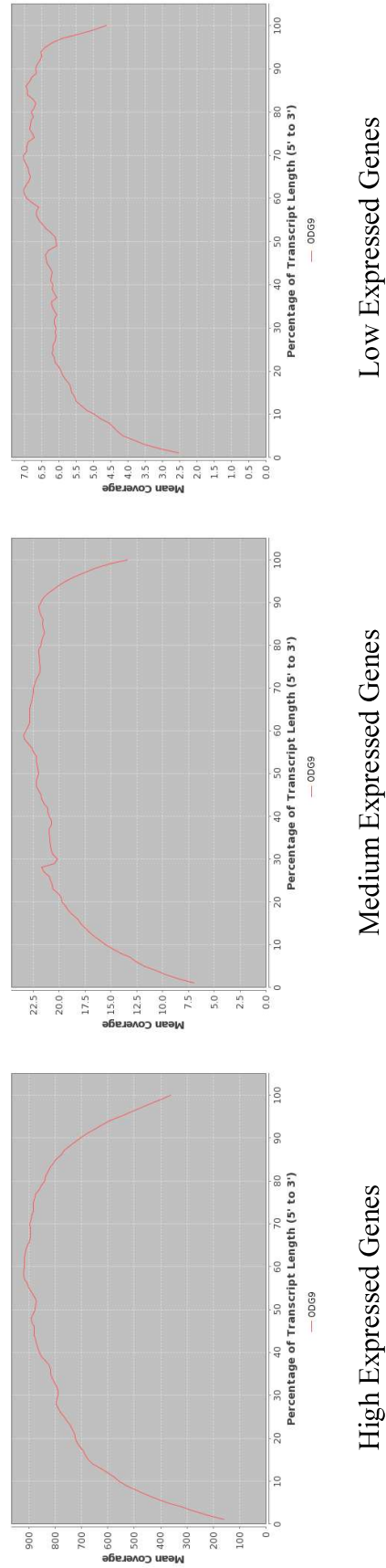


Figure 4.7 (Continued): Coverage of High, Medium and Low expressed transcripts in sample ODG9. The X axis describes percentage of transcript length (5' to 3') and the Y axis shows the mean coverage.

4.4.2.1 Differential Gene Expression Analysis Between *CIC*-mutated and *CIC*-wild type Oligodendroglial Tumour Samples from present study cohort Using EdgeR.

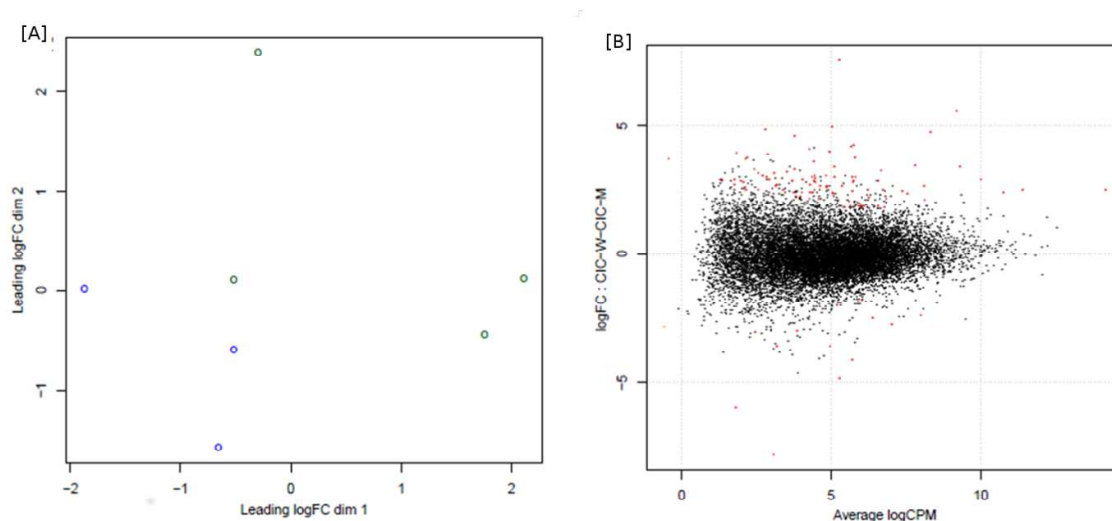


Figure 4.8: RNAseq differential gene expression plots of sample relations and MA plots – For present study (ACTREC) cohort by EdgeR. [A] Multidimensional scaling plot showing the relationship between all pairs of samples. Sample relations created by edgeR using a count-specific distance measure. Each sample is represented by ‘o’ [Green –*CIC*-mutated (n=4), Blue-*CIC*-wild type (n=3)] [B] Plot shows the log-fold change (i.e., the log ratio of normalized expression levels between two experimental conditions *CIC*-mutant and *CIC*-wild type) against the log counts per million (CPM).

From the Exome sequencing of eleven paired tumour-blood samples of the tumours of oligodendroglioma morphology, nine tumours were found to have *IDH1/2* mutation and 1p/19q codeletion. Of these nine *IDH1/2* mutated, 1p/19q codeleted tumour samples four samples were found to be *CIC* mutant (ODG1, ODG2, ODG3 and ODG4). The other five samples though were wild type for *CIC*, two of them had activating *KRAS* mutations (ODG8 (Q61L) and ODG9 (G12D)). *CIC* gene located on chromosome arm 19q encodes a transcriptional repressor protein. Capicua (*CIC*) is a key sensor of RTK signaling in both *Drosophila* and mammals. *CIC* functions as a repressor of RTK-responsive genes. *CIC* keeps RTK-responsive genes silent in the absence of RTK signaling. Following the activation of RTK signaling, *CIC* repression is relieved, and this allows the expression of the *CIC* repression target genes [109], [128]. To identify possible *CIC* repression target genes, differential gene expression analysis was carried out between two groups *CIC* mutated samples (ODG1, ODG2, ODG3 and ODG4) and *CIC* wild type samples (ODG5, ODG6 and ODG7) using R Bioconductor package EdgeR [129]. Samples in both groups had *IDH1/2* mutation and 1p-19q codeletion. Samples with activating *KRAS* mutations were excluded from the differential gene expression analysis.

Figure 4.8 shows quality of the differential gene expression analysis. Figure [A] shows multidimensional scaling plot of the relationship of all samples. The *CIC*-mutant vs *CIC*-wild type samples do not segregate in different groups indicating that all these tumors have considerable similarity since all of them have same background alterations i.e. *IDH1* mutation and 1p/19q codeletion. This also indicates that there is no batch effect that may result from processing of samples in batches. Figure [B] shows MA-plot, i.e. a scatter plot of logarithmic fold changes (on the y-axis) versus the mean of normalized counts (on the x-axis). Majority of the points (representing genes) are centered around log ratio of 0 indicating that the normalization across the samples is done appropriately.

The differential gene expression analysis identified 106 genes to be significantly differentially expressed at 5% false discovery rate. Potentially biologically significant genes that were identified to be significantly differentially expressed were *PDGFRA* (log fold change= -2.729123852, FDR 1.84E-02), *VEGFA* (log fold change= -4.130574787, FDR 2.16E-02), *ETV1* (log fold change= -2.401391061, FDR 3.81E-02), *SPRED2* (log fold change= -1.993761919, FDR 4.05E-02), *SPRED1* (log fold change= -1.87231971, FDR 4.14E-02). Detailed significantly differentially expressed genes list is given in Table 4.7 in Appendix VI.

4.4.2.2 Differential Gene Expression Analysis Between *CIC*-mutated and *CIC*-wild type Oligodendroglial Tumour Samples From TCGA Using EdgeR.

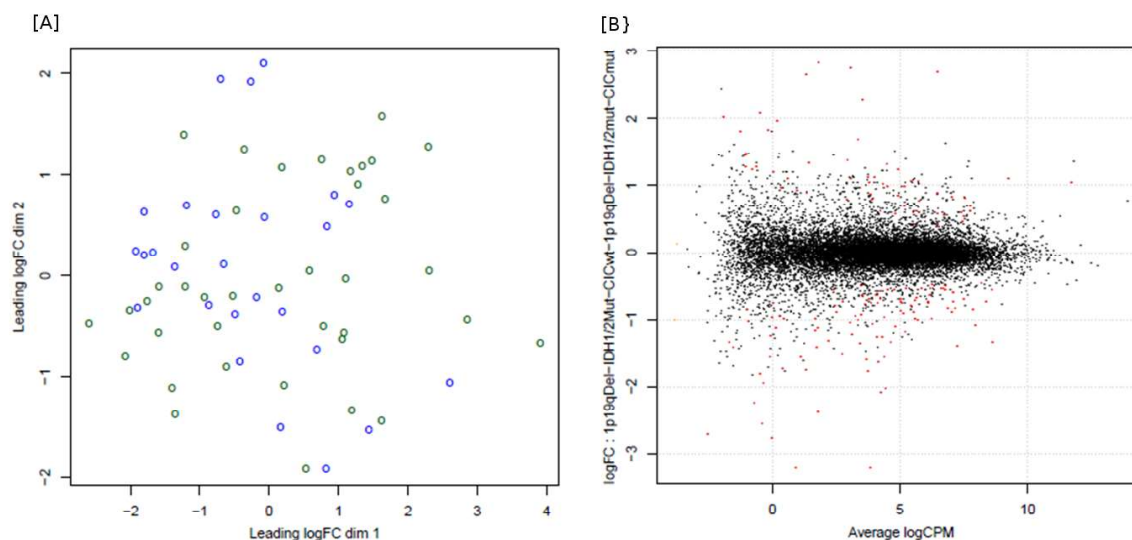


Figure 4.9: RNAseq differential gene expression plot of sample relations and MA plot for TCGA cohort by EdgeR. [A] Multidimensional scaling plot showing the relationship between all pairs of samples. Sample relations created by edgeR using a count-specific distance measure. Each sample is represented by 'o' [Green –*CIC*-mutated (n=39), Blue-*CIC*-wild type (n=26)] [B] MA Plot shows the log-fold change (i.e., the log ratio of normalized expression levels between two experimental conditions *CIC*-mutant and *CIC*-wild type) against the log counts per million (CPM).

Multiplatform genomic data for low grade glioma including oligodendroglioma is available from The Cancer Genome Atlas (TCGA) data portal [<https://tcga-data.nci.nih.gov/>]. Data for total of 65 samples for which somatic mutation, copy number alteration and mRNA seq data was available, were downloaded from the TCGA data portal. All 65 samples possessed *IDH1/2* mutation and chromosome 1p/19q codeletion. Of the 65 samples 39 were mutant for *CIC* gene while 26 were wild type for the *CIC* gene. This data was used as an independent dataset to confirm the differential gene expression results obtained from the study data cohort of seven samples (*CIC* mutant n=4 and *CIC* wild type n=3). The TCGA data was grouped into *CIC*

mutant and *CIC* wild type and differential gene expression analysis was carried out using R Bioconductor package EdgeR [129].

Figure 4.9 shows quality of the differential gene expression analysis. Figure [A] shows multidimensional scaling plot of the relationship of all samples. The *CIC*-mutant vs *CIC*-wild type samples do not segregate in different groups indicating that all these tumors have considerable similarity since all of them have same background alterations i.e. *IDH1* mutation and 1p/19q codeletion. This also indicates that there is no batch effect that may result from processing of samples in batches. Figure [B] shows MA-plot, i.e. a scatter plot of logarithmic fold changes (on the y-axis) versus the mean of normalized counts (on the x-axis). Majority of the points (representing genes) are centered around log ratio of 0 indicating that the normalization across the samples is done appropriately.

A total of 158 genes were found to be significantly differentially expressed at 5% false discovery rate from the differential expression analysis between *CIC* mutated (n=39) and *CIC* wild type (n=26) samples from TCGA cohort using EdgeR. *ETV1*, *SPRED2* and *SPRED1*, the genes that were found to be significantly differentially expressed in the differential expression analysis between *CIC* mutant and *CIC* wild type samples from the study cohort, were also found to be significantly differentially expressed between the groups in the TCGA cohort. [*ETV1* (log fold change= -1.34139, FDR 2.19E-04), *SPRED2* (log fold change= -0.75923, FDR 1.33E-03), *SPRED1* (log fold change= -0.55917, FDR 2.89E-02)]. Additionally two more ETV/PEA3 family transcription factor-encoding genes *ETV5* and *ETV4* were also found to be significantly differentially expressed. [*ETV5* (log fold change= -1.38861, FDR 2.00E-04), *ETV4* (log fold change= -2.02452, FDR 3.54E-03)]. Other biologically significant genes that were found to be statistically significantly differentially expressed were *SPRY4* (log fold change= -1.32875, FDR 4.55E-05), *DUSP6* (log fold change= -0.99789, FDR 9.16E-04), *DUSP4* (log fold change = -1.50918, FDR 1.99E-03), *ALK* (log fold change= -1.00131, FDR 1.74E-02), *SHC3* (log fold

change= -1.1737, FDR 1.74E-02). *SPRY4*, *SPRED1*, *SPRED2*, *DUSP4*, *DUSP6* are the negative regulators of the tyrosine kinase receptor signaling pathway. *ALK* is a receptor tyrosine kinase encoding gene. These genes were found to be upregulated in *CIC* mutant samples. CREB3L1, a member of the CREB/ATF family transcription factors that modulates unfolded protein response signaling, was also found to be upregulated in the *CIC*-mutant oligodendrogliomas. Detailed significantly differentially expressed genes list is given in Table 4.8 in Appendix VI.

4.4.2.3 Differential Gene Expression Analysis Between *CIC*-mutated and *CIC*-wild type Oligodendrglial Tumour Samples from TCGA Using DEseq.

The data downloaded from TCGA data portal of 65 low grade glioma samples harboring *IDH1/2* mutation and chromosome 1p/19q codeletion was also used to carry out differential gene expression analysis between groups *CIC* mutant (n=39) and *CIC* wild type (n=26) using R Bioconductor package DEseq [130].

Figure 4.10 shows quality of the differential gene expression analysis. Figure [A] shows MA-plot, i.e. a scatter plot of logarithmic fold changes (on the y-axis) versus the mean of normalized counts (on the x-axis). Majority of the points (representing genes) are centered around log ratio of 0 indicating that the normalization across the samples is done appropriately. Figure [B] shows a principal component (PC) plot of VST (variance-stabilizing transformation)-transformed count data. The *CIC*-mutant vs *CIC*-wild type samples do not segregate in different groups indicating that all these tumors may have considerable similarity since all of them have same background alterations i.e. *IDH1* mutation and 1p/19q codeletion. This also indicates that there is no batch effect that may result from processing of samples in batches. Figure [C] shows the typical features of a P value histogram resulting from a good data set: a sharp peak at the left side, containing genes with strong differential expression, a ‘floor’ of values that are approximately

uniform in the interval $[0, 1]$, corresponding to genes that are not differentially expressed, and a peak at the upper end at 1.

A total of 76 genes were found to be significantly differentially expressed at 10% false discovery rate from the differential expression analysis between *CIC* mutated and *CIC* wild type samples from TCGA cohort using DEseq. The biologically potentially significant genes which were identified using EdgeR were also found using DEseq. The genes included *ETV1* (log2 fold change=-1.39, FDR 2.58E-04), *SPRY4* (log2 fold change=-1.15, FDR 8.73E-04), *ETV5* (log2 fold change=-1.22, FDR 1.29E-03), *DUSP4* (log2 fold change=-1.45, FDR 3.06E-03), *ETV4* (log2 fold change=-1.84, FDR 3.36E-02), *SHC3* (log2 fold change=-1.20, FDR 3.95E-02), *SPRED2* (log2 fold change=-0.71, FDR 5.90E-02), *DUSP6* (log2 fold change=-0.81, FDR 6.07E-02) and *ALK* (log2 fold change=-0.98, FDR 6.99E-02). *CREB3L1* (log2 fold change=-1.65, FDR 6.20E-05), a member of the CREB/ATF family transcription factors that modulates unfolded protein response signaling was also found to be significantly differentially expressed. These genes were found to be upregulated in the *CIC*-mutant oligodendrogliomas. Detailed list of genes significantly differentially expressed is given in Table 4.9 in Appendix VI.

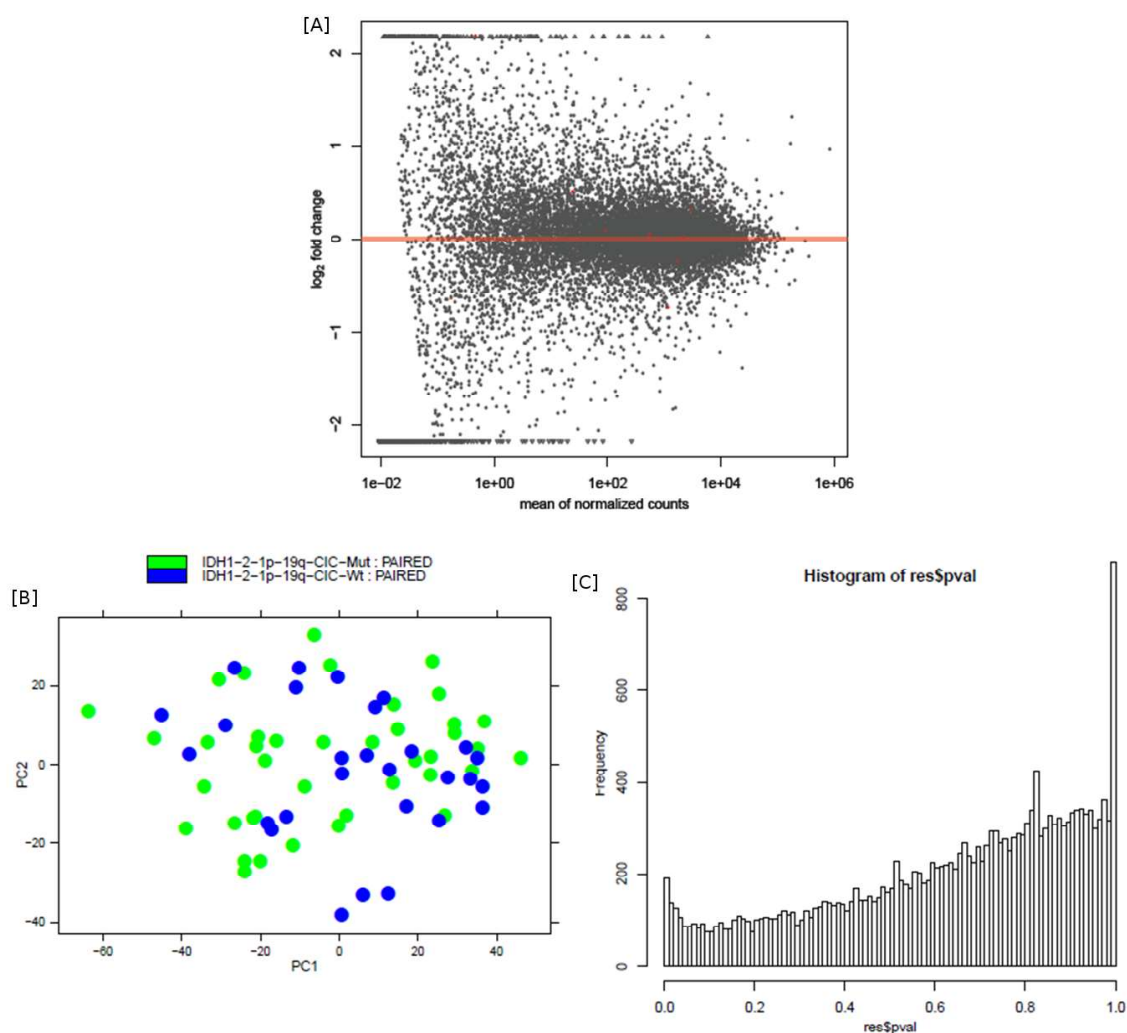


Figure 4.10: RNAseq differential gene expression MA plot, plot of sample relations and p-value histogram for TCGA cohort by DEseq. [A] MA Plot displays differential expression (log-fold changes) versus expression strength (log average read count). [B] A principal component (PC) plot of VST (variance-stabilizing transformation)-transformed count data showing sample relations. Each sample is represented by a dot [Green – *CIC*-mutated (n=39), Blue - *CIC*-wild type (n=26)] [C] Histogram of p values from gene-by-gene statistical tests for differential expression.

4.4.2.4 Differential Gene Expression Analysis Between *CIC*-mutated and *CIC*-wild type Oligodendroglial Tumour Samples Using SAM.

RNA-seq data on 65 oligodendrogliomas with 1p/19q codeletion from the TCGA cohort was analyzed for differential gene expression. SAM analysis identified 148 genes to be significantly differentially expressed in the 39 *CIC*-mutant oligodendrogliomas as compared with the 26 *CIC*-wild type oligodendrogliomas from the TCGA cohort at a False Discovery Rate of <5% (Figure 4.11 [A], Table 4.10, Appendix V - TCGA sample Information). The differential gene expression comparing the *CIC*-mutant and *CIC*-wild type oligodendrogliomas from our cohort as well as the TCGA cohort was also done using EdgeR analysis (Table 4.10). The genes identified to be significantly differentially expressed in the TCGA cohort showed differential expression in our cohort as well, although some genes did not reach statistical significance due to the small sample size (Figure 4.11 [B] Table 4.10). *ETV1*, *ETV4*, and *ETV5*, the three genes belonging to the ETS/PEA3 family of transcription factors, were found to be upregulated in the *CIC*-mutant tumors. The MAPK pathway genes upregulated in *CIC* mutant tumours included the dual specificity phosphatase genes *DUSP4*, *DUSP6*, and *DUSP19*, the Sprouty family members *SPRY4*, *SPRED1*, and *SPRED2*, and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, and *EPHB*. *CREB3L1*, a member of the CREB/ATF family transcription factors that modulates unfolded protein response signaling, was also found to be upregulated in the *CIC*-mutant oligodendrogliomas. Two oligodendrogliomas (ODG8 and ODG9) carried activating mutation in *KRAS* gene. These two cases had higher expression of some of the negative regulators of tyrosine kinase receptor signaling pathway such as *SPRY4*, *SPRED2*, and *DUSP6* (Figure 4.11 [B]). The TCGA data, however, contains only one tumor with an activating mutation in the *NRAS* gene out of the 65 oligodendrogliomas.

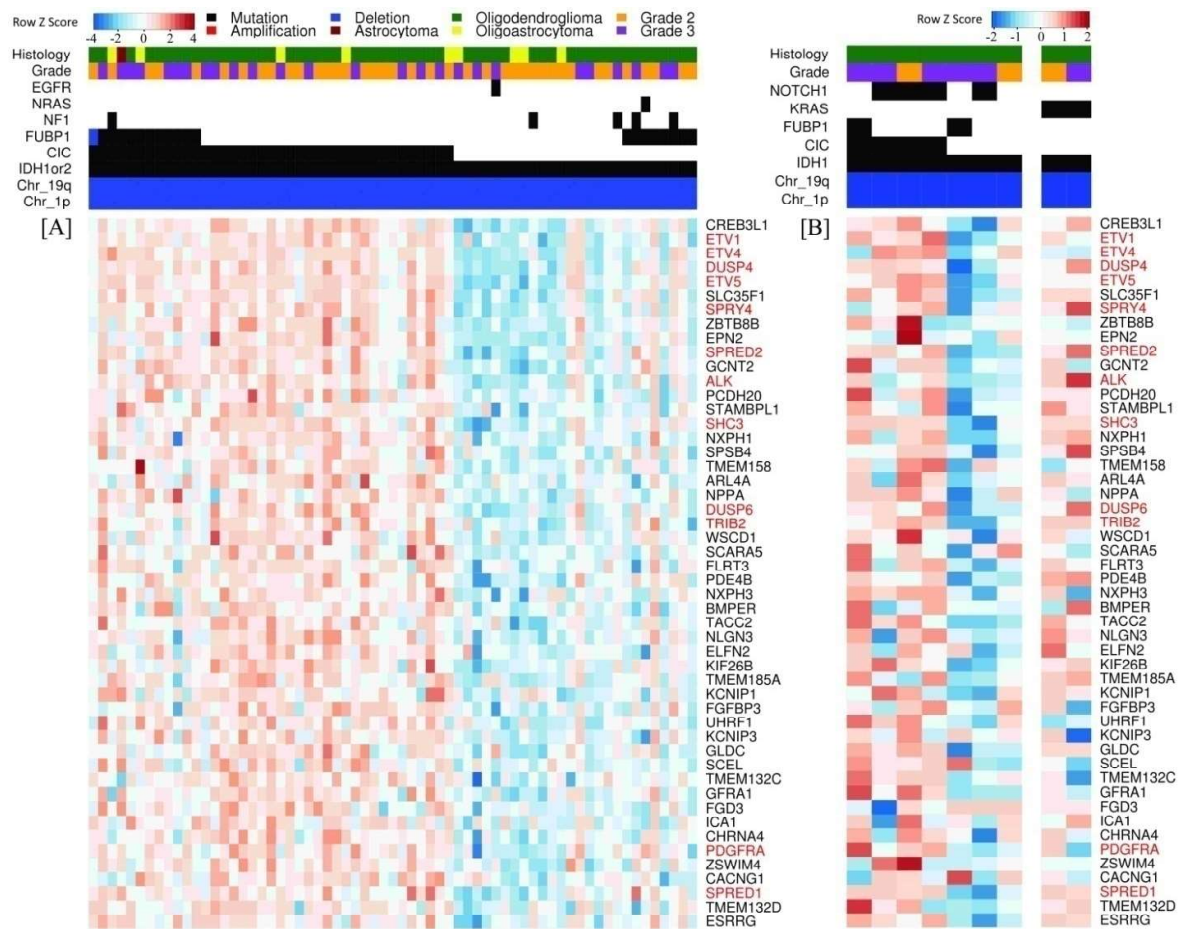


Figure 4.11: Heat map showing the top 50 genes most significantly differentially expressed in the 39 *CIC*-mutant vs. 26 *CIC*-wild type oligodendrogliomas from the TCGA cohort (A) and from the present study cohort of 9 oligodendrogliomas (B). The genes belonging to the MAPK signaling pathway are highlighted in red.

Table 4.10: List of the 148 genes significantly differentially expressed (FDR< 0.05) in the *CIC*-mutant as compared to the *CIC*-wild type, 1p/19q codeleted, *IDH1/IDH2*-mutant oligodendrogliomas from the TCGA cohort and the study cohort as identified by the SAM and EdgeR analysis. The genes known to play role in MAPK signaling are highlighted with dark background. NA: Not available; NS: Non-significant.

Sr No	GeneID	SAM analysis (FDR< 0.05) (TCGA cohort)			EdgeR analysis (TCGA cohort)		Edge R (Study cohort)	
		Expected score (dExp)	Observed score (d)	Fold change (Unlogged)	P value	FDR	P value	FDR
1	<i>CREB3L1</i>	-0.831	7.202	3.048	0.000	0.000	0.083	0.525
2	<i>ETV1</i>	-0.571	7.186	2.625	0.000	0.000	0.000	0.038
3	<i>ETV4</i>	-0.570	6.789	3.516	0.000	0.004	0.085	0.525
4	<i>DUSP4</i>	-0.646	6.392	2.619	0.000	0.002	0.003	0.141
5	<i>ETV5</i>	-0.570	6.381	2.326	0.000	0.000	0.008	0.223
6	<i>SLC35F1</i>	0.841	6.283	1.762	0.000	0.000	0.000	0.049
7	<i>SPRY4</i>	0.998	6.085	2.200	0.000	0.000	NS	-
8	<i>ZBTB8B</i>	1.775	5.772	2.485	0.000	0.000	NS	-
9	<i>EPN2</i>	-0.585	5.602	1.652	0.000	0.000	NS	-
10	<i>SPRED2</i>	0.994	5.587	1.632	0.000	0.001	0.000	0.041
11	<i>GCNT2</i>	-0.394	5.344	1.916	0.000	0.001	0.033	0.392
12	<i>ALK</i>	-1.922	5.304	1.934	0.000	0.017	0.008	0.224
13	<i>PCDH20</i>	0.343	5.273	2.995	0.000	0.000	0.000	0.042
14	<i>STAMBPL1</i>	1.027	5.165	1.375	0.000	0.007	NS	-
15	<i>NUDT9P1</i>	0.232	5.126	1.333	NA	NA	NA	NA
16	<i>SHC3</i>	0.790	5.006	2.285	0.000	0.017	0.001	0.098
17	<i>NXPH1</i>	0.238	4.932	2.100	0.000	0.015	0.001	0.081
18	<i>SPSB4</i>	0.999	4.931	1.694	0.000	0.011	NS	-
19	<i>TMEM158</i>	1.216	4.891	2.389	0.000	0.001	0.047	0.443
20	<i>ARL4A</i>	-1.676	4.864	2.137	0.000	0.001	0.025	0.361

Sr No	GeneID	SAM analysis (FDR< 0.05) (TCGA cohort)			EdgeR analysis (TCGA cohort)		Edge R (Study cohort)	
		Expected score (dExp)	Observed score (d)	Fold change (Unlogged)	P value	FDR	P value	FDR
21	<i>NPP4</i>	0.210	4.827	3.529	0.000	0.000	0.033	0.388
22	<i>DUSP6</i>	-0.646	4.784	1.743	0.000	0.001	0.001	0.103
23	<i>TRIB2</i>	1.327	4.774	1.556	0.000	0.031	0.010	0.254
24	<i>WSCD1</i>	1.691	4.751	1.843	0.000	0.002	NS	
25	<i>SCAR45</i>	0.720	4.745	4.488	0.000	0.000	NA	NA
26	<i>FLRT3</i>	-0.449	4.691	1.769	0.000	0.029	0.004	0.172
27	<i>PDE4B</i>	0.362	4.685	1.569	0.000	0.007	0.001	0.089
28	<i>NXPH3</i>	0.239	4.655	1.622	0.000	0.004	0.025	0.364
29	<i>BMPER</i>	-1.456	4.648	2.115	0.000	0.004	NS	
30	<i>LOC92659</i>	0.003	4.648	1.421	NA	NA	NA	NA
31	<i>TACC2</i>	1.081	4.600	1.444	0.000	0.013	0.002	0.106
32	<i>NLGN3</i>	0.193	4.596	1.407	0.000	0.017	0.097	0.547
33	<i>ELFN2</i>	-0.607	4.583	1.925	0.000	0.002	NS	
34	<i>KIF26B</i>	-0.101	4.550	2.360	0.000	0.001	0.003	0.150
35	<i>TMEM185A</i>	1.223	4.548	1.172	0.002	0.118	NS	0.601
36	<i>KCNIP1</i>	-0.136	4.498	1.780	0.000	0.010	0.076	0.513
37	<i>FGFBP3</i>	-0.471	4.481	1.906	0.000	0.013	NS	0.786
38	<i>UHRF1</i>	1.498	4.457	1.666	0.324	0.935	0.073	0.507
39	<i>KCNIP3</i>	-0.136	4.421	1.497	0.000	0.012	NS	0.994
40	<i>GLDC</i>	-0.373	4.408	1.433	0.001	0.057	0.001	0.065
41	<i>SCEL</i>	0.725	4.376	1.907	0.000	0.006	NA	
42	<i>TMEM132C</i>	1.209	4.337	1.638	0.000	0.037	NS	0.585
43	<i>GFR1</i>	-0.387	4.279	1.696	0.000	0.017	0.004	0.167
44	<i>FGD3</i>	-0.475	4.213	1.912	0.000	0.005	0.055	0.466
45	<i>MGC12982</i>	0.059	4.205	2.475	0.000	0.000	NA	
46	<i>ICAI</i>	-0.214	4.198	1.517	0.001	0.069	NS	0.988

Sr No	GeneID	SAM analysis (FDR < 0.05) (TCGA cohort)			EdgeR analysis (TCGA cohort)		Edge R (Study cohort)	
		Expected score (dExp)	Observed score (d)	Fold change (Unlogged)	P value	FDR	P value	FDR
47	CHRNA4	-0.922	4.171	1.736	0.000	0.007	0.265	0.715
48	PDGFRA	0.365	4.150	1.761	0.004	0.157	0.000	0.018
49	ZSWIM4	3.012	4.146	1.387	0.000	0.037	NS	0.759
50	CACNG1	-1.103	4.114	1.838	0.000	0.032	NS	0.736
51	SPRED1	0.994	4.085	1.438	0.000	0.029	0.000	0.041
52	TMEM132D	1.209	4.084	2.000	0.001	0.082	0.016	0.300
53	ESRRG	-0.573	4.076	1.499	0.001	0.069	NS	0.022
54	TMC3	1.194	4.074	3.226	0.000	0.005	NA	NA
55	ELFNI	-0.607	4.044	1.979	0.000	0.005	NS	0.895
56	LPPR5	0.006	4.032	1.552	0.000	0.025	0.016	0.300
57	ST3GAL5	1.020	4.017	1.417	0.000	0.034	NS	0.729
58	C8orf56	-1.135	3.985	1.945	0.000	0.017	NA	NA
59	KCND2	-0.139	3.965	1.670	0.001	0.059	0.001	0.106
60	ANKRD6	-1.819	3.932	1.267	0.002	0.102	NS	0.919
61	PPP1R14C	0.471	3.867	1.413	0.005	0.185	NS	0.497
62	SLC29A1	0.832	3.857	1.332	0.000	0.003	NS	0.578
63	BCL2	-1.490	3.851	1.361	0.005	0.181	NS	0.743
64	BACH2	-1.523	3.847	1.381	0.001	0.055	NS	0.876
65	LASP1	-0.058	3.841	1.219	0.002	0.104	NS	0.766
66	ALMSIP	-1.921	3.835	1.582	0.000	0.037	NA	NA
67	PDZD2	0.370	3.826	1.768	0.000	0.005	NS	0.544
68	DLL3	-0.692	3.824	1.914	0.000	0.004	NS	0.495
69	C3orf31	-1.198	3.778	1.154	0.051	0.548	NA	NA
70	VSIG10	1.600	3.777	1.286	0.001	0.082	NS	0.853
71	NEUROD1	0.177	3.768	1.769	0.001	0.057	NS	0.976
72	RNF216L	0.651	3.758	1.247	0.007	0.210	NA	NA

Sr No	GeneID	SAM analysis (FDR < 0.05) (TCGA cohort)			EdgeR analysis (TCGA cohort)		Edge R (Study cohort)	
		Expected score (dExp)	Observed score (d)	Fold change (Unlogged)	P value	FDR	P value	FDR
73	<i>FGFR1</i>	-0.470	3.757	1.345	0.000	0.006	0.038	0.407
74	<i>HS6ST2</i>	-0.231	3.749	1.597	0.000	0.039	NS	0.570
75	<i>SOX11</i>	0.964	3.748	1.610	0.006	0.205	0.026	0.364
76	<i>GPR123</i>	-0.342	3.724	1.409	0.000	0.032	NS	0.833
77	<i>MB</i>	0.039	3.714	1.768	0.002	0.126	NS	0.602
78	<i>STARD5</i>	1.029	3.704	1.471	0.000	0.017	NS	-
79	<i>NCAN</i>	0.149	3.691	1.417	0.004	0.166	NS	0.598
80	<i>CHST11</i>	-0.919	3.691	1.498	0.004	0.166	0.011	0.257
81	<i>TBC1D10A</i>	1.102	3.688	1.680	0.000	0.043	NS	0.652
82	<i>GLT25D2</i>	-0.369	3.677	1.492	0.001	0.069	NA	NA
83	<i>NRXN2</i>	0.221	3.660	1.308	0.003	0.139	NS	0.666
84	<i>C6orf118</i>	-1.168	3.622	1.639	0.000	0.026	NA	NA
85	<i>CSPG5</i>	-0.810	3.611	1.490	0.002	0.121	0.006	0.193
86	<i>ISM1</i>	-0.160	3.598	1.887	0.000	0.027	NS	0.820
87	<i>TMOD1</i>	1.257	3.589	1.364	0.000	0.023	NS	0.919
88	<i>LOC283392</i>	-0.016	3.588	1.663	0.000	0.031	NA	NA
89	<i>KCNK3</i>	-0.132	3.585	1.642	0.000	0.032	0.004	0.168
90	<i>FAM196A</i>	-0.527	3.584	1.796	0.001	0.089	NS	0.581
91	<i>DIAPH2</i>	-0.703	3.561	1.551	0.000	0.036	NS	0.770
92	<i>TRAF4</i>	1.315	3.552	1.411	0.000	0.039	NS	0.947
93	<i>ANKRD55</i>	-1.822	3.538	2.267	0.000	0.049	0.002	0.125
94	<i>COL20A1</i>	-0.864	3.526	1.987	0.001	0.077	NS	0.997
95	<i>IPO8</i>	-0.169	3.522	1.203	0.001	0.067	NS	0.801
96	<i>GRM5</i>	-0.318	3.512	1.868	0.001	0.081	NA	NA
97	<i>CDH3</i>	-0.980	3.505	1.898	0.000	0.034	NS	0.903
98	<i>THSD4</i>	1.167	3.493	1.942	0.000	0.028	NS	0.640

Sr No	GeneID	SAM analysis (FDR< 0.05) (TCGA cohort)				EdgeR analysis (TCGA cohort)		Edge R (Study cohort)	
		Expected score (dExp)	Observed score (d)	Fold change (Unlogged)	P value	FDR	P value	FDR	
99	WIP12	1.674	3.488	1.142	0.016	0.324	NS	0.995	
100	CDC47L	-0.984	3.484	1.671	0.001	0.075	0.006	0.195	
101	RAB31	0.560	3.471	1.358	0.001	0.097	NS	0.475	
102	DPP6	-0.664	3.470	1.270	0.003	0.130	NS	0.683	
103	FOXP4	-0.436	3.467	1.447	0.000	0.034	NS	0.670	
104	PRKG2	0.497	3.457	2.755	0.000	0.021	NS	0.876	
105	ABHD2	-2.510	3.456	1.411	0.004	0.166	0.000	0.047	
106	LBH	-0.056	3.453	1.376	0.001	0.065	0.008	0.226	
107	USP27X	1.532	3.451	1.229	0.010	0.256	Ns	0.891	
108	C2orf27A	-1.216	3.448	1.475	0.004	0.161	Ns	0.491	
109	GPR3	-0.333	3.430	1.346	0.005	0.186	NS	0.547	
110	SNCAIP	0.887	3.428	1.617	0.001	0.096	NS	0.937	
111	ACCN2	-2.398	3.427	1.444	0.001	0.085	NA	NA	
112	NRG1	0.218	3.426	2.144	0.000	0.015	NA	NA	
113	LMO1	-0.036	3.425	1.635	0.021	0.368	NS	0.696	
114	PCOLCE2	0.355	3.409	1.667	0.001	0.096	NS	0.830	
115	LRRN2	0.013	3.406	1.282	0.002	0.118	NS	0.537	
116	WNT7A	1.685	3.403	1.552	0.001	0.082	NS	0.535	
117	BAALC	-1.525	3.403	1.396	0.002	0.118	NS	0.571	
118	TRAPPC9	1.322	3.401	1.191	0.002	0.111	NS	0.732	
119	SATB2	0.716	3.375	1.350	0.002	0.102	NS	0.683	
120	CD82	-0.995	3.363	1.363	0.001	0.057	0.009	0.245	
121	SHROOM2	0.794	3.346	1.298	0.006	0.199	0.004	0.172	
122	BMP6	-1.457	3.345	1.697	0.002	0.123	NA	NA	
123	ROD1	0.658	3.341	1.266	0.001	0.069	NA	NA	
124	MFGES8	0.056	3.338	1.389	0.003	0.136	NS	-	

Sr No	GeneID	SAM analysis (FDR < 0.05) (TCGA cohort)			EdgeR analysis (TCGA cohort)		Edge R (Study cohort)	
		Expected score (dExp)	Observed score (d)	Fold change (Unlogged)	P value	FDR	P value	FDR
125	<i>CA10</i>	-1.113	3.330	1.554	0.009	0.247	NS	-
126	<i>MMD2</i>	0.072	3.314	1.604	0.004	0.166	NS	-
127	<i>STK17A</i>	1.035	3.311	1.249	0.001	0.073	NS	-
128	<i>STC2</i>	1.033	3.306	1.752	0.001	0.099	NS	-
129	<i>RHOBTB2</i>	0.627	3.298	1.282	0.005	0.186	NS	-
130	<i>FAM5C</i>	-0.515	3.286	1.517	0.001	0.082	NA	NA
131	<i>GPSM2</i>	-0.327	3.279	1.241	0.007	0.209	NS	-
132	<i>SF3A1</i>	0.769	3.277	1.193	0.014	0.304	0.033	0.388
133	<i>MRPS7</i>	0.098	3.276	1.157	0.003	0.136	NS	-
134	<i>GALR1</i>	-0.405	3.266	1.636	0.040	0.499	NS	-
135	<i>HUNK</i>	-0.217	3.263	1.343	0.002	0.103	NS	-
136	<i>RNF216</i>	0.651	3.257	1.156	0.019	0.352	NS	-
137	<i>CPVL</i>	-0.834	3.254	1.392	0.000	0.037	NS	-
138	<i>RAI2</i>	0.574	3.254	1.239	0.014	0.302	NS	-
139	<i>ELP3</i>	-0.603	3.248	1.128	0.042	0.505	NS	-
140	<i>ADAMTS2</i>	-2.175	3.248	1.719	0.002	0.118	NS	-
141	<i>PIWIL3</i>	0.408	3.235	1.386	NA	NA	NA	NA
142	<i>PANK1</i>	0.328	3.209	1.250	0.015	0.311	NS	-
143	<i>EPHB3</i>	-0.587	3.204	1.342	0.006	0.205	0.028	0.375
144	<i>Clorf106</i>	-1.276	3.202	1.489	0.001	0.091	NS	-
145	<i>ANKMY2</i>	-1.853	3.201	1.165	0.017	0.328	NS	-
146	<i>GPR17</i>	-0.336	3.198	1.596	0.003	0.130	0.014	0.286
147	<i>MYCL1</i>	0.122	3.194	1.625	0.004	0.158	NA	NA
148	<i>DUSP19</i>	-0.648	3.191	1.249	0.036	0.482	NS	NA

4.4.3 Gene Set Enrichment Analysis.

Gene set enrichment analysis GSEA was performed to identify gene sets significantly enriched between differential expression analysis of the *CIC*-mutant versus *CIC*-wild type samples from the TCGA cohort. The gene set enrichment analysis was performed using R Bioconductor package SeqGSEA [131]. The datasets used were the c5 GO gene sets and c6 oncogenic signature gene sets (Molecular Signatures Database v5.0) [132]. The genes involved in the negative regulation of the MAP kinase (MAPK) signaling pathway and those upregulated by the *KRAS* oncogene were found to be significantly enriched in the *CIC*-mutant tumors. The gene sets and corresponding p-value and FDR is shown the figure 4.12.

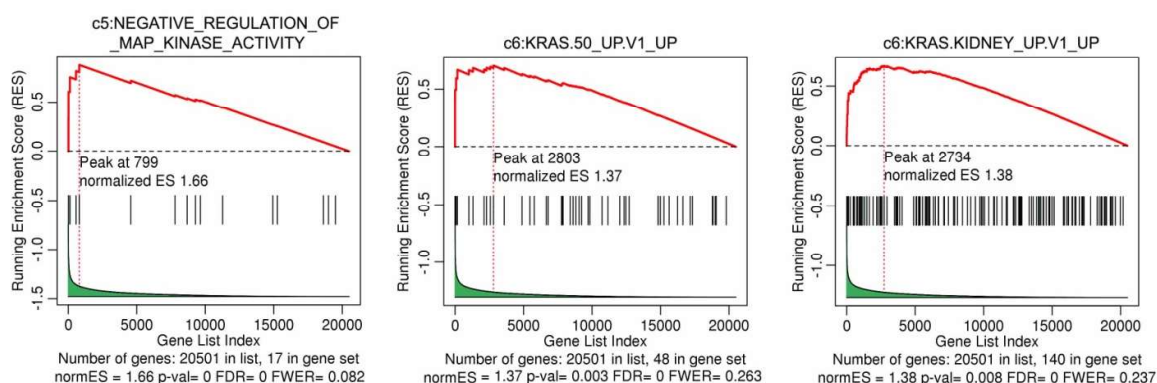


Figure 4.12: The gene sets significantly enriched in the differential expression analysis comparing the *CIC*-mutant and *CIC*-wild type tumors in the TCGA dataset done using the c5 GO gene sets and c6 oncogenic signature gene sets (Molecular Signatures Database v5.0).

Additional analyses were performed to identify potential biological significance of the upregulated genes. The KEGG pathway analysis of the gene set significantly differentially expressed between *CIC*-mutant and *CIC*-wild type tumors also identified enrichment of a number of genes in the MAPK signaling pathway ($P=0.0019$ and $FDR=0.0199$). These MAPK pathway genes included the dual specificity phosphatase genes *DUSP4*, *DUSP6*, and *DUSP19*, the Sprouty family members *SPRY4*, *SPRED1*, and *SPRED2*, and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, and *EPHB*.

The MAPK pathway genes significantly upregulated in the *CIC* mutant tumours included the dual specificity phosphatase genes *DUSP4*, *DUSP6*, and *DUSP19*, the Sprouty family members *SPRY4*, *SPRED1*, and *SPRED2*, and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, and *EPHB*. *ETV1*, *ETV4*, and *ETV5*, the three genes belonging to the ETS/PEA3 family of transcription factors, were found to be significantly upregulated in the *CIC*-mutant tumors. *CREB3L1*, a member of the CREB/ATF family transcription factors that modulates unfolded protein response signaling, was also found to be upregulated in the *CIC*-mutant oligodendrogliomas. Two oligodendrogliomas (ODG8 and ODG9) carried activating mutation in *KRAS* gene. These two cases had higher expression of some of the negative regulators of tyrosine kinase receptor signaling pathway such as *SPRY4*, *SPRED2*, and *DUSP6* (Figure 4.11 [B]). The TCGA data, however, contains only one tumor with an activating mutation in the *NRAS* gene out of the 65 oligodendrogliomas.

4.5 In Silico Analysis of HMG Domain

The *CIC* gene located on chromosome 19q13.2 encodes transcriptional repressor protein CIC. The *CIC* gene expresses two isoforms of the CIC protein CIC-S (encoded by exon 1 to exon 20, 1608 amino acids) and CIC-L (encoded by exon 0 to exon 20, 2517 amino acids, contains an extended N-terminal segment) [128]. *CIC* gene mutations in 1p/19q codeleted oligodendrogliomas from current study cohort and those reported by different groups were mapped on the *CIC* gene [47], [50], [103], [133]. The *CIC* gene mutations show a peculiar pattern (Figure 4.13). Protein truncating mutations such as frameshift insertion/deletion, stop gain/stop loss and splice site mutations occur throughout the coding portion of the gene. Interestingly nonsynonymous mutations appear to cluster in exon 5 and exon 19-20. CIC proteins share two highly conserved domains – the HMG-box which is encoded by exon 5 (involved in DNA binding) and a C-terminal motif C1 of unknown molecular function encoded by exon 19-20 [128]. Deletion of one copy of 19q and high frequency of protein truncating mutations

indicate loss of function of CIC repressor protein in oligodendrogliomas. An *in silico* analysis was performed to probe the impact of nonsynonymous mutations affecting DNA binding HMG domain of CIC.

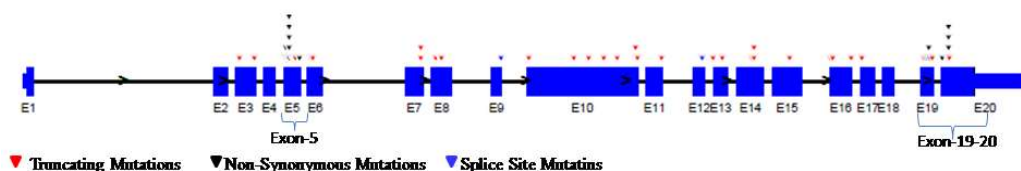


Figure 4.13 : *CIC* gene mutations from current study and those reported by different groups mapped on *CIC* gene ([47], [50], [103], [133]). Deleterious protein truncating mutations map to all exons whereas non synonymous mutations are restricted to exon 5 and exon 19-20.

The HMG box domain amino acid sequence from CIC was analyzed using NCBI tool BLAST [<https://blast.ncbi.nlm.nih.gov/Blast.cgi>]. The BLAST results identified the amino acid sequence to be part of HMG box superfamily and predicted DNA binding amino acids using conserved domain database (Figure 4.14).

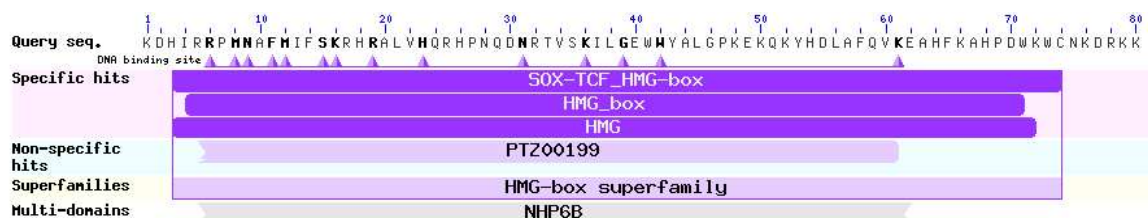


Figure 4.14: BLAST results for the CIC HMG box domain amino acid sequence. The triangular symbols point the predicted DNA binding amino acids.

The amino acid sequence of the CIC HMG box domain was analyzed using FRpred [<https://toolkit.tuebingen.mpg.de/frpred>] to assess sequence conservation of the domain. The amino acid sequence of the CIC HMG box domain was found to be conserved across species and across HMG superfamily genes (Figure 4.15). Some amino acids were found to be highly conserved. Many of these conserved amino acids from CIC HMG box domain were found to be

frequent targets of nonsynonymous mutations (201R, 202R, 203P, 210F, 215R, 228R, 234L, 238W and 253A) identified in oligodendrogloma tumor tissues (Appendix-IV, Appendix-V, Table 4.5) [47], [50], [103].

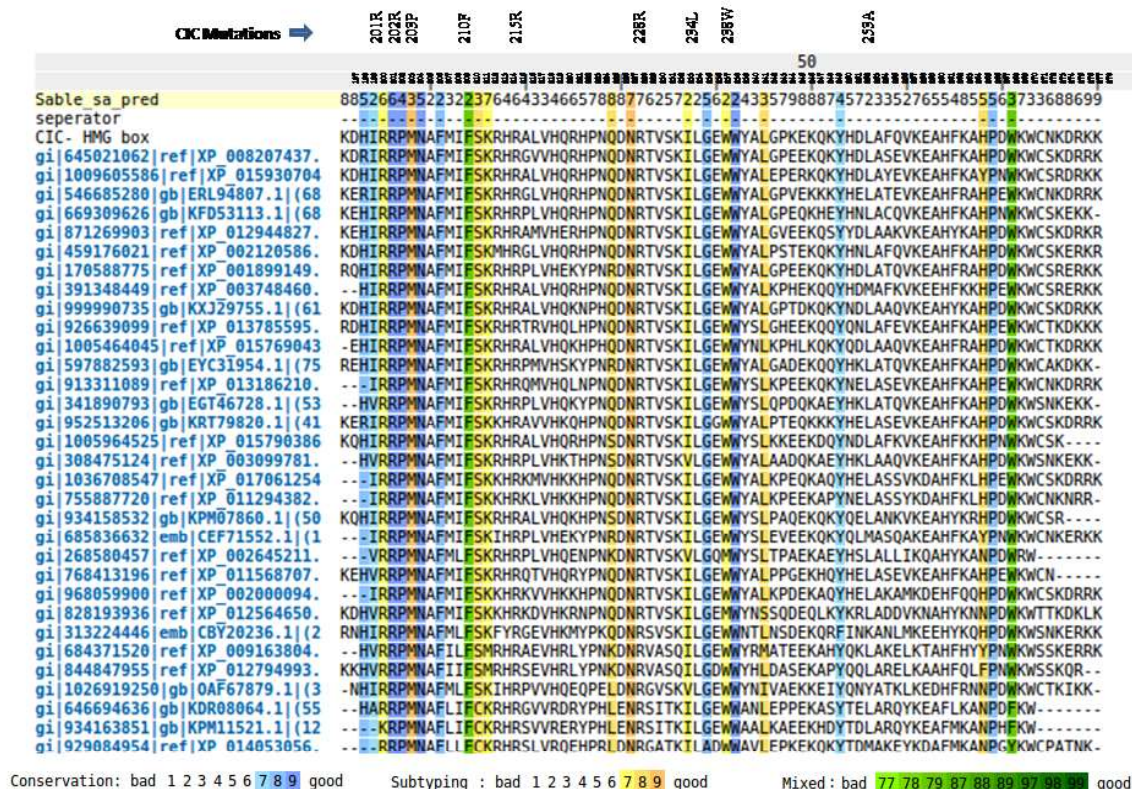


Figure 4.15: FRpred amino acid sequence conservation analysis results.

Seven X-ray crystallographically derived structures of HMG domains (Table 4.11) downloaded from protein data bank were analyzed for sequence and structure similarity. These structures include HMG domains of SOX2, SOX4, SOX9, SOX17 and SOX18. Multiple sequence alignment of the CIC HMG box domain amino acid sequence and the seven HMG domain structures was carried out using ClustalW in Multiseq. All the eight sequences were found to have considerable sequence similarity. Some of the amino acids common to the eight sequences

were found to undergo nonsynonymous mutations in chromosome 1p/19q codeleted gliomas. (Figure 4.16). In particular, the CIC amino acids 202P, 203P, 215R, 234L, 238W and 253A that have been identified to be mutated are common to the eight sequences.

Table 4.11: X-ray crystal structures of HMG domain from different genes and organisms.

PDB ID	Description	Organism	Method
4A3N	Crystal Structure of HMG-BOX of Human SOX17	<i>Homo sapiens</i>	X-RAY DIFFRACTION
3F27	Structure of SOX17 Bound to DNA	<i>Mus musculus</i>	X-RAY DIFFRACTION
1GT0	Crystal structure of a POU/HMG/DNA ternary complex	<i>Mus musculus</i> <i>Homo sapiens</i>	X-RAY DIFFRACTION
3U2B	Structure of the Sox4 HMG domain bound to DNA	<i>Mus musculus</i>	X-RAY DIFFRACTION
4S2Q	Crystal Structure of HMG domain of the chondrogenesis master regulator, Sox9 in complex with ChIP-Seq identified DNA element	synthetic construct <i>Mus musculus</i>	X-RAY DIFFRACTION
4Y60	Structure of SOX18-HMG/PROX1-DNA	<i>Mus musculus</i>	X-RAY DIFFRACTION
4EUW	Crystal structure of a HMG domain of transcription factor SOX-9 bound to DNA (SOX-9/DNA)	<i>Homo sapiens</i>	X-RAY DIFFRACTION

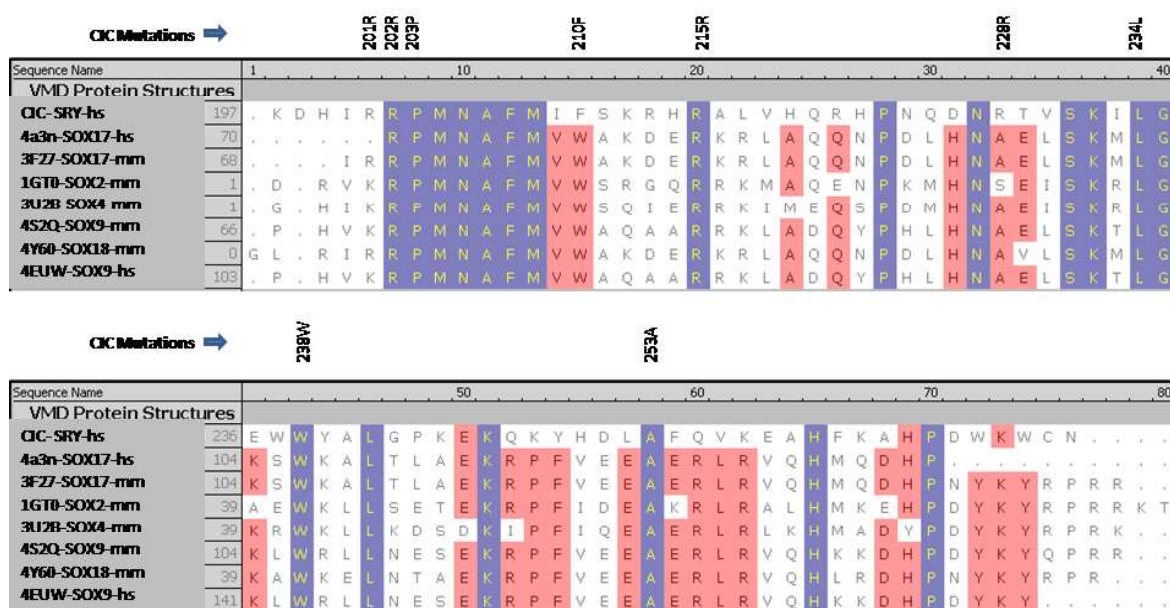


Figure 4.16: Multiseq Alignment of the CIC HMG box domain amino acid sequence and amino acid sequence from seven HMG box domain crystal structures.

The seven X-ray crystallographically derived structures of the HMG domains were aligned using STAMP structural alignment in Multiseq (Figure 4.17). The seven HMG domains aligned exhibiting high structural similarity.

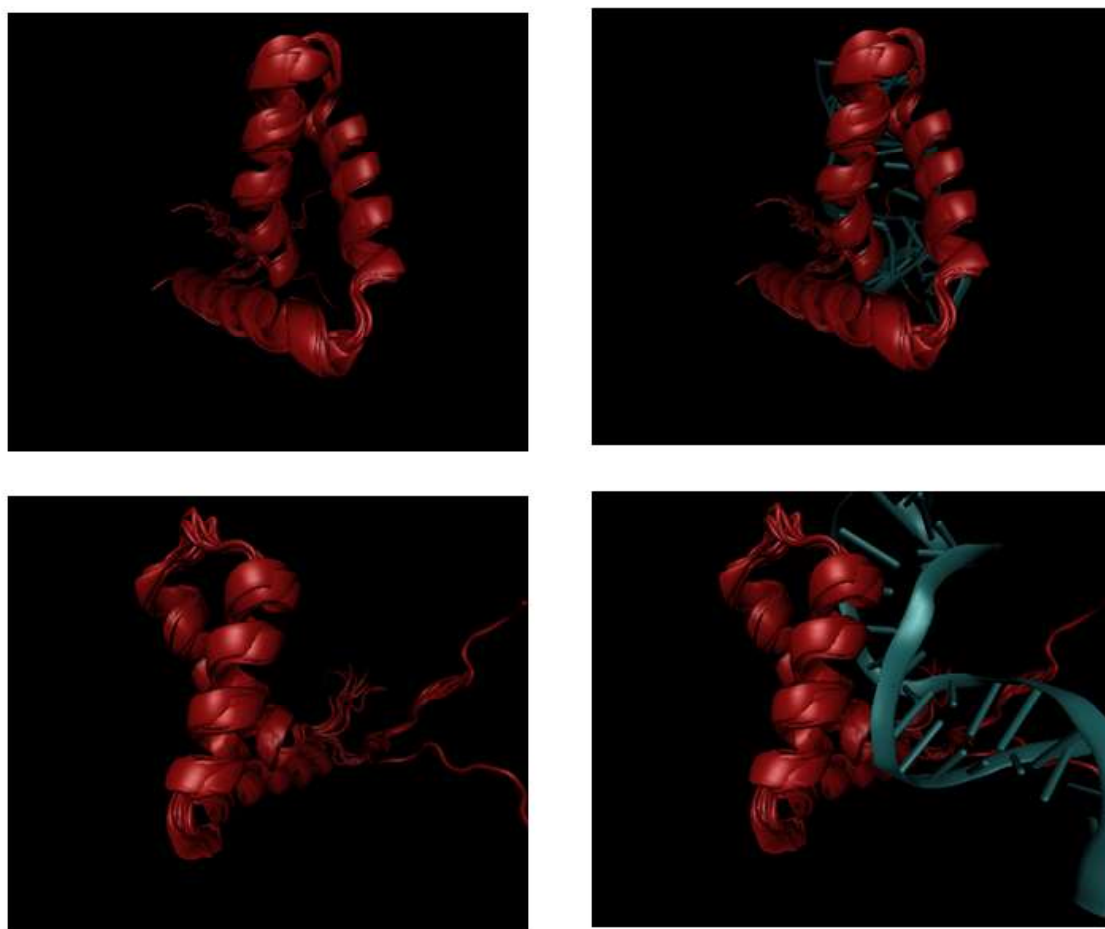


Figure 4.17: Cartoon representation of the structural alignment of seven HMG domains. Left images shows the seven aligned HMG domains. The right images show aligned HMG domain structures bound to the DNA from SOX17 HMG domain and DNA cocrystal structure [pdb 3F27].

As a co crystal structure of CIC HMG domain and DNA is not available, SOX17 HMG domain with DNA bound crystal structure was used as a surrogate for studying CIC HMG domain and DNA interaction due to high amino acid sequence and structural similarity in their HMG domains. Thus SOX17 Crystal, bound to DNA at the binding sequence ‘ACAATAGA’, was used to analyze the amino acid-DNA interaction. CIC HMG box domain recognizes octameric [T(G/C)AATG(A/G)A] sequence [128]. VMD visualization tool was used to find out the amino

acids that are in close proximity with the DNA. Table 4.12 gives the list of amino acids identified as being closest to the DNA.

Table 4.12: The amino acids in SOX17 HMG domain that are closest to the bound DNA.

Distance in Å of Amino Acid from DNA	Sox17 Bound to DNA 14-3F27.PDB Amino Acid	CIC Residue No.
2.7	R141	Not common with CIC HMG box
2.8	S99	S231
2.83	R70	R202
2.85	N95	N227
2.89	Y137	Not common with CIC HMG box
2.91	W106	W238
2.92	N73	N205
3.1	R69, R70, S99	R201,R202,S231
3.2	H94	H226
3.3	F75, R83	F207,R215
3.5	M76	M208

Eight of the ten most proximal amino acids within (R202, N205, F207, M208, R215, N227, S231, W238) were found to be conserved in all eight sequences (Figure 4.16). Two proximal amino acids R201 and H226 are not conserved in all eight sequences. Amino acids R141 and Y137 are not found in the corresponding 141 and 137 positions in CIC HMG domain sequence. Among the eight most proximal amino acids R202 and W238 were found to be frequent targets of nonsynonymous mutations in *CIC* (Appendix-IV, Appendix-V, Table 4.5) [47], [50], [103].

The highly conserved, most proximal amino acids likely play a significant role in the binding of the HMG box domain to the DNA.

4.6 Mutation Based Stratification of Copy Number Alterations of TCGA Low Grade Glioma Tumours.

The copy number alteration data of TCGA low grade glioma tumour samples were stratified according to the genes that have most frequently undergone somatic mutations. The data used for the analysis was downloaded from The Cancer Genome Atlas data portal [<http://tcga-data.nci.nih.gov/>]. A total of 251 low grade glioma (WHO grade II and III) tumour sample data, for which both somatic mutation and copy number alteration data was available, was downloaded from TCGA data portal [<http://tcga-data.nci.nih.gov/>]. The tumour samples had been histopathologically diagnosed by TCGA panel of histopathologists as Astrocytoma (n=87), Oligodendroglioma (n=98) and oligoastrocytoma (n=66). The copy number alteration data of the samples was stratified according to most frequent somatic mutations among the samples ie. *IDH1/2*, *TP53*, *ATRX*, *CIC* and *FUBP1*. The resultant data structure was visualized in Integrative Genomics Viewer (IGV) [134], [135].

The mutation based stratification of the copy number alteration resolved the data into three major molecular subgroups (Figure 4.19). (1) *IDH1/2* mutant without chromosome 1p/19q codeletion. (2) *IDH1/2* mutant with chromosome 1p/19q codeletion and (3) *IDH1/2* wild type. Eight samples showed overlapping molecular alterations.

TP53 gene was the most frequently (96.12%, 124/129) mutated gene among the *IDH1/2* mutant without chromosome 1p/19q codeletion subgroup. The next most frequent gene mutation in this subgroup was *ATRX* (79.84%, 103/129). Among the *IDH1/2* mutant with chromosome 1p/19q codeleted subgroup the most frequently occurring mutated gene was *CIC* (54.93%, 39/71). The next most frequently mutated gene in this subgroup was *FUBP1* (28.17%, 20/71). Positive

1p/19q codeletion status was strongly associated with Oligodendroglioma morphology (83.09%, 59/71) (Figure 4.18).

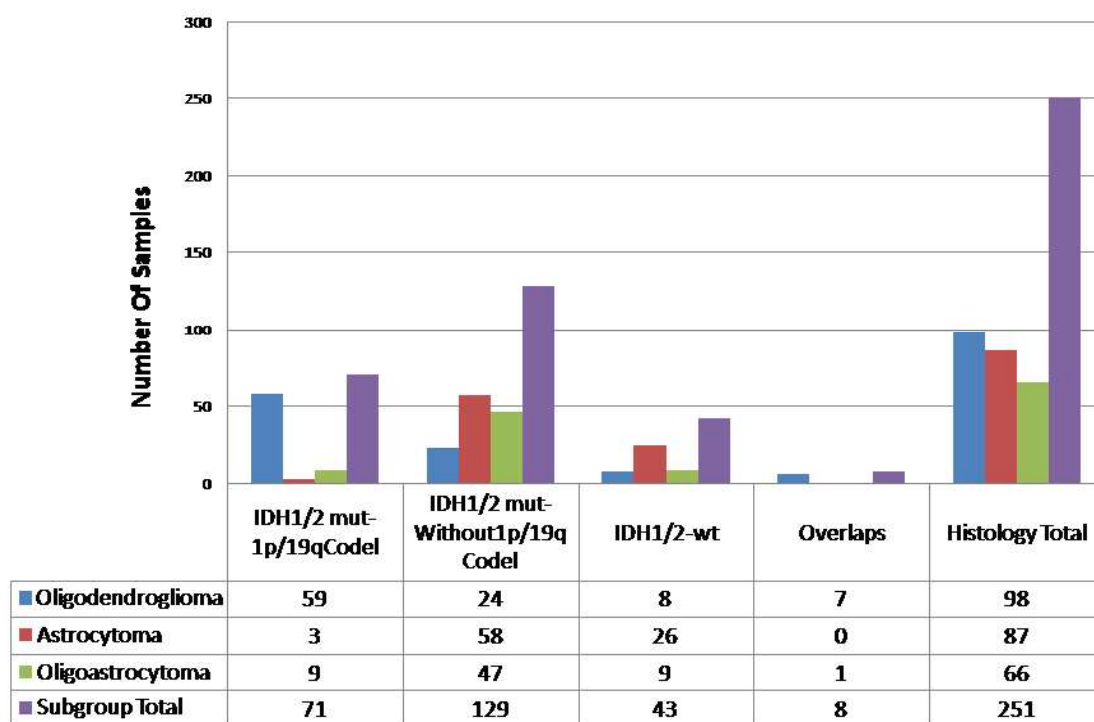


Figure 4.18: Breakup of 251 TCGA low grade glioma samples according to molecular subgroup and histology as a result of mutation based stratification of copy number alteration. The upper panel shows bar chart for each subgroup data.

Among the *IDH* wild type subgroup the frequent somatic alteration were *PDGFRA* amplification, *CDKN2A* deletion, *EGFR* amplification, *NF1* mutation and *PTEN* mutations. The most prominent chromosome level alterations were chromosome 7 amplification, partial chromosome 9p deletion containing *CDKN2A* gene and chromosome 10 deletion. This molecular alteration pattern is similar to that found in primary glioblastoma. According to the molecular alterations this *IDH1/2* wild type subgroup seems closer to primary glioblastoma than low grade glioma. 90.69% of the tumours in this subgroup were WHO grade III (39/43).

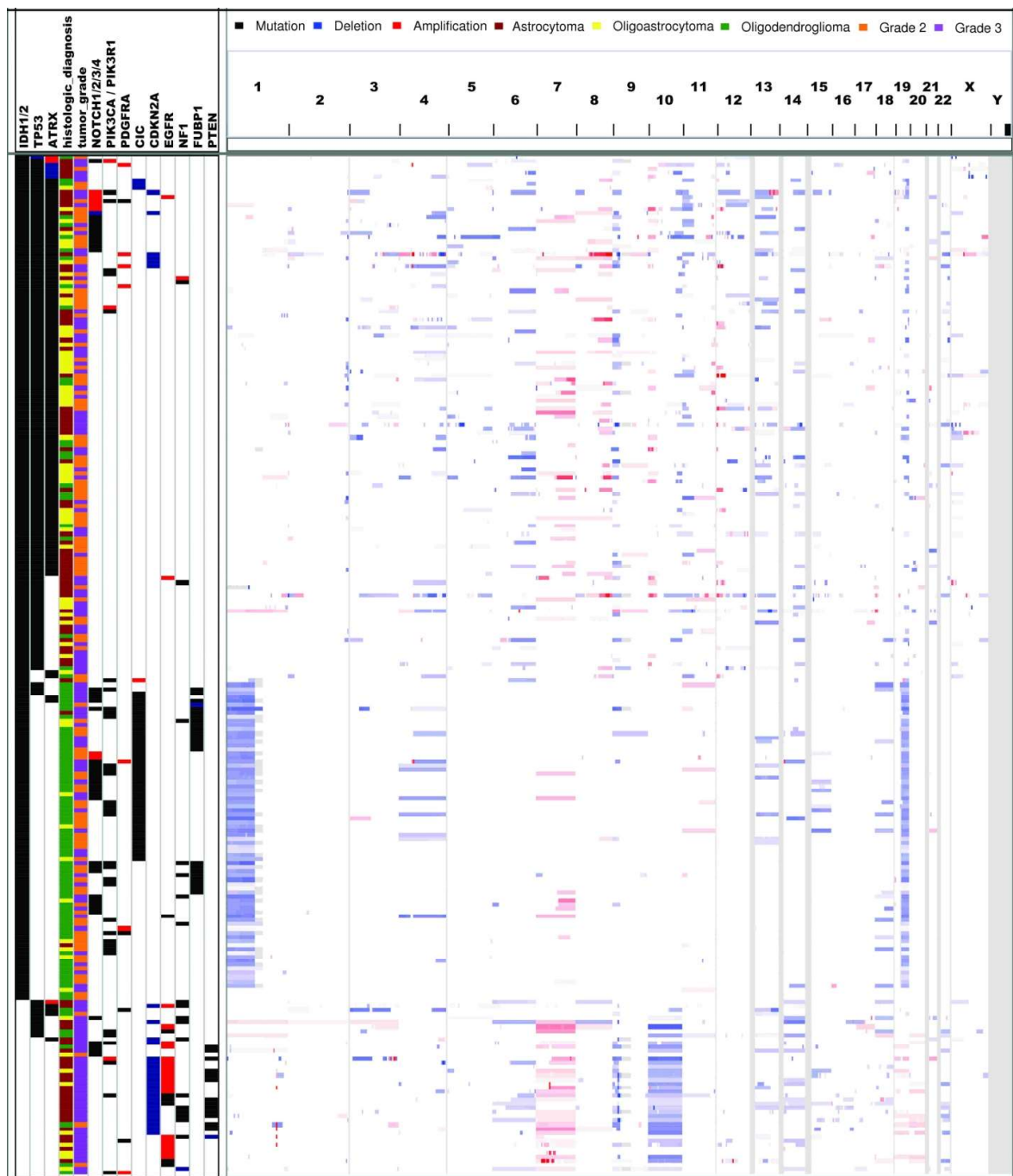


Figure 4.19: Integrated genomic view of mutation based stratification of copy number alteration of 251 TCGA low grade glioma (WHO grade II and III) samples. The three major subgroups are (1) *IDH1/2* mutant without chromosome 1p/19q codeletion. (2) *IDH1/2* mutant with chromosome 1p/19q codeletion and (3) *IDH1/2* wild type.

Eight tumour samples with overlapping molecular alterations included three *IDH1/2* mutated samples without chromosome 1p/19q codeletion which had homozygous deletion of *CIC* gene. Three *IDH1/2* mutated samples with chromosome 1p/19q codeletion which had *TP53* mutations. Two *IDH1/2* mutated samples with chromosome 1p/19q codeletion had *ATRX* mutations.

4.7 Targeted Sequencing of Oligodendroglial Tumour Samples.

To assess the suitability of the Ion torrent platform for targeted sequencing of a small panel of genes to identify nucleotide variants targeted sequencing of two samples from the present study cohort was performed. Ion AmpliSeq Cancer Hotspot Panel v2 was used to construct the amplicon libraries that were sequenced on Ion Torrent platform. The Ion AmpliSeq Cancer Hotspot Panel v2 targeted cancer hotspot regions, including ~2,800 COSMIC mutations of 50 oncogenes and tumor suppressor genes, with wide coverage of the *KRAS*, *BRAF*, and *EGFR* genes. The panel consisted of 207 primer pairs to amplify 207 amplicons ranging in length 111–187 bp (average 154 bp). The sequencing was performed using Ion 314 Chip by multiplexing two samples with barcodes. The nucleotide sequence variants were detected after alignment of the short sequence reads to the hg19 reference genome. Total of 56.1 million bases were sequenced constituting 506,377 total reads. The mean read length was 111 bp. Ion sphere particle loading was achieved at 76%. Table 4.13 shows the data statistics for the two samples. Table 4.14 and Table 4.15 enlist the nucleotide variants detected in sample ODG10 and ODG 11 respectively.

Table 4.13: Data statistics for Ion torrent run

Sample	Mapped Reads	On Target	>=Q20 Bases	Mean Depth	Variants Detected	Hotspot Variants
ODG10	226,000	98.33%	23,659,694	1,046	12	2
ODG11	261,961	98.37%	26,949,168	1,198	18	4

Table 4.14: Variants detected in sample ODG10

Chrom	Position	Gene ID	Ref	Variant	Allele Call	Allele Source	Allele Name
chr2	209113113	<i>IDH1</i>	G	A	Heterozygous	Hotspot	COSM28747
chr3	178917005	PIK3CA	A	G	Heterozygous	Non Hotspot	---
chr4	55593464	KIT	A	C	Heterozygous	Hotspot	COSM28026
chr4	55980239	KDR	C	T	Heterozygous	Non Hotspot	---
chr5	112175770	APC	G	A	Heterozygous	Non Hotspot	---
chr5	149433596	CSF1R	TG	GA	Heterozygous	Non Hotspot	---
chr7	55249063	<i>EGFR</i>	G	A	Heterozygous	Non Hotspot	---
chr13	28610153	FLT3	G	A	Heterozygous	Non Hotspot	---
chr17	7579472	<i>TP53</i>	G	C	Heterozygous	Non Hotspot	---
chr19	1220321	STK11	T	C	Heterozygous	Non Hotspot	---
chr4	1807894	FGFR3	G	A	Homozygous	Non Hotspot	---
chr4	55141055	<i>PDGFRA</i>	A	G	Homozygous	Non Hotspot	---

Table 4.15: variants detected in sample ODG11

Chrom	Position	Gene ID	Ref	Variant	Allele Call	Allele Source	Allele Name
chr2	212812097	ERBB4	T	C	Heterozygous	Non Hotspot	---
chr3	178917005	PIK3CA	A	G	Heterozygous	Non Hotspot	---
chr3	178952020	PIK3CA	C	T	Heterozygous	Hotspot	COSM21451
chr4	55152040	<i>PDGFRA</i>	C	T	Heterozygous	Hotspot	COSM22413
chr4	55972974	KDR	T	A	Heterozygous	Non Hotspot	---
chr7	55249063	<i>EGFR</i>	G	A	Heterozygous	Non Hotspot	---
chr10	43615633	RET	C	G	Heterozygous	Non Hotspot	---
chr11	108117809	ATM	C	A	Heterozygous	Non Hotspot	---
chr19	1223104	STK11	G	A	Heterozygous	Non Hotspot	---
chr19	1223125	STK11	C	G	Heterozygous	Hotspot	COSM21360
chr22	24176287	SMARCB1	G	A	Heterozygous	Hotspot	COSM1090
chr4	1807894	FGFR3	G	A	Homozygous	Non Hotspot	---
chr4	55141055	<i>PDGFRA</i>	A	G	Homozygous	Non Hotspot	---
chr5	112175770	APC	G	A	Homozygous	Non Hotspot	---
chr5	149433596	CSF1R	TG	GA	Homozygous	Non Hotspot	---
chr10	43613843	RET	G	T	Homozygous	Non Hotspot	---
chr13	28610183	FLT3	A	G	Homozygous	Non Hotspot	---
chr19	1220321	STK11	T	C	Homozygous	Non Hotspot	---

All the identified variants were checked in the Exome sequence data of the corresponding tumour. All these variants identified from targeted sequencing were found to be present in the Exome data as well.

DISCUSSION

5 DISCUSSION

Brain tumours are a major cause of deaths resulting from cancer in children and adults. Gliomas account for almost 80% of primary malignant brain tumours. Diagnosis of gliomas still primarily depends on histopathologic analysis of H & E stained slides of tumour tissue sections. Histopathologic diagnosis is particularly challenging when it comes to reliably distinguishing between oligodendroglial and astrocytic component in low grade to intermediate grade gliomas. Accurate diagnosis of glioma subtypes is not just of academic interest but is necessary for deciding treatment strategy and for prognostication. Oligodendrogliomas generally have slower growth rates and have better prognosis than astrocytomas of similar grade. Oligodendrogliomas especially those having combined loss of chromosome 1p and 19q are known to be particularly sensitive to chemotherapy with much longer progression free survival. Combined loss of chromosome 1p and 19q is known to occur in about 60-90% oligodendrogliomas. However genes residing on these chromosomal regions which are likely to have undergone mutational alterations and thereby contribute to pathogenesis of oligodendrogliomas were not identified. Identification of genetic alterations specific to oligodendroglial tumours was necessary for accurate diagnosis and prognostication of these relatively chemo/radiation sensitive tumours. Understanding of genetic alterations is a prerequisite to development of specific targeted therapies for effective treatment of oligodendrogliomas. In the present study therefore exome sequencing and transcriptome sequencing of oligodendroglioma tumor tissues was carried out with the aim of identifying genetic alterations and their functional role.

According to current WHO definition, oligodendroglioma is a well differentiated diffusely infiltrating tumour of adults typically located in cerebral hemispheres and composed predominantly of cells resembling oligodendroglia. Pure oligodendrogliomas (ODs) are composed of single cell type while mixed oligoastrocytomas have morphological characteristics

of both pure oligodendrogliomas and astrocytomas. Only pure oligodendrogliomas as diagnosed histopathologically were included in the study in order to identify oligodendroglioma specific genetic alterations. Oligodendrogliomas together with mixed oligoastrocytomas (OAs) constitute only 12-20% of all gliomas. Therefore only a small number of oligodendrogliomas that were prospectively recruited along with the paired blood specimen could be included in the study. Paired blood specimen needs to be sequenced in parallel due to scarcity of genome sequencing data from Indian population.

Sanger Sequence Analysis and Copy Number Variation Analysis shows presence of *IDH1/IDH2* mutation and chromosome 1p/19q codeletion in majority of the oligodendroglioma tumor tissues studied

Mutations in *IDH1* encoding Isocitrate dehydrogenase were first identified in a genome-wide mutational analysis of glioblastomas (WHO grade IV glioma) in a small fraction of such tumors, most of which were known to have evolved from lower-grade gliomas [69]. *IDH1* gene encodes cytosolic NADP⁺ dependent Isocitrate Dehydrogenase while *IDH2* encodes mitochondrial NADP⁺ dependent Isocitrate Dehydrogenase. Somatic mutations in the *IDH1* gene were subsequently reported to occur in majority of astrocytomas, oligodendrogliomas and oligoastrocytomas of WHO grades II and III [22], [57, p. 1], [86]. *IDH2* is also mutated in these gliomas but at much lower frequencies and was found to be mutually exclusive with mutation in *IDH1* gene [22]. Oligodendrogliomas recruited in the present study were therefore analyzed for mutation in *IDH1/IDH2* gene. Of the 11 oligodendroglioma tumour tissues analyzed by Sanger sequencing for *IDH1/IDH2* mutation status, 10 tumor tissues were found to be *IDH1/IDH2* mutated. In a study done on 302 grade II/III oligodendrogliomas, 105 grade II tumors (82.0%), 128 grade III tumors (69.5%) were found to carry mutation in *IDH1* gene while 6 grade II (4.7%) and 9 grade III (5.2%) tumors carried mutation in *IDH2* gene [86]. Other group studied *IDH1* mutations in 105 oligodendrogliomas (51 grade II and 54 grade III) for and found *IDH1*

mutations in 71% of grade II and 67% of grade III oligodendrogliomas [57]. In a study reported by Yan et al., of the 51 grade II oligodendrogliomas analyzed 41 carried mutations in *IDH1* while 2 carried mutations in *IDH2*. In the same study, of the 36 grade III oligodendrogliomas 31 harbored mutation in *IDH1* while 3 carried mutation in *IDH2* [22]. Thus, presence of *IDH1* and *IDH2* mutation in 9 out of 11 (81%) and 1 out of 11 (9.0%) oligodendroglial tumors respectively in the present cohort is consistent with the reported studies. The predominant amino acid sequence alteration in *IDH1* is known to be R132H accounting for 92.7% of the detected mutations while R172K in *IDH2*, accounts for 65% of the detected mutations [86]. Among the nine *IDH1* mutated tumour tissues in the present study, eight (88.9%) were found to carry R132H mutation while one was found to have R132C mutation and one tumor was found to carry R172K mutation in *IDH2* gene consistent with the published data.

Combined loss of short arm of chromosome 1 (1p) and long arm of chromosome 19 (19q) has been reported to occur in 60-90% of pure oligodendrogliomas and 10-20% of mixed Oligoastrocytomas [75], [136]–[138]. Most commonly used technique for detection of chromosome 1p/19q codeletion is Fluorescence In situ Hybridization (FISH) [139], [140]. FISH analysis was performed by pathology department of Tata Memorial Hospital while loss of heterozygosity (LOH) of microsatellite markers was used to assess the deletion status of a chromosomal arms in the present study [141]–[144]. LOH of microsatellite markers can be used for copy number variation analysis provided polymorphism exists in the length of microsatellite repeats on two alleles in germ line (i.e. in the paired blood DNA used in the present study). Therefore multiple markers need to be used for the LOH analysis so that at least some are polymorphic in an individual specimen. Exome sequencing data of the 11 oligodendroglioma tumor tissues was also analyzed for copy number variation using both loss of heterozygosity analysis using Control-FREEC algorithm as well as based on the coverage using FishingCNV software. Both these analysis showed combined loss of chromosome 1p and 19q in 9 out of 11

tumor tissues (ODG1 to ODG9) and the loss included almost the entire arms. In microsatellite marker analysis, D1S548 was found to be most polymorphic and showed deletion in 4 out of 8 tumor tissues tested while in the rest 4 cases it was non-informative. In case of two tumor tissues, 3 or 4 out of 5 microsatellite markers on chromosome 1p were not informative while the remaining markers did not show deletion. FISH analysis of these two tumor tissues however did show deletion of chromosome 1p arm in accordance with the exome sequence data. Microsatellite marker D19S178 on chromosome 19q arm, on the other hand was found to be polymorphic and deleted in 6 out of 8 tumor tissues studied in accordance with the exome sequence data. However, FISH analysis was not interpretable in some cases particularly in the case of chromosome 19q arm most likely due to the small size of the arm and poor quality of paraffin embedded tumor tissue DNA. Thus, copy number variation analysis was most accurate dependable from analysis of the exome sequence data since it analyzed the entire chromosomes while FISH analysis although accurate, at times was not interpretable due to poor quality paraffin blocks. Microsatellite marker analysis is a technique that costs the least but lacks sensitivity due to lack of polymorphism in the markers. It is also less accurate since multiple bands may appear even in the absence of polymorphism as DNA polymerase stutters during amplification of the repeat sequences.

Genetic alterations identified in *IDH*-mutant chromosome 1p/19q codeleted oligodendrogliomas and their functional significance

Nine out of 11 oligodendrogliomas were found to carry 1p/19q codeletion and *IDH1* mutation, genetic alterations known to be characteristic of oligodendrogliomas [75], [136]–[138]. The total number of somatic mutations in these tumors ranged from 10 – 46 in these tumor tissues. The rate of mutation was found to be 0.46/MB and transition to transversion ratio (Ti/Tv) was found to be 3.54. *IDH1* gene is the most commonly mutated gene in oligodendrogliomas, followed by *CIC*

and *NOTCH1* gene. *IDH1* gene mutations however, are known to occur in most grade II/III gliomas including astrocytomas. The gene *CIC* located on chromosome 19q arm was found to be the most commonly mutated gene in these chromosome 1p/19q codeleted oligodendrogliomas with four out of nine tumors carrying 3 missense and two frameshift deletion mutations.

Two tumors carried mutation introducing a stop codon and a non-frameshift deletion in *FUBP1* gene located on chromosome 1p arm indicating tumor suppressor role of this gene since one copy is deleted and other allele carries loss of function mutation. Mutations in *FUBP1* gene have been reported in chromosome 1p/19q codeleted gliomas although at much lower frequency (less than or equal to 30 %) [47], [49], [50] than that in the *CIC* gene. *FUBP1* encodes Far Upstream Element Binding Protein 1. FUBP1 is a single stranded DNA binding protein that is known to bind multiple DNA elements including FUSE (Far upstream element) located upstream of *MYC* oncogene. *FUBP1* has been shown to activate *MYC* promoter by stimulating TFIIH helicase activity. However, functional role of FUBP1 in oligodendroglioma is not known. Both, *FUBP1* gene located on chromosome 1p arm and *CIC* gene located on chromosome 19q arm are known to be altered only in chromosome 1p/19q codeleted oligodendrogliomas and not in astrocytomas.

Missense/Non-frame shift deletion mutations in *NOTCH1* gene were identified in four oligodendrogliomas while one tumor carried frame-shift mutation in *MAML3* gene. Notch signaling pathway is a well known signaling pathway involved in development various tissues and organs [145]–[147]. Alterations in Notch signaling pathway genes leading to both activation and inactivation have been reported in various cancers [147]. *MAML3* gene is known to activate NOTCH signaling by acting as a co-activator [148]. Loss of function mutation in *MAML3* gene suggests inactivation of Notch signaling pathway in oligodendrogliomas. In two large scale exome sequencing studies of grade II/grade III adult gliomas, one done by The Cancer Genome Atlas Project and another a study from Japan have identified frequent alterations in NOTCH signaling genes, most commonly (31%) in *NOTCH1* gene in chromosome 1p/19q co-deleted

gliomas. These mutations are located in hotspots similar to those identified as inactivating mutations in lung, Head & Neck and cervical cancers suggesting that mutations in Notch signaling genes lead to inactivation of Notch signaling in chromosome 1p/19q co-deleted gliomas. Inactivation of Notch signaling pathway however remains to be demonstrated experimentally in these gliomas.

Two out of the 11 oligodendrogliomas studied were found to carry frame-shift deletion mutation in chromatin modifier *ARID1A* gene. Large scale exome sequencing studies of various cancers in last 4-5 years have identified mutations in various epigenetic modifier genes including genes of SWI/SNF chromatin remodeling complex like *SMARCA4*, *SMARCB1* in atypical teratoid/rhabdoid tumors, genes encoding histones like *H3F3A* in pediatric glioblastomas and histone modifier genes like *MLL1-MLL5* genes in leukemias and medulloblastomas. *ARID1A* gene belongs to SWI/SNF chromatin remodeling complex and inactivating mutations in this gene have been reported in various cancers including lung cancer.

Comprehensive analysis of exome/genome sequencing studies has suggested that a tumor contains 2 to 8 ‘driver gene’ mutations and driver genes can be classified in 12 major signaling pathways that regulate three core regulatory processes namely; cell fate, cell survival and genome maintenance. Notch pathway genes and *ARID1A*, chromatin modifier gene as well as *IDH1* gene that were found to be mutated in chromosome 1p/19q codeleted oligodendrogliomas control the cell fate determination core process. Other genes including *CIC*, *KRAS*, *IDH1* and *FUBP1* that are altered in these gliomas contribute to cell proliferation/cell survival core process as discussed below.

***CIC*, the most commonly altered gene in oligodendrogliomas and likely impact of the mutations identified on *CIC*’s role as a transcription factor**

CIC appears to act as a tumor suppressor gene in oligodendrogliomas, with one of the two copies of the gene deleted and the other copy carrying either protein truncating mutation or potentially deleterious missense mutation, as indicated by the present study as well as other reports [47], [149]. The missense mutations in the *CIC* gene in the TCGA study as well in our study were predominantly localized in the exons 5 and 20 of the *CIC* gene, as has been reported before [50]. These exons of the *CIC* gene are known to be involved in the HMG box DNA binding domain and protein–protein interaction domain of the *CIC* protein, respectively.

CIC belongs to HMG domain superfamily of DNA bending proteins [150]. *CIC* is known to act as transcriptional repressor and depends on the HMG domain to recognize and bind the DNA target sites [128]. To function as transcription factors, DNA chaperones, and DNA repair agents, the ability of the HMG box proteins to bend DNA is essential. Most of the current understanding of the bending mechanism is based on studies of single HMG boxes [150]–[152]. HMG box domains are characterized by three α helices forming an ‘L’ shaped structure [153] (Figure 4.17). The HMG box severely bends and underwinds DNA, using electrostatic and hydrophobic interactions to widen the minor groove and induce a bend towards the major groove. HMG box residues that intercalate DNA also help in stabilizing the distorted DNA structure [154]. The bending induced by HMG box can be seen in the co-crystal structure of the HMG domain and DNA (Figure 4.17).

Taking into account the high sequence conservation, structural similarity of HMG domain and close proximity of highly conserved amino acids of HMG domain to the DNA it can be assumed that the conserved amino acids play an essential role in binding of the HMG domain to the DNA. Majority of the nonsynonymous mutations in the *CIC* exon 5 identified in oligodendrogliomas affect the highly conserved amino acids which may be essential for *CIC* repressor’s HMG domain binding to the DNA. Thus, the loss of *CIC* repression of *CIC* target genes in the *CIC* mutant tumours may be due to loss of *CIC* HMG domain’s ability to bind to the DNA.

Role of *CIC* in the RTK-RAS-MAPK signaling pathway and in the expression of the ETV transcription factors

In order to identify role of *CIC*, the most commonly and specifically mutated gene in chromosome 1p/19q codeleted oligodendrogliomas, we compared expression profiles of *CIC*-mutant to *CIC*-wild type tumors from the present study cohort as well as from the TCGA cohort. *ETV1*, *ETV4* and *ETV5*, the three genes belonging to the ETS/PEA3 family of transcription factors were found to be upregulated in the *CIC*-mutant tumours in both cohorts. The gene set enrichment analysis identified a number of genes involved in the MAP kinase (MAPK) signaling pathway to be significantly enriched ($P = 0.0039$) in the *CIC*-mutant oligodendrogliomas. These MAPK pathway genes included *DUSP4*, *DUSP6*, *DUSP19*, the dual specificity phosphatase genes, Sprouty family members *SPRY4*, *SPRED1* and *SPRED2*, and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, *EPHB*.

Cic, Capicua meaning head-and-tail in Catalan, was identified in developmental studies of *Drosophila* [104]. *Drosophila* Cic plays an essential role downstream of the TORISO and the epidermal growth factor receptors, two tyrosine kinases that transmit the signaling via the RAS-RAF-MAP kinase pathway [107], [108], [128]. Apart from Cic's role in the cell fate determination downstream of Receptor Tyrosine Kinase (RTK) pathways in *Drosophila*, Cic is also known to play a role in regulating growth of imaginal discs downstream of the RTK/RAS-MAPK pathway. Enrichment of MAP kinase signaling pathway genes in expression profile of *CIC*-mutant oligodendrogliomas is consistent with Cic's known role downstream of MAP Kinase signaling pathway during *Drosophila* development.

CIC's role downstream RTK-MAPK signaling pathway has been experimentally demonstrated in human melanoma cells [109]. EGF stimulation of melanoma cells results in the phosphorylation of *CIC* at multiple sites and upregulation of the ETV4/ETV5 transcription factors. SiRNA

mediated knock-down of *CIC* in a melanoma cell line with constitutively activated ERK signaling was found to result in upregulation of *ETV1*, *ETV4*, and *ETV5* indicating *CIC*'s role in the suppression of PEA3 family transcription factors [109]. Some Ewing's sarcomas contain a *CIC-DUX4* translocation resulting in a fusion protein that converts the *CIC* protein from transcriptional repressor to an activator [106]. Expression profile of Ewing's sarcomas containing the *CIC-DUX4* translocation showed upregulation of the *ETV1/ETV5* transcription factors. Furthermore, the *CIC-DUX4* fusion protein binds to the promoter region of *ETV5*, indicating upregulation of *ETV5* gene as a result of direct binding of the *CIC* repressor turned into activator. Ewing's sarcomas carrying *CIC-DUX4* translocation have also been reported to have distinct transcription profiles, with overexpression of ETV family members as compared to the *EWSR1-FLII* fusion [155]. Upregulation of the ETS/PEA3 transcription factors in the *CIC*-mutant oligodendrogliomas is thus consistent with their upregulation in Ewing's sarcomas carrying the *CIC-DUX4* translocation and in melanoma cells on siRNA mediated knock-down of the *CIC* gene expression.

DUSP4, *DUSP6*, *DUSP19*, the dual specificity phosphatase encoding genes which are known to inactivate MAP kinases like ERK1/ERK2/JNK [156] and Sprouty family members *SPRY4*, *SPRED1* and *SPRED2*, which are known to be inhibitors of the Receptor Tyrosine Kinase (RTK)/Mitogen Activated Protein Kinase (MAPK) signaling pathway [157] were also found to be significantly upregulated in the *CIC*-mutant oligodendrogliomas. Upregulation of these negative regulators of the RTK signaling pathway is likely to be due to the constitutive activation of the pathway resulting from the inactive *CIC* mutant protein. Whether downregulation of *CIC* alone is sufficient for the upregulation of the RTK/MAPK signaling target genes like PEA3 transcription factors in oligodendroglioma tumour tissues needs to be investigated further.

In some of the oligodendrogliomas lacking mutations in the *CIC* gene, other components of the RTK/RAS/MAPK signaling pathway appear to be activated by inactivating mutations in the negative regulators of the pathway like *NF1* and by activating mutations in *KRAS*, *NRAS*, and

EGFR based on the analysis of genetic alterations identified in the TCGA and the present study cohort of chromosome 1p/19q codeleted gliomas. PDGF expression in neural progenitor cells or overexpression of mutant *EGFR* under S100 beta promoter has been found to induce oligodendrogliomas in mouse models [158]–[160]. Thus, activation of the RTK/RAS/MAPK signaling pathway appears to be a major driver of the oligodendroglioma pathogenesis.

Oncogenic potential of the ETS/ETV/PEA3 subfamily transcription factors upregulated in the *CIC*-mutant oligodendrogliomas

A number of studies indicate oncogenic potential of the ETS/PEA3 family transcription factors upregulated in the *CIC*-mutant oligodendrogliomas [161], [162]. The first *ETS* (E26 Transformation Specific) transcription factor encoding gene was identified as a transforming gene in the avian E26 erythroblastosis virus. Twenty-eight human ETS family members are known that share ~ 85 amino acid long DNA binding ETS domain. The PEA3 (Polyoma virus enhancer activator 3) subfamily of the ETS family transcription factors includes *ETV1*, *ETV4* and *ETV5* [163], all of which were found to be upregulated in the *CIC*-mutant oligodendrogliomas. ETS family genes including *ETV1* and *ETV4* are known to be involved in chromosomal translocations in Ewing's sarcoma and in peripheral primitive neuroectodermal tumors resulting in EWS-ETS fusion protein [164], [165]. The EWS-ETS fusion protein contains a N-terminal transactivation domain of the EWS gene and a C-terminal DNA binding domain of the *ETS* gene suggesting a role for ETS regulated genes in pathogenesis of these tumors. Majority of the prostate cancers carry a chromosomal translocation involving ETS family gene *ERG* or *ETV1/4/5* genes and the prostate organ specific, androgen-inducible *TMPRSS2* gene resulting in TMPRSS2-ETS fusion protein that is overexpressed in androgen inducible manner in prostate tumors [166]–[168].

Three recent publications [169]–[171] subsequent to our publication in *Genes, Chromosomes and Cancer* (2015) [133] have corroborated our findings of activation of RAS-MAPK signaling pathway and upregulation of ETS/ETV/PEA3 transcription factors in *CIC*-mutant chromosome 1p/19q codeleted gliomas and have cited our paper. Tirosh et al. identified a signature of expression changes between the 28 mutant *CIC* and 27 *CIC* wild-type cells that included increased expression of *ETV1* and *ETV5* in mutant *CIC* cells [170]. Venteicher et al. Estimated that 10 % expression difference between IDH-A (Astrocytoma) and IDH-O (Oligodendroglioma) tumours, using either bulk tumour samples or single malignant cells, is accounted for by loss of function of the transcriptional repressor *CIC*, which is specific to IDH-O, as inferred from a *CIC* expression signature identified from multiple reports including from current study [171]. LeBlanc et al. performed microarray gene expression analyses on *CIC* knockout cell lines to identify genes whose expression was affected by *CIC* loss. HEK-derived *CIC* knockout lines showed increased expression of *ETV1*, *ETV4*, and *ETV5*. They also showed that loss of *CIC* leads to overexpression of downstream members of the mitogen-activated protein kinase (MAPK) signalling cascade [169].

Molecular markers essential for accurate diagnosis of adult gliomas

The glioma subtype diagnosis has historically been based on the histologic appearance of these tumors. Molecular markers have played secondary role to that of histology with respect to subtype diagnosis. Histology based classification of gliomas suffers from high intraobserver and interobserver variability [10]. Histology based classification also fails to accurately predict clinical outcomes [9], [10]. Accurate subtype diagnosis of gliomas is essential for appropriate patient management and for the interpretation of basic and clinical investigations. Diagnostic accuracy and reproducibility are compromised by the subjective histologic criteria currently used to classify and grade gliomas [9].

Nine out of 11 tumors with oligodendroglioma histopathological diagnosis were found to carry chromosome 1p/19q codeletion and *IDH1* mutation, genetic alterations characteristic of oligodendrogliomas. However, one tumor ODG10 lacking chromosome 1p/19q codeletion was found to carry a mutation in *IDH1*, *ATRX* and *TP53* gene. Mutations in the *ATRX* gene have been found to be restricted to *IDH1/IDH2* mutated gliomas, mutually exclusive with chromosome 1p/19q codeletion and correlating with astrocytic morphology [52], [53]. Therefore, although ODG10 was diagnosed as oligodendroglioma based on the characteristic histological appearance, it is identified as an astrocytoma based on the genetic alterations. One out of the 11 tumors lacking chromosome 1p/19q codeletion as well as *IDH1/IDH2* mutation lacked mutation in *ATRX* or *TP53* gene as well and hence cannot be classified as an oligodendroglioma or astrocytoma. This tumor was found to carry some of the genetic alterations that are known to occur in glioblastomas like *CDKN2A* deletion, *PDGFRA* amplification and mutation in *NF1* gene. Thus, based on the genetic alterations this tumor is closer to glioblastomas than low grade gliomas. Integrated DNA methylation and copy-number profiling study on a cohort of 228 anaplastic gliomas has also identified three similar molecular types [172]. The three subtypes consisted of a group of *IDH1/2* mutated CpG island Methylator Phenotype (CIMP) positive tumors with chr 1p/19q codeletion having the best prognosis, CIMP positive tumors lacking chr 1p/19q codeletion having intermediate prognosis (likely to correspond to astrocytomas) and glioblastoma-like CIMP negative tumors having copy number alterations similar to those in glioblastomas having the worst prognosis [172]. Thus, in addition to the histopathological characterization molecular characterization of adult gliomas is necessary for accurate diagnosis. Integrated diagnosis based on histopathology and molecular markers has been recently recommended for inclusion in the WHO guidelines for diagnosis of the central nervous system tumors by the International society of Neuropathologists [67].

The two large scale exome/genome sequencing of total 747 adult grade II/grade III gliomas have now conclusively demonstrated three distinct subtypes of gliomas based on the genetic alterations [103], [173]. One subtype (Type I) carries *IDH1/IDH2* mutation and chromosome 1p/19q codeletion. Majority of the tumors belonging to oligodendroglioma morphology belong to this molecular subtype. Type II subtype gliomas also carry *IDH1/IDH2* mutation, lack chromosome 1p/19q codeletion and carry mutations in *TP53* gene and/or *ATRX* gene. Majority of gliomas with astrocytic and mixed oligoastrocytic morphology belong to this subtype. The third subtype does not carry mutation in *IDH1/IDH2* gene and shows some of the genetic alterations characteristic of grade IV glioblastomas like loss of chromosome 9p, chromosome 10, mutation in *TP53/PTEN/RB/NF1* gene, deletion of *CDKN2A* locus, amplification of *EGFR/PDGFR/CDK4* etc [103]. Multiple biopsies from a single tumor also showed that the tumor is homogenous in the truncal alterations like mutation in *IDH1/IDH2* or *TP53/ATRX* gene and chromosome 1p/19q deletion although heterogeneity exists in mutations in other genes [173]. Loss of chromosome 1p/19q codeletion was found to be mutually exclusive with mutations in *TP53/ATRX* gene and these alterations belonging to the two tumor types were not found in the same tumor even in multiple biopsies from distinct locations or in recurrent tumors. The three subtypes of grade III histology show distinct differences in their overall survival with the Type I subtype have the best overall survival, followed by Type II tumors while Type III having the worst survival rate. Type III tumors can be described as glioblastoma-like and although have the worst survival rates among the three grade III subtypes overall survival is better than that of glioblastoma cases. ODG1 to ODG9 in our study cohort belong to the Type I glioma while ODG 10 belongs to Type II and ODG11 belongs to type III glioma while all 11 of them have pure oligodendroglioma morphology. Thus, genetic alterations based diagnosis of gliomas is superior to histopathology based diagnosis and has now been included in 2016 revision to the WHO classification of brain tumors.

Genetic alterations in oligodendrogliomas and their possible impact on the future treatment strategies

While oligodendrogliomas are known to be more sensitive to chemotherapy than astrocytomas, the tumors recur despite multimodal treatment including surgery, radiation and conventional chemotherapy [174]. Development of novel treatment strategies based on the knowledge of the underlying genetic alterations is necessary for the effective treatment of these tumors. The presence of mutations in chromatin modifier genes like *ARID1A* and Notch signaling pathway genes like *NOTCH1*, *MAML3* may make these tumors amenable to treatment using chromatin modifying drugs and Notch signaling modulators [175]. Higher expression of the oncogenic ETV transcription factors in the *CIC*-mutant oligodendrogliomas may make these tumors more aggressive than the *CIC*-wild type oligodendrogliomas. Loss of *CIC* expression has been reported to correlate with shorter progression free survival in a study done on 55 oligodendroglioma tumors [176]. Further, targeted treatment using RTK/RAS/MAPK inhibitors is likely to be effective only in the case of *CIC*-wild type oligodendrogliomas some of which were found to carry mutations in genes like *KRAS*, *NRAS*, *NF1* and *EGFR* that are involved in the RTK/RAS/MAPK signaling (Figure 5.1). Doxorubicin has been found to inhibit cancer cell proliferation by stimulating proteolytic cleavage of *CREB3L1* [177], a gene found to be upregulated in the *CIC*-mutant oligodendrogliomas. *CREB3L1* gene overexpression was shown to make human breast cancer and hepatoma cell lines sensitive to doxorubicin. While doxorubicin due to its poor ability to cross blood brain barrier, is not used for the treatment of oligodendrogliomas, its derivatives that can cross blood brain barrier may be effective in treatment of *CIC*-mutant oligodendrogliomas. Further validation of these findings using established oligodendroglioma cell lines and/or *in vivo* human tumor xenografts is necessary to translate these genetic alterations identified in

oligodendrogliomas into the development of effective targeted treatment strategies for this presently incurable malignant brain tumor.

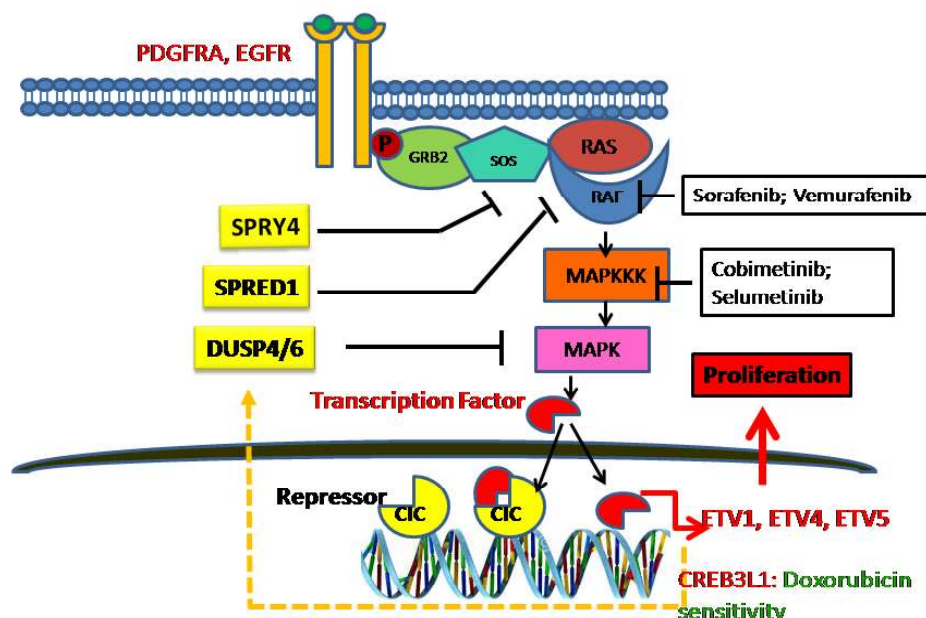


Figure 5.1: A schematic representation of RTK/RAS/MAPK pathway mediated control over downstream CIC repressor protein and inhibitors of components of RTK/RAS/MAPK pathway (Figure by Dr. Neelam Shirsat).

Targeted Sequencing as the most effective strategy for accurate molecular diagnosis of gliomas and identifying effective treatment strategy

Comprehensive methods to identify mutations such as exome sequencing are suitable for discovery cohort in studies where the aim is to identify frequent yet unknown genetic alterations in the given disease type. In low grade gliomas, it has been established that only few selective genes such as *IDH1*, *IDH2*, *TP53*, *ATRX*, *CIC*, *FUBP1*, *NOTCH1*, *NF1* etc. are affected by somatic mutations frequently among samples [47], [52], [103], [149] and every tumour sample has a combination of these selective genes somatically altered. Thus it is unnecessary to assess all 20,000 genes for each glioma sample. Evaluation of a panel of few selected frequently mutated genes is sufficient to help in molecular classification of gliomas. This may be more economical

than Exome sequencing. The results obtained from the AmpliSeq Cancer Hotspot Panel v2 on Ion Torrent platform indicate the suitability of this approach to find genetic alterations in the glioma tumour samples for routine molecular classification of the glioma samples.

SUMMARY AND CONCLUDING REMARKS

6 SUMMARY AND CONCLUDING REMARKS

Tumours of the central nervous system are a major cause of deaths resulting from cancer in children and adults. Gliomas are classified as astrocytic, oligodendroglial, or ependymal depending on the resemblance of tumour cell morphology to the three glial cell types *viz.* astrocytes, oligodendrocytes, and ependymal cells which are presumed to be cell of origin of these tumours. Oligodendroglioma is a histologically defined subtype of diffuse glioma along with Astrocytoma and Oligoastrocytoma. Oligodendrogliomas are graded as per WHO 2007 classification of brain tumours as oligodendroglioma (grade II) and anaplastic oligodendroglioma grade (III) [178]. Pure oligodendrogliomas are composed of single cell type while mixed oligoastrocytomas have morphological characteristics of both pure oligodendrogliomas and astrocytomas. Complete neurosurgical resection of diffuse gliomas is not possible because of their highly invasive nature [8]. The existence of residual tumour often results in recurrence and malignant progression. Along with surgical resection as primary treatment, chemotherapy and/or radiation therapy may be given as a post-surgical adjuvant therapy. Studies have suggested that astrocytomas show poor response to chemotherapy regimens, whereas oligodendrogliomas are sensitive to PCV [Procarbazine, Lomustine (CCNU), Vincristine] chemotherapy[14]–[18]. Oligodendrogliomas generally have slower growth rates and have better prognosis than astrocytomas of similar grade[20]. Patients with oligodendroglioma in general survive longer than patients with astrocytomas of similar grade[2]. Diagnosis of gliomas primarily depends on histopathologic analysis of H & E (Hematoxylin and Eosin) stained slides of tumour tissue sections [7], [8]. Histopathologic diagnosis is particularly challenging when it comes to reliably distinguishing between oligodendroglial and astrocytic component in non-classic low grade to intermediate grade gliomas [8]. Several studies have shown considerable intra-observer variation in the diagnosis of astrocytomas, oligodendrogliomas, and oligoastrocytomas [8], [9], [11], [12]. Accurate

pathologic diagnosis needs the ability to distinguish astrocytic from oligodendroglial differentiation in histologic sections. As treatment strategies and prognostication is different for different glioma subtypes accurate diagnosis of glioma subtypes is necessary for prognostication and designing treatment strategy. Specific molecular markers are needed for accurate diagnosis of glioma subgroups.

In order to identify molecular markers specific for oligodendrogliomas, Exome sequencing of glioma tumours histopathologically diagnosed as oligodendrogliomas was carried out. Exomes of eleven oligodendroglioma tumor tissue DNA and their paired normal blood DNA were sequenced by high throughput DNA sequencing on Illumina Hiseq 1500. Copy number variation analysis of the exome sequence data of the 11 tumors was done using two algorithms FishingCNV, and Control-FREEC taking into account both the coverage and minor allele frequency. The analysis identified concurrent loss of a copy of chromosome 1p and chromosome 19q, a known characteristic of oligodendrogliomas, in 9 out of the 11 tumor tissues. Other than the 1p/19q codeletion, recurrent chromosomal level copy number alteration was noted as chromosome 14 deletion in three chromosome 1p/19q codeleted tumors. The number of non-synonymous somatic variants per tumor exome was found to range from 10 to 46. R132H/R132C or R172K mutation in the *IDH1* or *IDH2* gene respectively was identified in 10 tumors while one tumor lacked mutation in the *IDH1* as well as *IDH2* gene. Four out of the 9 tumors having 1p/19q codeletion were found to carry a missense or a frame-shift deletion mutation in the *CIC* gene, located on chromosome 19q while two tumors carried mutations in *FUBP1* gene located on chromosome 1p. Two tumors with 1p/19q codeletion but no somatic alteration in the *CIC* gene were found to carry an activating mutation (Q61L, G12D) in *KRAS* gene. Recurrent mutations were identified in the Notch signaling pathway genes including four tumors with mutation in *NOTCH1* and, one tumor with a mutation in *MAML3* gene. Two tumors were found to carry a mutation in the chromatin modifier *ARID1A* gene. One tumor with

IDH1 mutation but without 1p/19q codeletion was found to have mutation in *TP53* and *ATRX* gene. One of the 11 tumors was found to lack mutation in *IDH1/IDH2* gene as well as chromosome 1p/19q codeletion. This *IDH* wild type tumor was found to carry mutation in *NF1* gene as well as amplification of *PDGFRA* gene and deletion of *CDKN2A* locus.

In the present study, four out of 9 tumors with chromosome 1p/19q codeletion were found to harbor mutation in the *CIC* gene. Transcriptome sequencing was performed on nine chromosome 1p-19q codeleted tumor tissues. *CIC* gene located on chromosome 19q13.2 encodes a HMG domain transcription factor. In order to understand role of the *CIC* gene in oligodendroglioma pathogenesis, differential gene expression analysis was carried out comparing the *CIC*-mutant (four tumors) tumor profile with that of *CIC*-wild type (3 tumors) *IDH1/2*-mutant, chromosome 1p/19q codeleted tumour tissues from study cohort of 7 tumors as well as for 65 tumours (39 *CIC*-mutant and 26 *CIC* wild type) from the TCGA cohort. The differential gene expression analysis in both the data sets revealed upregulation of *ETV/Pea3* family transcription factor-encoding genes *ETV1*, *ETV4* and *ETV5* in the *CIC*- mutant tumours. The gene set enrichment analysis identified a number of genes involved in the MAP kinase (MAPK) signaling pathway to be significantly enriched in the *CIC*-mutant oligodendrogliomas. These MAPK pathway genes included *DUSP4*, *DUSP6*, *DUSP19*, the dual specificity phosphatase genes, Sprouty family members *SPRY4*, *SPRED1* and *SPRED2*, and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, *EPHB*. Upregulation of these negative regulators of the RTK signaling pathway is likely to be due to the constitutive activation of the pathway resulting from the loss of function of the *CIC* repressor protein. In some of the oligodendrogliomas lacking mutations in the *CIC* gene, other components of the RTK/RAS/MAPK signaling pathway appear to be activated by inactivating mutations in the negative regulators of the pathway like *NF1* and by activating mutations in *KRAS*, *NRAS*, and

EGFR. Thus, activation of the RTK/RAS/MAPK signaling pathway appears to be a major driver of the oligodendroglioma pathogenesis.

Higher expression of oncogenic ETV transcription factors in the *CIC*-mutant oligodendrogliomas may make these tumours more aggressive than *CIC*-wild type tumours. Further, targeted treatment using RTK/RAS/MAPK inhibitors is likely to be effective only in the case of *CIC*-wild type oligodendrogliomas some of which were found to carry mutations in genes like *KRAS*, *NRAS*, *NF1* and *EGFR* that are involved in the RTK/RAS/MAPK signaling. Doxorubicin has been found to inhibit cancer cell proliferation by stimulating proteolytic cleavage of *CREB3L1*, a gene found to be upregulated in the *CIC*-mutant oligodendrogliomas. While doxorubicin due to its poor ability to cross blood brain barrier, is not used for the treatment of oligodendrogliomas, its derivatives that can cross blood brain barrier may be effective in treatment of *CIC*-mutant oligodendrogliomas. The presence of mutations in chromatin modifier genes like *ARID1A* and Notch signaling pathway genes like *NOTCH1*, *MAML3* may make these tumors amenable to treatment using chromatin modifying drugs and Notch signaling modulators.

Majority of the nonsynonymous mutations in the *CIC* gene (identified from present study cohort and TCGA cohort) were found to be specifically clustered on exon 5 and exon 19-20 unlike truncating mutations which occurred throughout the gene. Exon 5 encodes highly conserved HMG-box which functions as DNA binding domain. *In silico* analysis was performed to assess the effect of the nonsynonymous mutations in exon 5 on HMG domain. The analysis identified that the nonsynonymous mutations in exon 5 affected highly conserved amino acids of the HMG domain. X-ray crystallographically derived structure of *CIC* HMG domain is not available. Analysis of cocrystal structure of DNA bound -SOX17 HMG domain, which has high amino acid sequence and structural similarity to *CIC* HMG domain, revealed that the highly conserved amino acids which are frequently affected with nonsynonymous mutations in *CIC* HMG domain

were closest to the bound DNA in SOX 17 HMG domain cocrystal structure. This suggests that the exon 5 nonsynonymous mutations may affect the amino acids in HMG domain involved in DNA binding which indicates that loss of repressor function of the CIC protein as a result of nonsynonymous mutations in the *CIC* exon 5.

Mutation based stratification of copy number alterations was carried out on 251 low grade glioma (WHO grade II and III) tumour samples data from TCGA., for which both somatic mutation and copy number alteration data was available. The 251 samples segregated in three major molecular subgroups 1) *IDH1/2* mutant tumors with 1p/19q codeletion 2) *IDH1/2* mutant tumors without 1p/19q codeletion and 3) *IDH1/2* wild type tumors. *IDH1/2* wild type tumours exhibited mutations and copy number alterations similar to that of glioblastomas. These three tumor types were also found in the present study cohort that included only histologically pure oligodendroglioma tumors. Nine tumours with oligodendroglioma morphology carried *IDH* mutation with 1p/19q codeletion. Two tumours with oligodendroglioma morphology showed molecular alterations that are known to be associated primarily with astrocytoma (*IDH* mutation along with *TP53* and *ATRX* mutation in absence of 1p/19q codeletion) and glioblastoma (*IDH1/2* wild type, *NF1* mutation, *PDGFRA* amplification and *CDKN2A* deletion). In the TCGA cohort while most histopathologically diagnosed oligodendrogliomas belonged to *IDH* mutated 1p/19q codeleted group, and most histopathologically diagnosed astrocytomas belonged to molecular *IDH* mutated without 1p/19q codeleted group and *IDH* wild type group, a considerable histological heterogeneity was found in each molecular subtype.

In conclusion, *CIC*-mutant 1p/19q codeleted oligodendrogliomas show upregulation of the ETV/Pea3 transcription factor family genes and the negative regulators of tyrosine kinase receptor signaling pathway. Upregulation of the negative regulators of the RTK pathway is likely to be due to the constitutive activation of the pathway resulting from the loss of function of the CIC repressor protein. Higher expression of oncogenic ETV transcription factors in the

CIC-mutant oligodendrogliomas may make these tumours more aggressive than *CIC*-wild type tumours. Nonsynonymous mutations in *CIC* exon 5 affect highly conserved amino acids in the DNA binding HMG domain which is indicated to lead to loss of repressor function of *CIC* protein. Targeted treatment using RTK/RAS/MAPK inhibitors is likely to be effective only in the case of *CIC*-wild type oligodendrogliomas some of which were found to carry mutations in genes like *KRAS*, *NRAS*, *NF1* and *EGFR* that are involved in the RTK/RAS/MAPK signaling. Molecular marker based classification was found to be superior than histology based classification of grade II-III gliomas. Molecular marker based classification segregated the grade II-III gliomas in three major classes 1) *IDH* mutated with 1p/19q codeletion, 2) *IDH* mutated without 1p/19q codeletion and 3) *IDH* wild type. Molecular alterations provide objective criteria for glioma classification which can resolve the dilemma in glioma subtype diagnosis.

Based on the mutations, copy number alterations and comparative transcriptome analysis the study indicates *RTK/RAS/MAPK* pathway activation as a driver of Oligodendroglioma pathogenesis and suggest various possibilities of targeted treatment options based on the presence of identified alterations in the *CIC* as well as NOTCH signaling pathway and chromatin modifier genes. Further the study shows the necessity of molecular classification of gliomas for accurate diagnosis wherein three distinct molecular subtypes need to be identified based upon the presence of *IDH1/IDH2* mutation and 1p/19q codeletion.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] T. A. Dolecek, J. M. Propp, N. E. Stroup, and C. Kruchko, “CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009,” *Neuro-Oncol.*, vol. 14 Suppl 5, pp. v1-49, Nov. 2012.
- [2] V. P. Collins, “Brain tumours: classification and genes,” *J. Neurol. Neurosurg. Psychiatry*, vol. 75 Suppl 2, pp. ii2-11, Jun. 2004.
- [3] T. D. Bourne and D. Schiff, “Update on molecular findings, management and outcome in low-grade gliomas,” *Nat. Rev. Neurol.*, vol. 6, no. 12, pp. 695–701, Dec. 2010.
- [4] D. N. Louis *et al.*, “The 2007 WHO classification of tumours of the central nervous system,” *Acta Neuropathol. (Berl.)*, vol. 114, no. 2, pp. 97–109, Aug. 2007.
- [5] C. L. Gladson, R. A. Prayson, and W. (Michael) Liu, “The Pathobiology of Glioma Tumors,” *Annu. Rev. Pathol.*, vol. 5, pp. 33–50, 2010.
- [6] Q. T. Ostrom *et al.*, “CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009-2013,” *Neuro-Oncol.*, vol. 18, no. suppl_5, pp. v1–v75, Oct. 2016.
- [7] P. Bailey and P. C. Bucy, “Oligodendrogliomas of the brain,” *J. Pathol. Bacteriol.*, vol. 32, no. 4, pp. 735–751, Oct. 1929.
- [8] M. Gupta, A. Djalilvand, and D. J. Brat, “Clarifying the diffuse gliomas: an update on the morphologic features and markers that discriminate oligodendroglioma from astrocytoma,” *Am. J. Clin. Pathol.*, vol. 124, no. 5, pp. 755–768, Nov. 2005.
- [9] S. W. Coons, P. C. Johnson, B. W. Scheithauer, A. J. Yates, and D. K. Pearl, “Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary gliomas,” *Cancer*, vol. 79, no. 7, pp. 1381–1393, Apr. 1997.

- [10] M. J. van den Bent, "Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective," *Acta Neuropathol. (Berl.)*, vol. 120, no. 3, pp. 297–304, Sep. 2010.
- [11] C. E. Fuller *et al.*, "Clinical utility of fluorescence in situ hybridization (FISH) in morphologically ambiguous gliomas with hybrid oligodendroglial/astrocytic features," *J. Neuropathol. Exp. Neurol.*, vol. 62, no. 11, pp. 1118–1128, Nov. 2003.
- [12] M. A. Mittler, B. C. Walters, and E. G. Stopa, "Observer reliability in histological grading of astrocytoma stereotactic biopsies," *J. Neurosurg.*, vol. 85, no. 6, pp. 1091–1094, Dec. 1996.
- [13] C. Watts and N. Sanai, "Surgical approaches for the gliomas," in *Handbook of Clinical Neurology*, vol. 134, Elsevier, 2016, pp. 51–69.
- [14] J. G. Cairncross *et al.*, "Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas," *J. Natl. Cancer Inst.*, vol. 90, no. 19, pp. 1473–1479, Oct. 1998.
- [15] D. Fortin, G. J. Cairncross, and R. R. Hammond, "Oligodendroglioma: an appraisal of recent data pertaining to diagnosis and treatment," *Neurosurgery*, vol. 45, no. 6, p. 1279–1291; discussion 191, Dec. 1999.
- [16] D. Fortin, D. R. Macdonald, L. Stitt, and J. G. Cairncross, "PCV for oligodendroglial tumors: in search of prognostic factors for response and survival," *Can. J. Neurol. Sci. J. Can. Sci. Neurol.*, vol. 28, no. 3, pp. 215–223, Aug. 2001.
- [17] P. Pytel and R. V. Lukas, "Update on diagnostic practice: tumors of the nervous system," *Arch. Pathol. Lab. Med.*, vol. 133, no. 7, pp. 1062–1077, Jul. 2009.
- [18] M. J. van den Bent *et al.*, "Phase II study of first-line chemotherapy with temozolomide in recurrent oligodendroglial tumors: the European Organization for Research and

- Treatment of Cancer Brain Tumor Group Study 26971,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 21, no. 13, pp. 2525–2528, Jul. 2003.
- [19] J. E. C. Bromberg and M. J. van den Bent, “Oligodendrogliomas: molecular biology and treatment,” *The Oncologist*, vol. 14, no. 2, pp. 155–163, Feb. 2009.
- [20] M. J. van den Bent, “Advances in the biology and treatment of oligodendrogliomas,” *Curr. Opin. Neurol.*, vol. 17, no. 6, pp. 675–680, Dec. 2004.
- [21] Cancer Genome Atlas Research Network, “Comprehensive genomic characterization defines human glioblastoma genes and core pathways,” *Nature*, vol. 455, no. 7216, pp. 1061–1068, Oct. 2008.
- [22] H. Yan *et al.*, “IDH1 and IDH2 mutations in gliomas,” *N. Engl. J. Med.*, vol. 360, no. 8, pp. 765–773, Feb. 2009.
- [23] H. H. Engelhard, A. Stelea, and E. J. Cochran, “Oligodendroglioma: pathology and molecular biology,” *Surg. Neurol.*, vol. 58, no. 2, p. 111–117; discussion 117, Aug. 2002.
- [24] M. R. Gilbert and F. F. Lang, “Anaplastic oligodendroglial tumors: a tale of two trials,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 24, no. 18, pp. 2689–2690, Jun. 2006.
- [25] M. E. Alonso *et al.*, “Mutational study of the 1p located genes p18ink4c, Patched-2, RIZ1 and KIF1B in oligodendrogliomas,” *Oncol. Rep.*, vol. 13, no. 3, pp. 539–542, Mar. 2005.
- [26] M. Benetkiewicz *et al.*, “NOTCH2 is neither rearranged nor mutated in t(1;19) positive oligodendrogliomas,” *PloS One*, vol. 4, no. 1, p. e4107, 2009.
- [27] T. Tsujimoto *et al.*, “The p73 gene is not mutated in oligodendrogliomas which frequently have a deleted region at chromosome 1p36.3,” *Anticancer Res.*, vol. 20, no. 4, pp. 2495–2497, Aug. 2000.
- [28] M. C. M. Kouwenhoven *et al.*, “Molecular analysis of anaplastic oligodendroglial tumors in a prospective randomized study: A report from EORTC study 26951,” *Neuro-Oncol.*, vol. 11, no. 6, pp. 737–746, Dec. 2009.

- [29] L. A. M. Gravendeel *et al.*, “Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology,” *Cancer Res.*, vol. 69, no. 23, pp. 9065–9072, Dec. 2009.
- [30] F. Ducray *et al.*, “Anaplastic oligodendrogliomas with 1p19q codeletion have a proneural gene expression profile,” *Mol. Cancer*, vol. 7, p. 41, May 2008.
- [31] G. Cairncross and R. Jenkins, “Gliomas with 1p/19q codeletion: a.k.a. oligodendroglioma,” *Cancer J. Sudbury Mass*, vol. 14, no. 6, pp. 352–357, Dec. 2008.
- [32] M. J. Walter, T. A. Graubert, J. F. Dipersio, E. R. Mardis, R. K. Wilson, and T. J. Ley, “Next-generation sequencing of cancer genomes: back to the future,” *Pers. Med.*, vol. 6, no. 6, p. 653, Nov. 2009.
- [33] P. Bailey and H. Cushing, *A Classification of the Tumours of the Glioma Group on a Histogenetic Basis With a Correlated Study of Prognosis*. Philadelphia: J. B. Lippincott Co, 1926.
- [34] M. Vitucci, D. N. Hayes, and C. R. Miller, “Gene expression profiling of gliomas: merging genomic and histopathological classification for personalised therapy,” *Br. J. Cancer*, vol. 104, no. 4, pp. 545–553, Feb. 2011.
- [35] D. J. MacKenzie, “A Classification of the Tumours of the Glioma Group on a Histogenetic Basis With a Correlated Study of Prognosis,” *Can. Med. Assoc. J.*, vol. 16, no. 7, p. 872, Jul. 1926.
- [36] K. J. Zulch, Ed., *Histological typing of tumors of the central nervous system*. Geneva: World Health Organization, 1979.
- [37] P. Kleihues, P. C. Burger, and B. W. Scheithauer, Eds., *Histological typing of tumours of the central nervous system. World Health Organization international histological classification of tumours*. Heidelberg: Springer, 1993.

- [38] P. Kleihues and W. K. Cavenee, Eds., *World Health Organization Classification of Tumours. Pathology and genetics of tumours of the nervous system*. Lyon: IARC Press, 2000.
- [39] D. N. Louis, H. Ohgaki, O. D. Wiestler, and W. K. Cavenee, Eds., *WHO Classification of tumours of the central nervous system*. Lyon: International Agency for Research on Cancer, 2007.
- [40] D. N. Louis *et al.*, “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary,” *Acta Neuropathol. (Berl.)*, vol. 131, no. 6, pp. 803–820, Jun. 2016.
- [41] F. J. Rodriguez, K. S. Lim, D. Bowers, and C. G. Eberhart, “Pathological and molecular advances in pediatric low-grade astrocytoma,” *Annu. Rev. Pathol.*, vol. 8, pp. 361–379, Jan. 2013.
- [42] A. Olar and E. P. Sulman, “Molecular Markers in Low-Grade Glioma-Toward Tumor Reclassification,” *Semin. Radiat. Oncol.*, vol. 25, no. 3, pp. 155–163, Jul. 2015.
- [43] C. L. Appin and D. J. Brat, “Biomarker-driven diagnosis of diffuse gliomas,” *Mol. Aspects Med.*, vol. 45, pp. 87–96, Nov. 2015.
- [44] M. J. van den Bent, T. J. Snijders, and J. E. C. Bromberg, “Current treatment of low grade gliomas,” *Memo - Mag. Eur. Med. Oncol.*, vol. 5, no. 3, pp. 223–227, Sep. 2012.
- [45] C. Leighton *et al.*, “Supratentorial low-grade glioma in adults: an analysis of prognostic factors and timing of radiation,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 15, no. 4, pp. 1294–1301, Apr. 1997.
- [46] G. Cairncross *et al.*, “Chemotherapy for anaplastic oligodendroglioma. National Cancer Institute of Canada Clinical Trials Group,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 12, no. 10, pp. 2013–2021, Oct. 1994.

- [47] C. Bettegowda *et al.*, “Mutations in CIC and FUBP1 Contribute to Human Oligodendroglioma,” *Science*, vol. 333, no. 6048, pp. 1453–1455, Sep. 2011.
- [48] P. J. Killela *et al.*, “TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 15, pp. 6021–6026, Apr. 2013.
- [49] F. Sahm *et al.*, “CIC and FUBP1 mutations in oligodendrogliomas, oligoastrocytomas and astrocytomas,” *Acta Neuropathol. (Berl.)*, vol. 123, no. 6, pp. 853–860, Jun. 2012.
- [50] S. Yip *et al.*, “Concurrent CIC mutations, IDH mutations, and 1p/19q loss distinguish oligodendrogliomas from other cancers,” *J. Pathol.*, vol. 226, no. 1, pp. 7–16, Jan. 2012.
- [51] E. Alshail, J. T. Rutka, L. E. Becker, and H. J. Hoffman, “Optic chiasmatic-hypothalamic glioma,” *Brain Pathol. Zurich Switz.*, vol. 7, no. 2, pp. 799–806, Apr. 1997.
- [52] Y. Jiao *et al.*, “Frequent ATRX, CIC, FUBP1 and IDH1 mutations refine the classification of malignant gliomas,” *Oncotarget*, vol. 3, no. 7, pp. 709–722, Jul. 2012.
- [53] K. Kannan *et al.*, “Whole-exome sequencing identifies ATRX mutation as a key molecular determinant in lower-grade glioma,” *Oncotarget*, vol. 3, no. 10, pp. 1194–1203, Oct. 2012.
- [54] X.-Y. Liu *et al.*, “Frequent ATRX mutations and loss of expression in adult diffuse astrocytic tumors carrying IDH1/IDH2 and TP53 mutations,” *Acta Neuropathol. (Berl.)*, vol. 124, no. 5, pp. 615–625, Nov. 2012.
- [55] B. M. McCormack, D. C. Miller, G. N. Budzilovich, G. J. Voorhees, and J. Ransohoff, “Treatment and survival of low-grade astrocytoma in adults--1977-1988,” *Neurosurgery*, vol. 31, no. 4, p. 636–642; discussion 642, Oct. 1992.
- [56] H. Ohgaki and P. Kleihues, “Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas,” *J. Neuropathol. Exp. Neurol.*, vol. 64, no. 6, pp. 479–489, Jun. 2005.

- [57] J. Balss, J. Meyer, W. Mueller, A. Korshunov, C. Hartmann, and A. von Deimling, "Analysis of the IDH1 codon 132 mutation in brain tumors," *Acta Neuropathol. (Berl.)*, vol. 116, no. 6, pp. 597–602, Dec. 2008.
- [58] K. Ichimura, "Molecular pathogenesis of IDH mutations in gliomas," *Brain Tumor Pathol.*, vol. 29, no. 3, pp. 131–139, Jul. 2012.
- [59] K. Masui, T. F. Cloughesy, and P. S. Mischel, "Molecular pathology in adult high-grade gliomas: from molecular diagnostics to target therapies," *Neuropathol. Appl. Neurobiol.*, vol. 38, no. 3, pp. 271–291, Jun. 2012.
- [60] H. Ohgaki and P. Kleihues, "Genetic alterations and signaling pathways in the evolution of gliomas," *Cancer Sci.*, vol. 100, no. 12, pp. 2235–2241, Dec. 2009.
- [61] M. J. Riemenschneider, M. E. Hegi, and G. Reifenberger, "MGMT promoter methylation in malignant gliomas," *Target. Oncol.*, vol. 5, no. 3, pp. 161–165, Sep. 2010.
- [62] M. Weller *et al.*, "MGMT promoter methylation in malignant gliomas: ready for personalized medicine?," *Nat. Rev. Neurol.*, vol. 6, no. 1, pp. 39–51, Jan. 2010.
- [63] C. W. Brennan *et al.*, "The somatic genomic landscape of glioblastoma," *Cell*, vol. 155, no. 2, pp. 462–477, Oct. 2013.
- [64] J. He *et al.*, "Glioblastomas with an oligodendroglial component: a pathological and molecular study," *J. Neuropathol. Exp. Neurol.*, vol. 60, no. 9, pp. 863–871, Sep. 2001.
- [65] M. Salvati *et al.*, "Cerebral glioblastoma with oligodendroglial component: analysis of 36 cases," *J. Neurooncol.*, vol. 94, no. 1, pp. 129–134, Aug. 2009.
- [66] D. Vordermark *et al.*, "Glioblastoma multiforme with oligodendroglial component (GBMO): favorable outcome after post-operative radiotherapy and chemotherapy with nimustine (ACNU) and teniposide (VM26)," *BMC Cancer*, vol. 6, p. 247, Oct. 2006.

- [67] D. N. Louis *et al.*, “International Society of Neuropathology-Haarlem Consensus Guidelines for Nervous System Tumor Classification and Grading,” *Brain Pathol.*, vol. 24, no. 5, pp. 429–435, Sep. 2014.
- [68] X. Xu *et al.*, “Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity,” *J. Biol. Chem.*, vol. 279, no. 32, pp. 33946–33957, Aug. 2004.
- [69] D. W. Parsons *et al.*, “An integrated genomic analysis of human glioblastoma multiforme,” *Science*, vol. 321, no. 5897, pp. 1807–1812, Sep. 2008.
- [70] L. Dang *et al.*, “Cancer-associated IDH1 mutations produce 2-hydroxyglutarate,” *Nature*, vol. 462, no. 7274, pp. 739–744, Dec. 2009.
- [71] M. E. Figueroa *et al.*, “Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation,” *Cancer Cell*, vol. 18, no. 6, pp. 553–567, Dec. 2010.
- [72] S. Turcan *et al.*, “IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype,” *Nature*, vol. 483, no. 7390, pp. 479–483, Feb. 2012.
- [73] P. Mur *et al.*, “Codeletion of 1p and 19q determines distinct gene methylation and expression profiles in IDH-mutated oligodendroglial tumors,” *Acta Neuropathol. (Berl.)*, vol. 126, no. 2, pp. 277–289, Aug. 2013.
- [74] H. Noushmehr *et al.*, “Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma,” *Cancer Cell*, vol. 17, no. 5, pp. 510–522, May 2010.
- [75] J. Reifenberger, G. Reifenberger, L. Liu, C. D. James, W. Wechsler, and V. P. Collins, “Molecular genetic analysis of oligodendroglial tumors shows preferential allelic deletions on 19q and 1p,” *Am. J. Pathol.*, vol. 145, no. 5, pp. 1175–1190, Nov. 1994.

- [76] R. B. Jenkins *et al.*, “A t(1;19)(q10;p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma,” *Cancer Res.*, vol. 66, no. 20, pp. 9852–9861, Oct. 2006.
- [77] J. S. Smith *et al.*, “Alterations of chromosome arms 1p and 19q as predictors of survival in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 18, no. 3, pp. 636–645, Feb. 2000.
- [78] C. A. Griffin *et al.*, “Identification of der(1;19)(q10;p10) in five oligodendrogliomas suggests mechanism of concurrent 1p and 19q loss,” *J. Neuropathol. Exp. Neurol.*, vol. 65, no. 10, pp. 988–994, Oct. 2006.
- [79] C. Giannini *et al.*, “Anaplastic oligodendroglial tumors: refining the correlation among histopathology, 1p 19q deletion and clinical outcome in Intergroup Radiation Therapy Oncology Group Trial 9402,” *Brain Pathol. Zurich Switz.*, vol. 18, no. 3, pp. 360–369, Jul. 2008.
- [80] A. A. Kanner *et al.*, “The impact of genotype on outcome in oligodendroglioma: validation of the loss of chromosome arm 1p as an important factor in clinical decision making,” *J. Neurosurg.*, vol. 104, no. 4, pp. 542–550, Apr. 2006.
- [81] K. J. Hatanpaa, P. C. Burger, J. R. Eshleman, K. M. Murphy, and K. D. Berg, “Molecular diagnosis of oligodendroglioma in paraffin sections,” *Lab. Invest. J. Tech. Methods Pathol.*, vol. 83, no. 3, pp. 419–428, Mar. 2003.
- [82] G. Reifenberger and D. N. Louis, “Oligodendroglioma: toward molecular definitions in diagnostic neuro-oncology,” *J. Neuropathol. Exp. Neurol.*, vol. 62, no. 2, pp. 111–126, Feb. 2003.
- [83] M. Labussière *et al.*, “TERT promoter mutations in gliomas, genetic associations and clinico-pathological correlations,” *Br. J. Cancer*, vol. 111, no. 10, pp. 2024–2032, Nov. 2014.

- [84] M. J. van den Bent *et al.*, “A hypermethylated phenotype is a better predictor of survival than MGMT methylation in anaplastic oligodendroglial brain tumors: a report from EORTC study 26951,” *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 17, no. 22, pp. 7148–7155, Nov. 2011.
- [85] J. G. Cairncross *et al.*, “Benefit from procarbazine, lomustine, and vincristine in oligodendroglial tumors is associated with mutation of IDH,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 32, no. 8, pp. 783–790, Mar. 2014.
- [86] C. Hartmann *et al.*, “Type and frequency of IDH1 and IDH2 mutations are related to astrocytic and oligodendroglial differentiation and age: a study of 1,010 diffuse gliomas,” *Acta Neuropathol. (Berl.)*, vol. 118, no. 4, pp. 469–474, Oct. 2009.
- [87] C. Houillier *et al.*, “IDH1 or IDH2 mutations predict longer survival and response to temozolomide in low-grade gliomas,” *Neurology*, vol. 75, no. 17, pp. 1560–1566, Oct. 2010.
- [88] Intergroup Radiation Therapy Oncology Group Trial 9402 *et al.*, “Phase III trial of chemotherapy plus radiotherapy compared with radiotherapy alone for pure and mixed anaplastic oligodendroglioma: Intergroup Radiation Therapy Oncology Group Trial 9402,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 24, no. 18, pp. 2707–2714, Jun. 2006.
- [89] M. J. van den Bent *et al.*, “Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 31, no. 3, pp. 344–350, Jan. 2013.
- [90] G. Cairncross *et al.*, “Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 31, no. 3, pp. 337–343, Jan. 2013.

- [91] A. B. Lassman *et al.*, “International retrospective study of over 1000 adults with anaplastic oligodendroglial tumors,” *Neuro-Oncol.*, vol. 13, no. 6, pp. 649–659, Jun. 2011.
- [92] R. Stupp *et al.*, “Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma,” *N. Engl. J. Med.*, vol. 352, no. 10, pp. 987–996, Mar. 2005.
- [93] A. M. Donson, S. O. Addo-Yobo, M. H. Handler, L. Gore, and N. K. Foreman, “MGMT promoter methylation correlates with survival benefit and sensitivity to temozolomide in pediatric glioblastoma,” *Pediatr. Blood Cancer*, vol. 48, no. 4, pp. 403–407, Apr. 2007.
- [94] M. E. Hegi *et al.*, “MGMT gene silencing and benefit from temozolomide in glioblastoma,” *N. Engl. J. Med.*, vol. 352, no. 10, pp. 997–1003, Mar. 2005.
- [95] A. A. Brandes *et al.*, “Correlations between O6-methylguanine DNA methyltransferase promoter methylation status, 1p and 19q deletions, and response to temozolomide in anaplastic and recurrent oligodendroglioma: a prospective GICNO study,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 24, no. 29, pp. 4746–4753, Oct. 2006.
- [96] S. M. Dong *et al.*, “Concurrent hypermethylation of multiple genes is associated with grade of oligodendroglial tumors,” *J. Neuropathol. Exp. Neurol.*, vol. 60, no. 8, pp. 808–816, Aug. 2001.
- [97] M. Möllemann, M. Wolter, J. Felsberg, V. P. Collins, and G. Reifenberger, “Frequent promoter hypermethylation and low expression of the MGMT gene in oligodendroglial tumors,” *Int. J. Cancer*, vol. 113, no. 3, pp. 379–385, Jan. 2005.
- [98] J. E. Eckel-Passow *et al.*, “Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors,” *N. Engl. J. Med.*, vol. 372, no. 26, pp. 2499–2508, Jun. 2015.
- [99] S. Horn *et al.*, “TERT promoter mutations in familial and sporadic melanoma,” *Science*, vol. 339, no. 6122, pp. 959–961, Feb. 2013.

- [100] F. W. Huang, E. Hodis, M. J. Xu, G. V. Kryukov, L. Chin, and L. A. Garraway, “Highly recurrent TERT promoter mutations in human melanoma,” *Science*, vol. 339, no. 6122, pp. 957–959, Feb. 2013.
- [101] Y. Okamoto *et al.*, “Population-based study on incidence, survival rates, and genetic alterations of low-grade diffuse astrocytomas and oligodendrogliomas,” *Acta Neuropathol. (Berl.)*, vol. 108, no. 1, pp. 49–56, Jul. 2004.
- [102] T. Watanabe, S. Nobusawa, P. Kleihues, and H. Ohgaki, “IDH1 mutations are early events in the development of astrocytomas and oligodendrogliomas,” *Am. J. Pathol.*, vol. 174, no. 4, pp. 1149–1153, Apr. 2009.
- [103] Cancer Genome Atlas Research Network *et al.*, “Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas,” *N. Engl. J. Med.*, vol. 372, no. 26, pp. 2481–2498, Jun. 2015.
- [104] G. Jiménez, A. Guichet, A. Ephrussi, and J. Casanova, “Relief of gene repression by torso RTK signaling: role of capicua in Drosophila terminal and dorsoventral patterning,” *Genes Dev.*, vol. 14, no. 2, pp. 224–231, Jan. 2000.
- [105] L. Ajuria *et al.*, “Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila,” *Development*, vol. 138, no. 5, pp. 915–924, Mar. 2011.
- [106] M. Kawamura-Saito *et al.*, “Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation,” *Hum. Mol. Genet.*, vol. 15, no. 13, pp. 2125–2137, Jul. 2006.
- [107] F. Roch, G. Jiménez, and J. Casanova, “EGFR signalling inhibits Capicua-dependent repression during specification of Drosophila wing veins,” *Dev. Camb. Engl.*, vol. 129, no. 4, pp. 993–1002, Feb. 2002.

- [108] A.-S. K. Tseng *et al.*, “Capicua regulates cell proliferation downstream of the receptor tyrosine kinase/ras signaling pathway,” *Curr. Biol. CB*, vol. 17, no. 8, pp. 728–733, Apr. 2007.
- [109] K. Dissanayake *et al.*, “ERK/p90(RSK)/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicúa,” *Biochem. J.*, vol. 433, no. 3, pp. 515–525, Feb. 2011.
- [110] R. Duncan *et al.*, “A sequence-specific, single-strand binding protein activates the far upstream element of c-myc and defines a new DNA-binding motif,” *Genes Dev.*, vol. 8, no. 4, pp. 465–480, Feb. 1994.
- [111] J. M. Bruner, L. Inouye, G. N. Fuller, and L. A. Langford, “Diagnostic discrepancies and their clinical impact in a neuropathology referral practice,” *Cancer*, vol. 79, no. 4, pp. 796–803, Feb. 1997.
- [112] D. Gorovets *et al.*, “IDH mutation and neuroglial developmental features define clinically distinct subclasses of lower grade diffuse astrocytic glioma,” *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 18, no. 9, pp. 2490–2501, May 2012.
- [113] A. Li *et al.*, “Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes,” *Cancer Res.*, vol. 69, no. 5, pp. 2091–2099, Mar. 2009.
- [114] W. Yan *et al.*, “Molecular classification of gliomas based on whole genome gene expression: a systematic report of 225 samples from the Chinese Glioma Cooperative Group,” *Neuro-Oncol.*, vol. 14, no. 12, pp. 1432–1440, Dec. 2012.
- [115] T. Sørlie *et al.*, “Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 19, pp. 10869–10874, Sep. 2001.
- [116] P. J. M. Valk *et al.*, “Prognostically useful gene-expression profiles in acute myeloid leukemia,” *N. Engl. J. Med.*, vol. 350, no. 16, pp. 1617–1628, Apr. 2004.

- [117] P. J. French *et al.*, “Gene expression profiles associated with treatment response in oligodendrogliomas,” *Cancer Res.*, vol. 65, no. 24, pp. 11335–11344, Dec. 2005.
- [118] H. S. Phillips *et al.*, “Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis,” *Cancer Cell*, vol. 9, no. 3, pp. 157–173, Mar. 2006.
- [119] R. G. W. Verhaak *et al.*, “Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1,” *Cancer Cell*, vol. 17, no. 1, pp. 98–110, Jan. 2010.
- [120] E. A. Howe, R. Sinha, D. Schlauch, and J. Quackenbush, “RNA-Seq analysis in MeV,” *Bioinformatics*, vol. 27, no. 22, pp. 3209–3210, Nov. 2011.
- [121] A. Edwards, A. Civitello, H. A. Hammond, and C. T. Caskey, “DNA typing and genetic mapping with trimeric and tetrameric tandem repeats,” *Am. J. Hum. Genet.*, vol. 49, no. 4, pp. 746–756, Oct. 1991.
- [122] J. L. Weber and P. E. May, “Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction,” *Am. J. Hum. Genet.*, vol. 44, no. 3, pp. 388–396, Mar. 1989.
- [123] Y. Shi and J. Majewski, “FishingCNV: a graphical software package for detecting rare copy number variations in exome-sequencing data,” *Bioinforma. Oxf. Engl.*, vol. 29, no. 11, pp. 1461–1462, Jun. 2013.
- [124] V. Boeva *et al.*, “Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data,” *Bioinforma. Oxf. Engl.*, vol. 28, no. 3, pp. 423–425, Feb. 2012.
- [125] H. Li and R. Durbin, “Fast and accurate short read alignment with Burrows-Wheeler transform,” *Bioinforma. Oxf. Engl.*, vol. 25, no. 14, pp. 1754–1760, Jul. 2009.

- [126] D. C. Koboldt *et al.*, “VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing,” *Genome Res.*, vol. 22, no. 3, pp. 568–576, Mar. 2012.
- [127] K. Wang, M. Li, and H. Hakonarson, “ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data,” *Nucleic Acids Res.*, vol. 38, no. 16, p. e164, Sep. 2010.
- [128] G. Jiménez, S. Y. Shvartsman, and Z. ’ev Paroush, “The Capicua repressor—a general sensor of RTK signaling in development and disease,” *J. Cell Sci.*, vol. 125, no. Pt 6, pp. 1383–1391, Mar. 2012.
- [129] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, “edgeR: a Bioconductor package for differential expression analysis of digital gene expression data,” *Bioinforma. Oxf. Engl.*, vol. 26, no. 1, pp. 139–140, Jan. 2010.
- [130] S. Anders and W. Huber, “Differential expression analysis for sequence count data,” *Genome Biol.*, vol. 11, no. 10, p. R106, 2010.
- [131] X. Wang and M. J. Cairns, “SeqGSEA: a Bioconductor package for gene set enrichment analysis of RNA-Seq data integrating differential expression and splicing,” *Bioinforma. Oxf. Engl.*, vol. 30, no. 12, pp. 1777–1779, Jun. 2014.
- [132] A. Subramanian *et al.*, “Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 43, pp. 15545–15550, Oct. 2005.
- [133] V. Padul, S. Epari, A. Moiyadi, P. Shetty, and N. V. Shirsat, “ETV/Pea3 family transcription factor-encoding genes are overexpressed in CIC-mutant oligodendrogliomas,” *Genes. Chromosomes Cancer*, vol. 54, no. 12, pp. 725–733, Dec. 2015.
- [134] J. T. Robinson *et al.*, “Integrative genomics viewer,” *Nat. Biotechnol.*, vol. 29, no. 1, pp. 24–26, Jan. 2011.

- [135] H. Thorvaldsdóttir, J. T. Robinson, and J. P. Mesirov, “Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration,” *Brief. Bioinform.*, vol. 14, no. 2, pp. 178–192, Mar. 2013.
- [136] J. G. Cairncross and D. R. Macdonald, “Successful chemotherapy for recurrent malignant oligodendroglioma,” *Ann. Neurol.*, vol. 23, no. 4, pp. 360–364, Apr. 1988.
- [137] T. Nagasaka *et al.*, “FISH 1p/19q deletion/imbalance for molecular subclassification of glioblastoma,” *Brain Tumor Pathol.*, vol. 24, no. 1, pp. 1–5, 2007.
- [138] R. Stupp and M. E. Hegi, “Neuro-oncology: oligodendroglioma and molecular markers,” *Lancet Neurol.*, vol. 6, no. 1, pp. 10–12, Jan. 2007.
- [139] A. Perry, C. E. Fuller, R. Banerjee, D. J. Brat, and B. W. Scheithauer, “Ancillary FISH analysis for 1p and 19q status: preliminary observations in 287 gliomas and oligodendroglioma mimics,” *Front. Biosci. J. Virtual Libr.*, vol. 8, pp. a1-9, Jan. 2003.
- [140] K. S. Reddy, “Assessment of 1p/19q deletions by fluorescence in situ hybridization in gliomas,” *Cancer Genet. Cytogenet.*, vol. 184, no. 2, pp. 77–86, Jul. 2008.
- [141] L. Cawkwell, S. M. Bell, F. A. Lewis, M. F. Dixon, G. R. Taylor, and P. Quirke, “Rapid detection of allele loss in colorectal tumours using microsatellites and fluorescent DNA technology,” *Br. J. Cancer*, vol. 67, no. 6, pp. 1262–1267, Jun. 1993.
- [142] P. Jha *et al.*, “Detection of allelic status of 1p and 19q by microsatellite-based PCR versus FISH: limitations and advantages in application to patient management,” *Diagn. Mol. Pathol. Am. J. Surg. Pathol. Part B*, vol. 20, no. 1, pp. 40–47, Mar. 2011.
- [143] M. D. Johnson, C. L. Vnencak-Jones, S. A. Toms, P. M. Moots, and R. Weil, “Allelic losses in oligodendroglial and oligodendroglioma-like neoplasms: analysis using microsatellite repeats and polymerase chain reaction,” *Arch. Pathol. Lab. Med.*, vol. 127, no. 12, pp. 1573–1579, Dec. 2003.

- [144] C. Ramirez *et al.*, “Loss of 1p, 19q, and 10q heterozygosity prospectively predicts prognosis of oligodendroglial tumors--towards individualized tumor treatment?,” *Neuro-Oncol.*, vol. 12, no. 5, pp. 490–499, May 2010.
- [145] S. Artavanis-Tsakonas, M. D. Rand, and R. J. Lake, “Notch signaling: cell fate control and signal integration in development,” *Science*, vol. 284, no. 5415, pp. 770–776, Apr. 1999.
- [146] S. J. Bray, “Notch signalling: a simple pathway becomes complex,” *Nat. Rev. Mol. Cell Biol.*, vol. 7, no. 9, pp. 678–689, Sep. 2006.
- [147] A. P. Weng *et al.*, “Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia,” *Science*, vol. 306, no. 5694, pp. 269–271, Oct. 2004.
- [148] T. Oyama *et al.*, “Mastermind-like 1 (MamL1) and mastermind-like 3 (MamL3) are essential for Notch signaling in vivo,” *Dev. Camb. Engl.*, vol. 138, no. 23, pp. 5235–5246, Dec. 2011.
- [149] F. Sahm *et al.*, “CIC and FUBP1 mutations in oligodendrogliomas, oligoastrocytomas and astrocytomas,” *Acta Neuropathol. (Berl.)*, vol. 123, no. 6, pp. 853–860, Jun. 2012.
- [150] F. V. Murphy and M. E. Churchill, “Nonsequence-specific DNA recognition: a structural perspective,” *Struct. Lond. Engl. 1993*, vol. 8, no. 4, pp. R83–89, Apr. 2000.
- [151] P. L. Privalov, A. I. Dragan, and C. Crane-Robinson, “The cost of DNA bending,” *Trends Biochem. Sci.*, vol. 34, no. 9, pp. 464–470, Sep. 2009.
- [152] M. Stros, “HMGB proteins: interactions with DNA and chromatin,” *Biochim. Biophys. Acta*, vol. 1799, no. 1–2, pp. 101–113, Feb. 2010.
- [153] H. M. Weir, P. J. Kraulis, C. S. Hill, A. R. Raine, E. D. Laue, and J. O. Thomas, “Structure of the HMG box motif in the B-domain of HMG1,” *EMBO J.*, vol. 12, no. 4, pp. 1311–1319, Apr. 1993.

- [154] M. E. A. Churchill, J. Klass, and D. L. Zoetewey, "Structural Analysis of HMGD-DNA Complexes Reveal Influence of Intercalation on Sequence Selectivity and DNA Bending," *J. Mol. Biol.*, vol. 403, no. 1, pp. 88–102, Oct. 2010.
- [155] K. Specht, Y.-S. Sung, L. Zhang, G. H. S. Richter, C. D. Fletcher, and C. R. Antonescu, "Distinct transcriptional signature and immunoprofile of *CIC-DUX4* fusion-positive round cell tumors compared to *EWSR1* -rearranged ewing sarcomas: Further evidence toward distinct pathologic entities: Gene Expression in *CIC-DUX4* Sarcomas," *Genes. Chromosomes Cancer*, vol. 53, no. 7, pp. 622–633, Jul. 2014.
- [156] K. I. Patterson, T. Brummer, P. M. O'Brien, and R. J. Daly, "Dual-specificity phosphatases: critical regulators with diverse cellular targets," *Biochem. J.*, vol. 418, no. 3, pp. 475–489, Mar. 2009.
- [157] J. M. Mason, D. J. Morrison, M. A. Basson, and J. D. Licht, "Sprouty proteins: multifaceted negative-feedback regulators of receptor tyrosine kinase signaling," *Trends Cell Biol.*, vol. 16, no. 1, pp. 45–54, Jan. 2006.
- [158] I. Appolloni *et al.*, "PDGF-B induces a homogeneous class of oligodendrogliomas from embryonic neural progenitors," *Int. J. Cancer*, vol. 124, no. 10, pp. 2251–2259, May 2009.
- [159] C. Dai, J. C. Celestino, Y. Okada, D. N. Louis, G. N. Fuller, and E. C. Holland, "PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes in vivo," *Genes Dev.*, vol. 15, no. 15, pp. 1913–1925, Aug. 2001.
- [160] W. A. Weiss *et al.*, "Genetic determinants of malignancy in a mouse model for oligodendroglioma," *Cancer Res.*, vol. 63, no. 7, pp. 1589–1595, Apr. 2003.
- [161] T. Oikawa and T. Yamada, "Molecular biology of the Ets family of transcription factors," *Gene*, vol. 303, pp. 11–34, Jan. 2003.

- [162] A. D. Sharrocks, “The ETS-domain transcription factor family,” *Nat. Rev. Mol. Cell Biol.*, vol. 2, no. 11, pp. 827–837, Nov. 2001.
- [163] S. Oh, S. Shin, and R. Janknecht, “ETV1, 4 and 5: an oncogenic subfamily of ETS transcription factors,” *Biochim. Biophys. Acta*, vol. 1826, no. 1, pp. 1–12, Aug. 2012.
- [164] R. Janknecht, “EWS-ETS oncoproteins: the linchpins of Ewing tumors,” *Gene*, vol. 363, pp. 1–14, Dec. 2005.
- [165] I. S. Jeon *et al.*, “A variant Ewing’s sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1,” *Oncogene*, vol. 10, no. 6, pp. 1229–1234, Mar. 1995.
- [166] B. E. Helgeson *et al.*, “Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer,” *Cancer Res.*, vol. 68, no. 1, pp. 73–80, Jan. 2008.
- [167] S. A. Tomlins *et al.*, “Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer,” *Science*, vol. 310, no. 5748, pp. 644–648, Oct. 2005.
- [168] S. A. Tomlins *et al.*, “TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer,” *Cancer Res.*, vol. 66, no. 7, pp. 3396–3400, Apr. 2006.
- [169] V. G. LeBlanc *et al.*, “Comparative transcriptome analysis of isogenic cell line models and primary cancers links capicua (CIC) loss to activation of the MAPK signalling cascade,” *J. Pathol.*, vol. 242, no. 2, pp. 206–220, Jun. 2017.
- [170] I. Tirosh *et al.*, “Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma,” *Nature*, vol. 539, no. 7628, pp. 309–313, 10 2016.
- [171] A. S. Venteicher *et al.*, “Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq,” *Science*, vol. 355, no. 6332, p. eaai8478, Mar. 2017.
- [172] B. Wiestler *et al.*, “Integrated DNA methylation and copy-number profiling identify three clinically and biologically relevant groups of anaplastic glioma,” *Acta Neuropathol. (Berl.)*, vol. 128, no. 4, pp. 561–571, Oct. 2014.

- [173] H. Suzuki *et al.*, “Mutational landscape and clonal architecture in grade II and III gliomas,” *Nat. Genet.*, vol. 47, no. 5, pp. 458–468, May 2015.
- [174] H. H. Engelhard, A. Stelea, and A. Mundt, “Oligodendroglioma and anaplastic oligodendroglioma: clinical features, treatment, and prognosis,” *Surg. Neurol.*, vol. 60, no. 5, pp. 443–456, Nov. 2003.
- [175] K. N. Bhalla, “Epigenetic and chromatin modifiers as targeted therapy of hematologic malignancies,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 23, no. 17, pp. 3971–3993, Jun. 2005.
- [176] A. K.-Y. Chan *et al.*, “Loss of CIC and FUBP1 expressions are potential markers of shorter time to recurrence in oligodendroglial tumors,” *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc.*, vol. 27, no. 3, pp. 332–342, Mar. 2014.
- [177] B. Denard, C. Lee, and J. Ye, “Doxorubicin blocks proliferation of cancer cells through proteolytic activation of CREB3L1,” *eLife*, vol. 1, p. e00090, Dec. 2012.
- [178] D. N. Louis *et al.*, “The 2007 WHO classification of tumours of the central nervous system,” *Acta Neuropathol. (Berl.)*, vol. 114, no. 2, pp. 97–109, Aug. 2007.

APPENDIX

APPENDIX – I: PCR Primer List

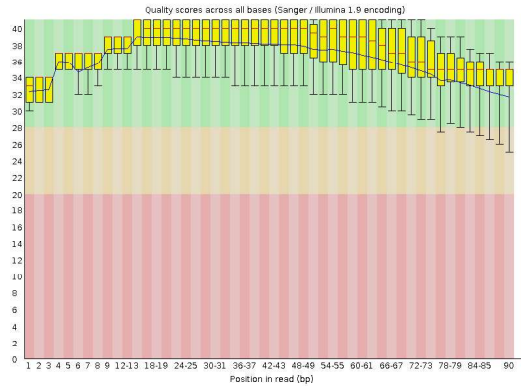
Primers sequences used for PCR amplification

Name	Forward	Reverse
IDH1 Exon 4	5'-GAGAAAGAGGGTTGAGGAGTTC-3'	5'-TGTGTTGAGATGGACGCCCTATTTG-3'
IDH2 Exon 4	5'-GAGGATGGCTAGGCGAGGAG-3'	5'-CGTCTGGCTGTGTTGTTGCTTG-3'
CIC Exon 5	5'-CCCTCACTGTCCCTGCTGCC-3'	5'-CGTCTCCTTGTGCCCCCCTCTG-3'
Exome Library Quantification	5'-AATGATACGGCGACCACCGA-3'	5'-CAAGCAGAAGACGGCATACGA-3'
Microsatellite-D1S468	5'-AATTAAACCGTTTGGTCCT-3'	5'-GCGACACACACTTCCCC-3'
Microsatellite-D1S548	5'-GAACTCATTTGGCAAAGGAA-3'	5'-GCCCTCTTTGTTGCAGTGATT-3'
Microsatellite-D1S214	5'-CCGAATGACAAAGTGAGACT-3'	5'-AATGTTGTTTCCAAAAGTGGC-3'
Microsatellite-D1S1597	5'-TTTATTGAGATATATTGACATGCA-3'	5'-AAGGAGGAAAAGCTTTTGGGA-3'
Microsatellite-D1S552	5'-TTCATGCAGCATCATCCCC-3'	5'-TGTGGGCAGGTGTAAAAGAGT-3'
Microsatellite-D1S178	5'-CACAAACACTGTTTCATTGTGC-3'	5'-TTTCAGTAGAATTCAGGCC-3'
Microsatellite-D1S180	5'-CTAAATATCCATCAAGGAATG-3'	5'-TACTCCATTTTCATTCAGGT-3'

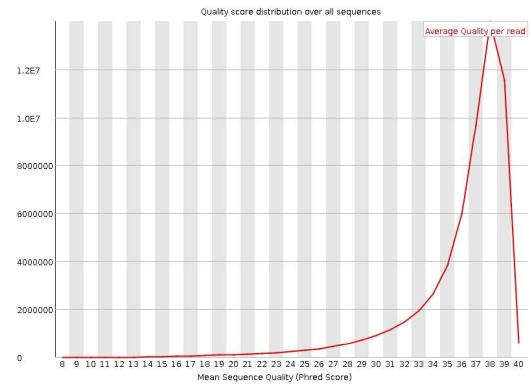
Appendix – II: Exome Sequencing Data Quality

ODG01-BLOOD

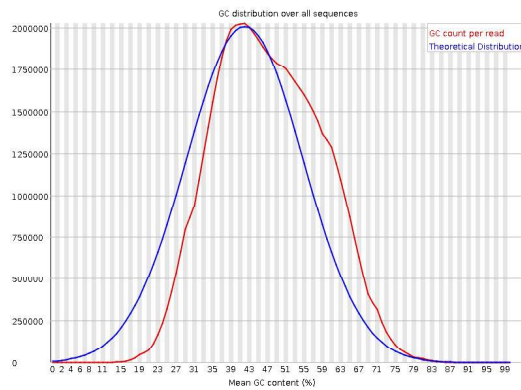
Quality Scores Across All Bases



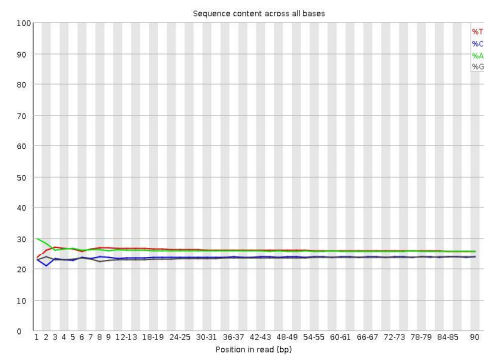
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

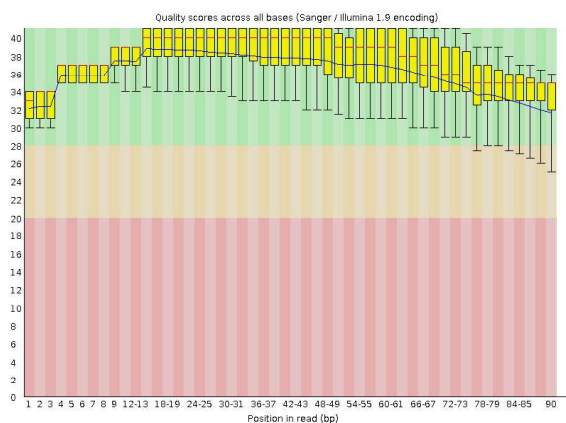


Sequence Content Across All Bases

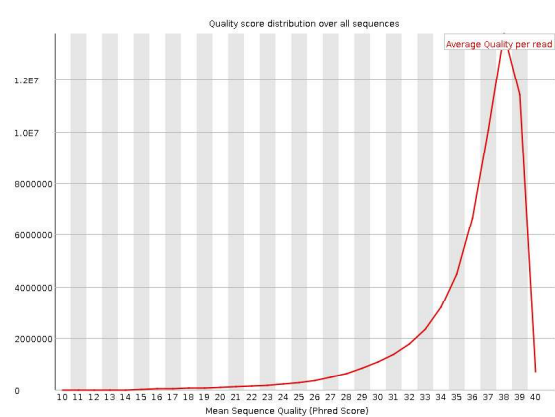


ODG01-TUMOUR

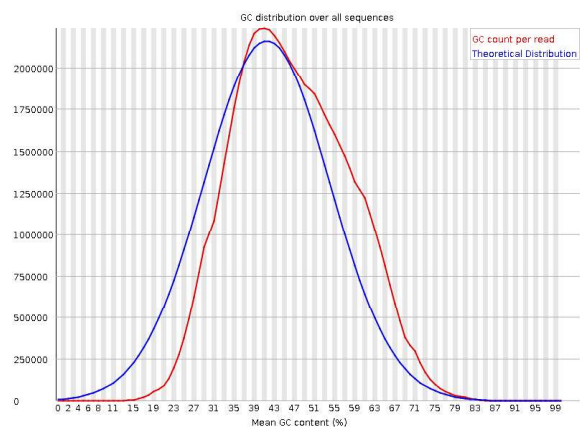
Quality Scores Across All Bases



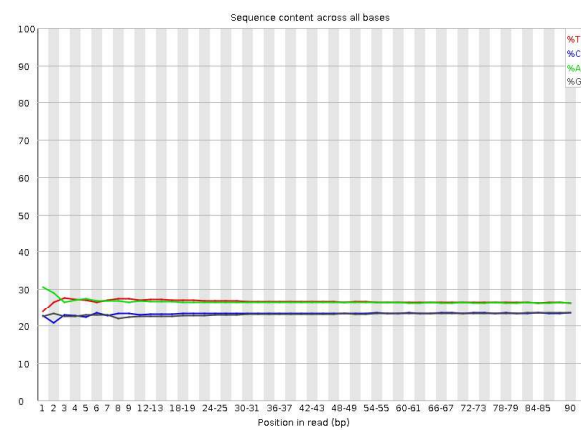
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

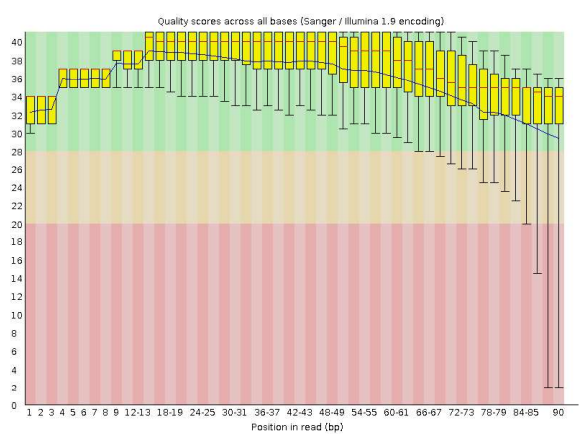


Sequence Content Across All Bases

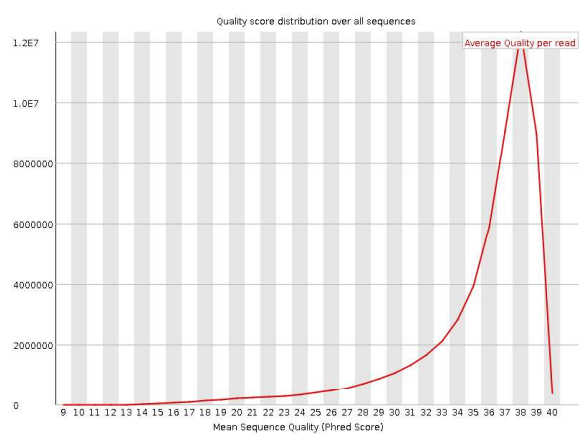


ODG2-BLOOD

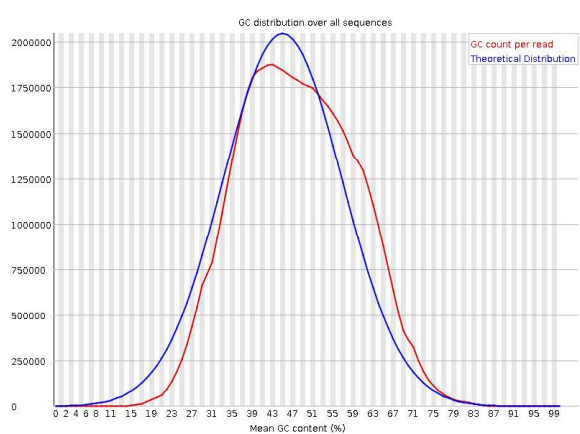
Quality Scores Across All Bases



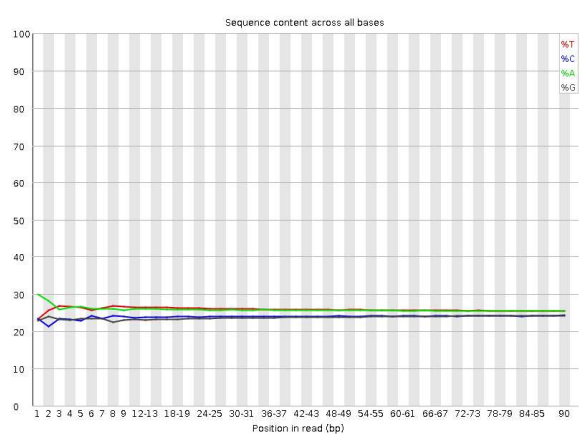
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

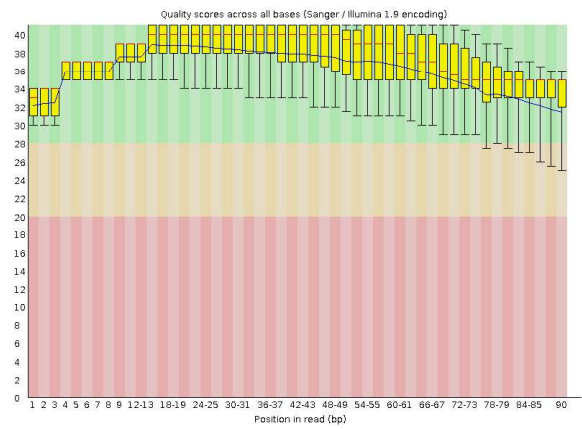


Sequence Content Across All Bases

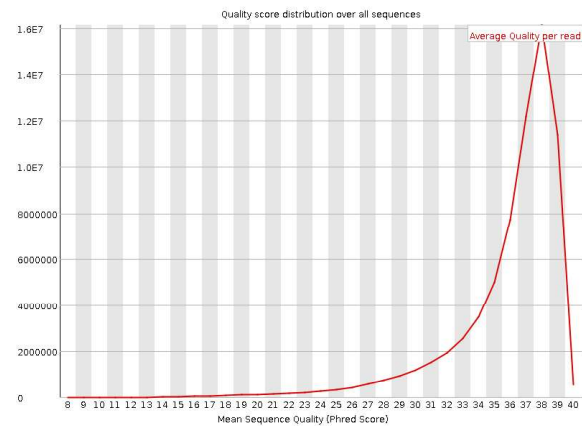


ODG2-TUMOUR

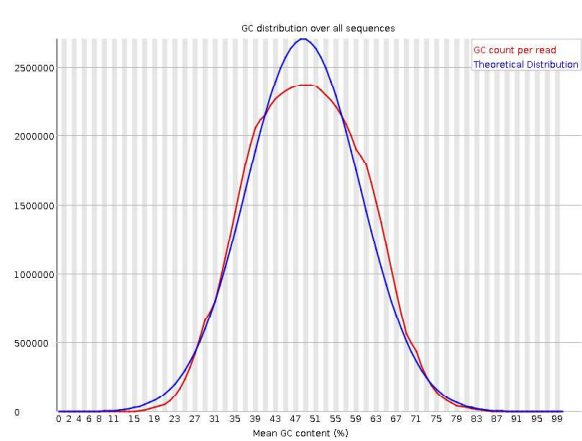
Quality Scores Across All Bases



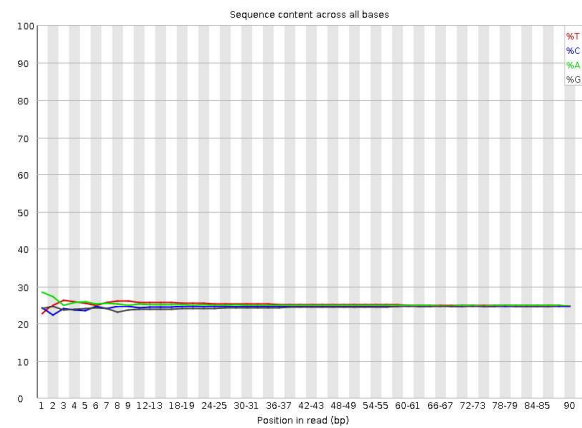
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

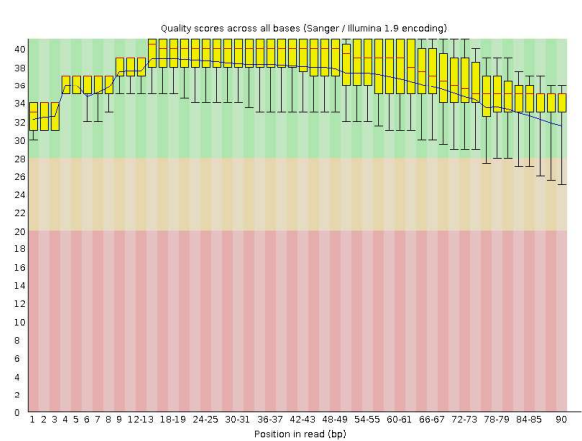


Sequence Content Across All Bases

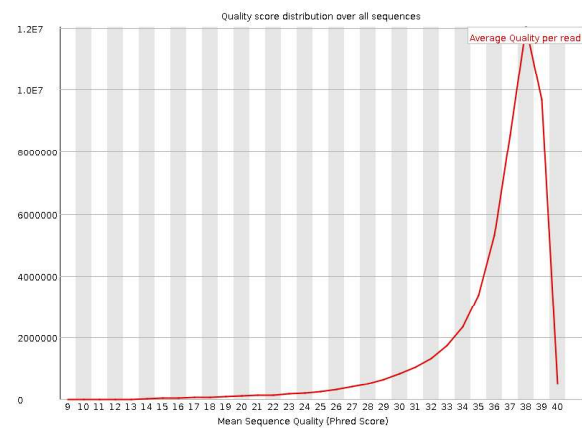


ODG3-BLOOD

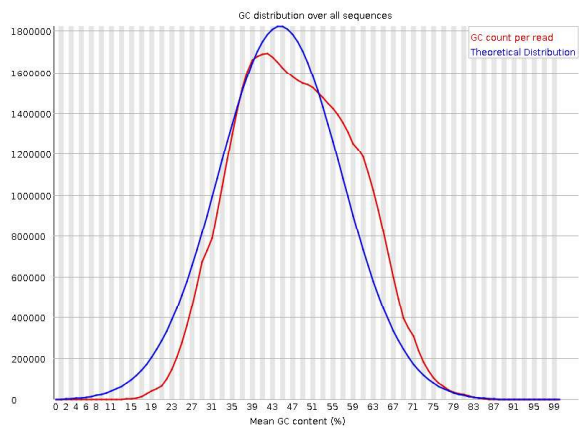
Quality Scores Across All Bases



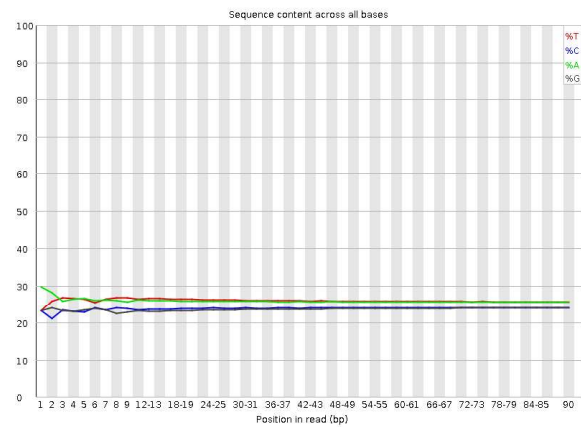
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

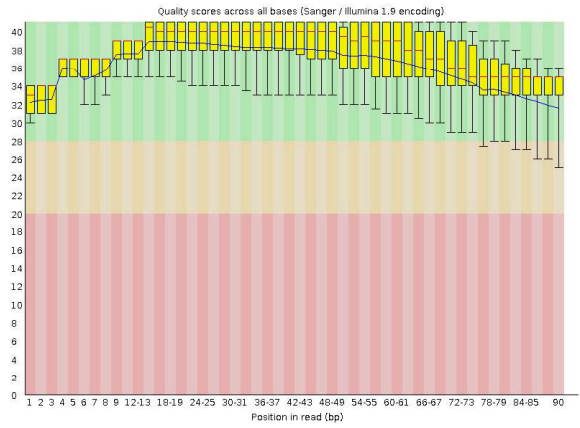


Sequence Content Across All Bases

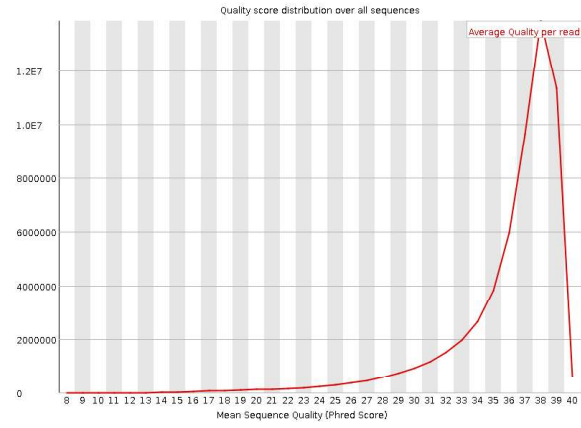


ODG3-TUMOUR

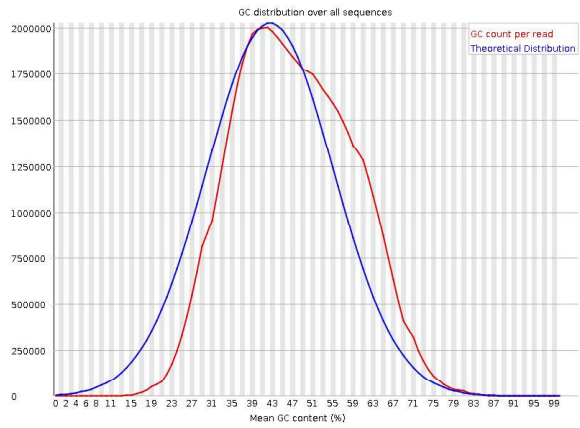
Quality Scores Across All Bases



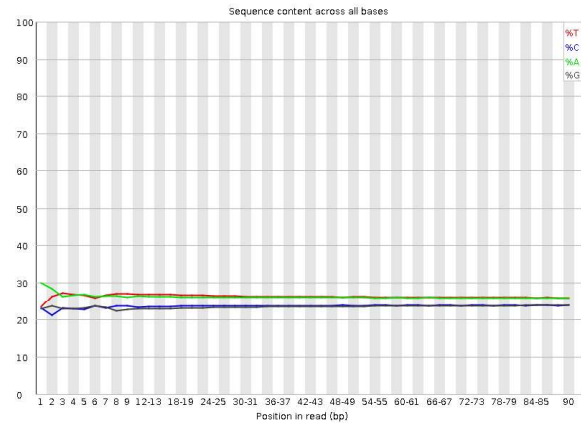
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

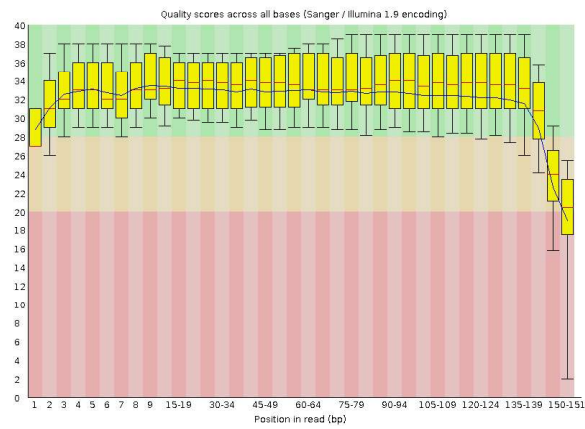


Sequence Content Across All Bases

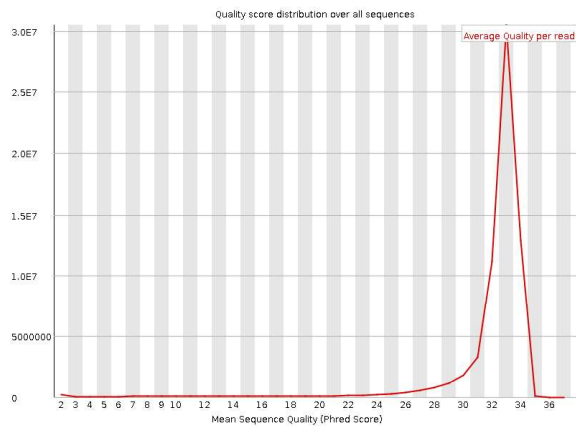


ODG4-BLOOD

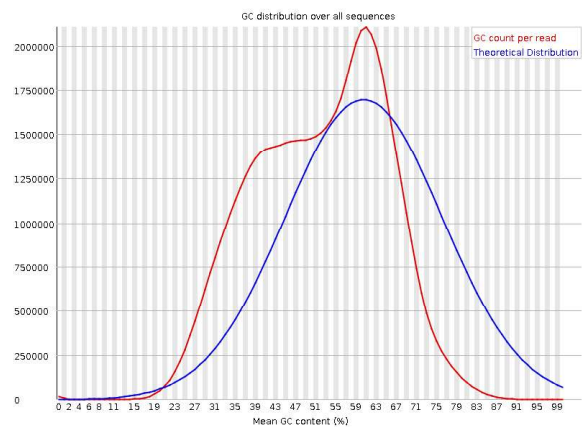
Quality Scores Across All Bases



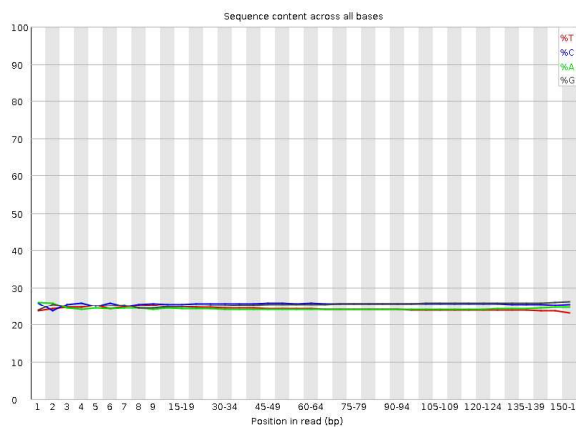
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

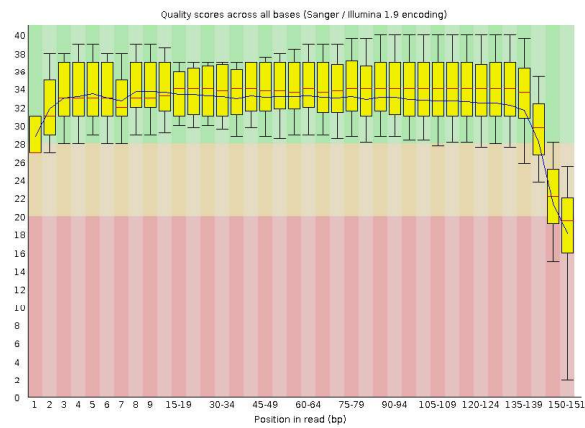


Sequence Content Across All Bases

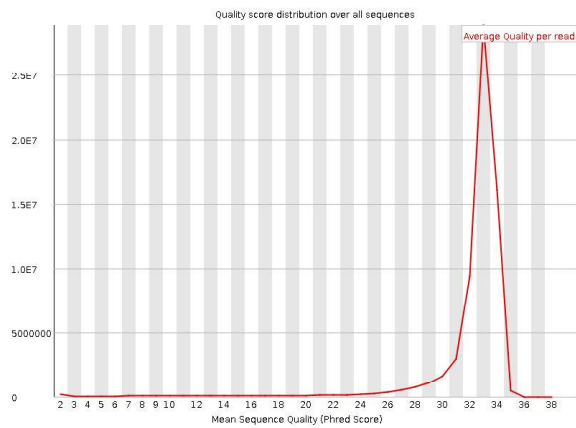


ODG4-TUMOUR

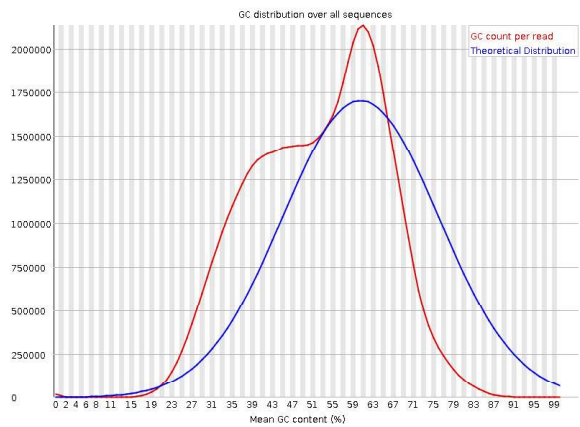
Quality Scores Across All Bases



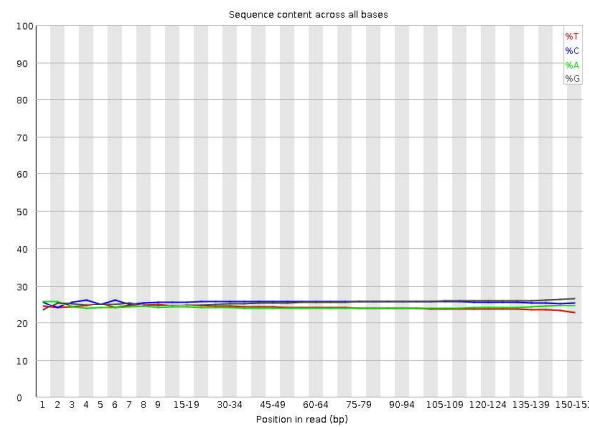
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

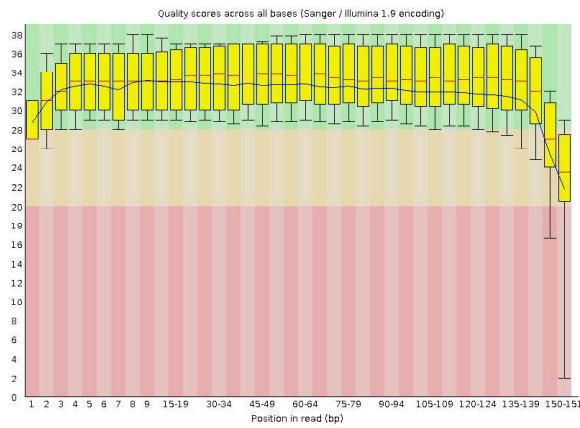


Sequence Content Across All Bases

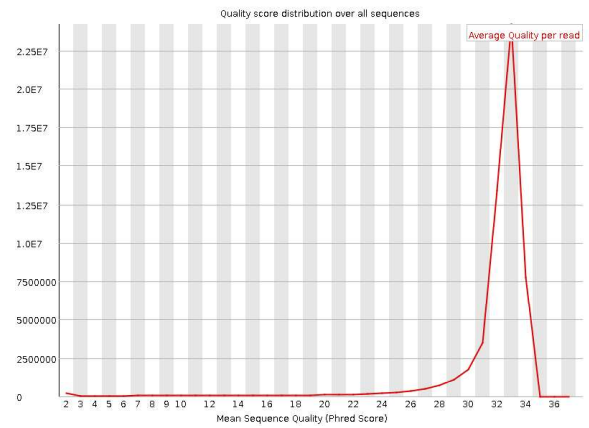


ODG5-BLOOD

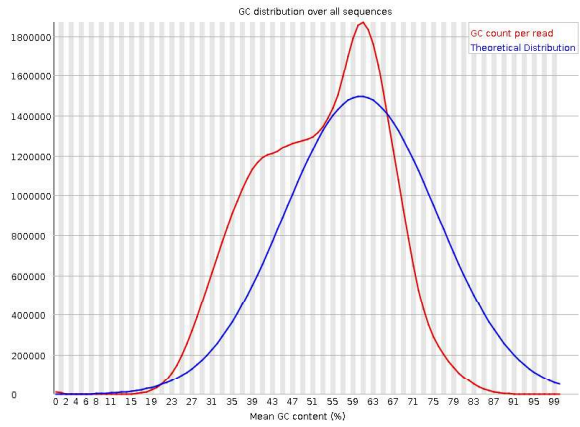
Quality Scores Across All Bases



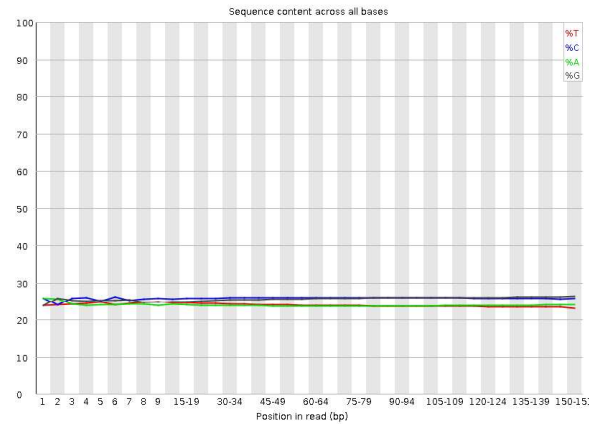
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

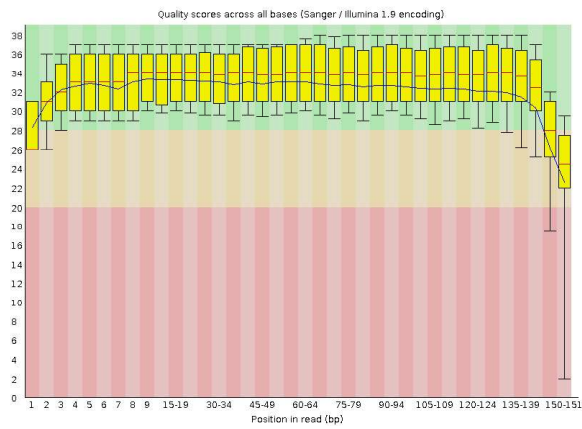


Sequence Content Across All Bases

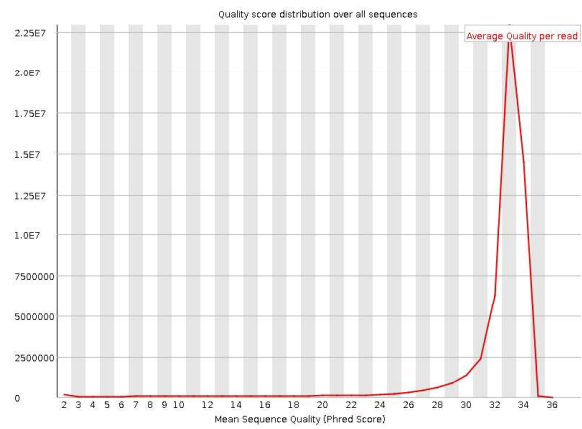


ODG5-TUMOUR

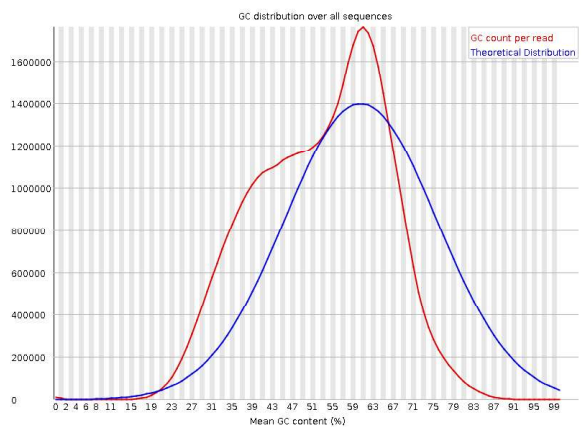
Quality Scores Across All Bases



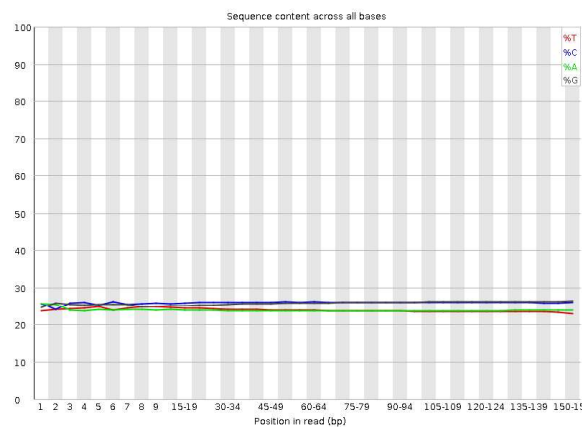
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

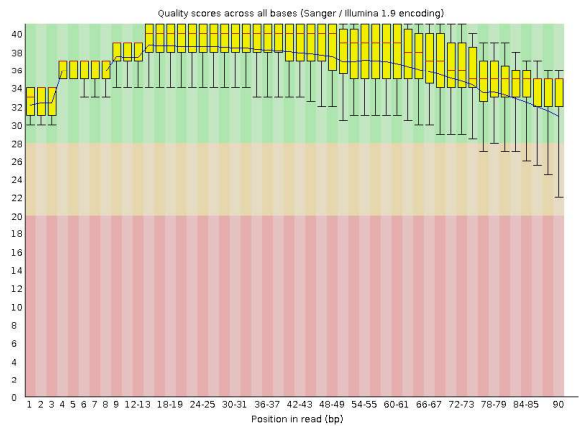


Sequence Content Across All Bases

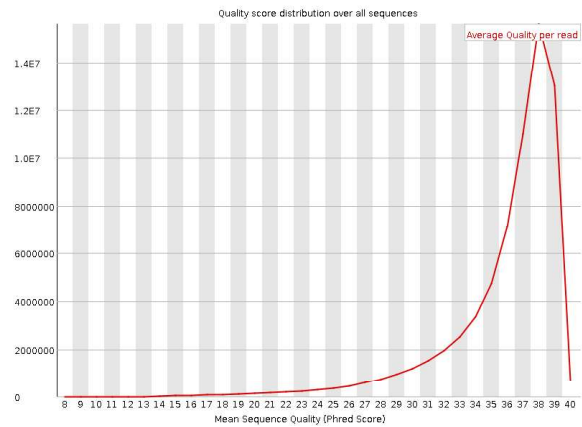


ODG6-BLOOD

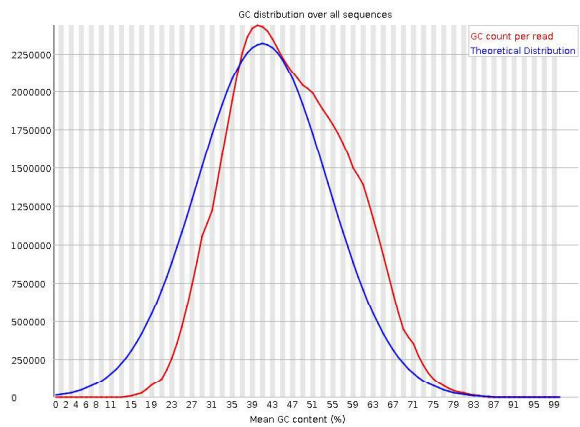
Quality Scores Across All Bases



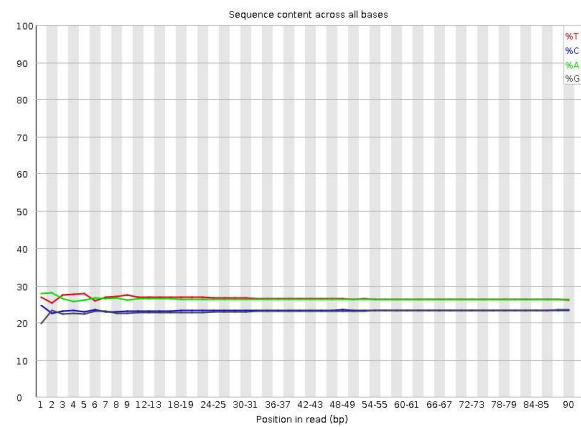
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

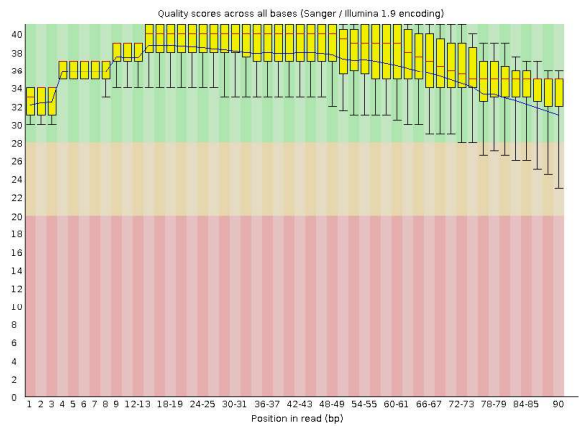


Sequence Content Across All Bases

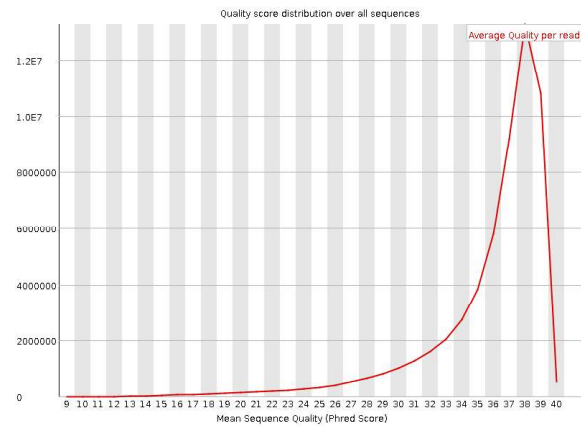


ODG6-TUMOUR

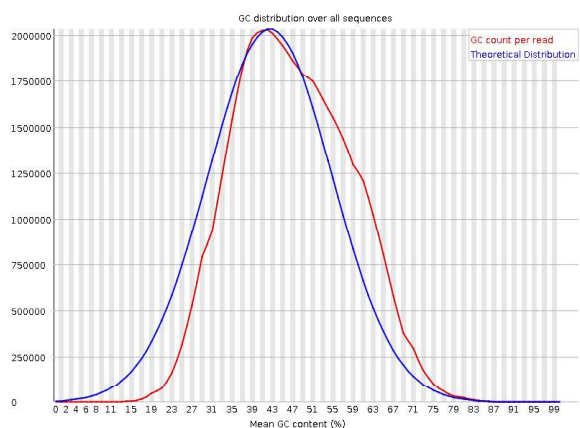
Quality Scores Across All Bases



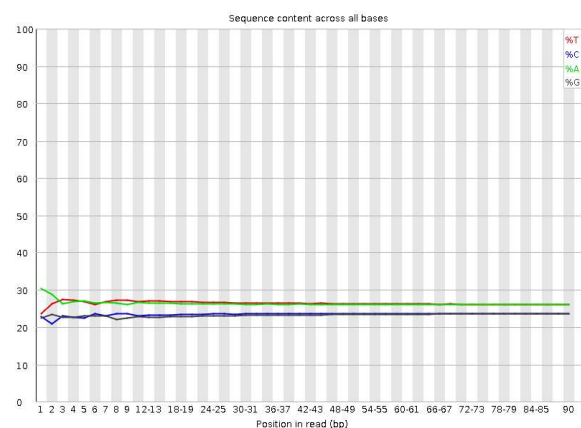
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

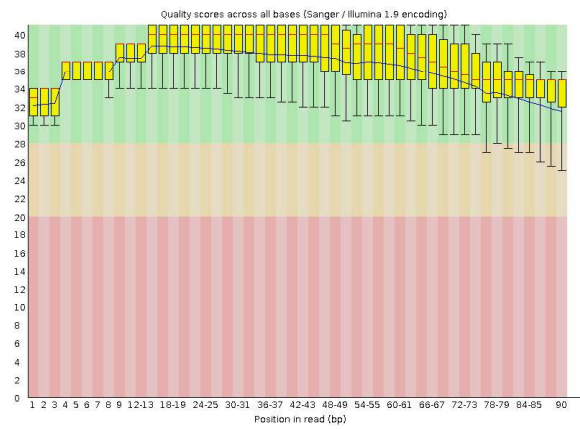


Sequence Content Across All Bases

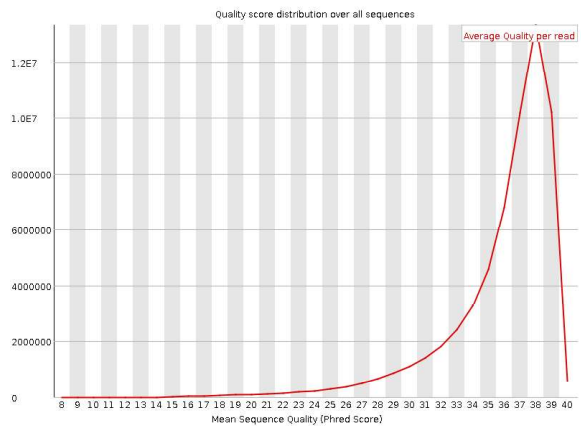


ODG7-BLOOD

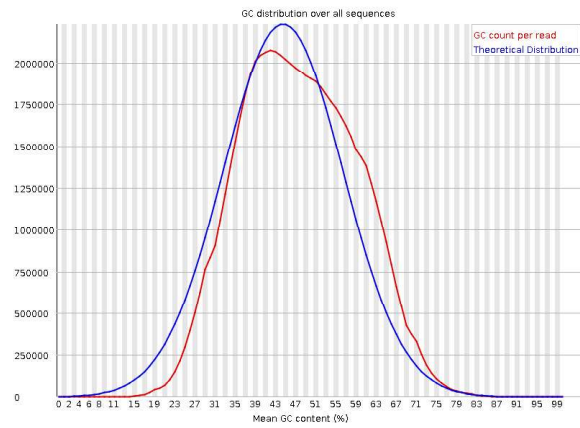
Quality Scores Across All Bases



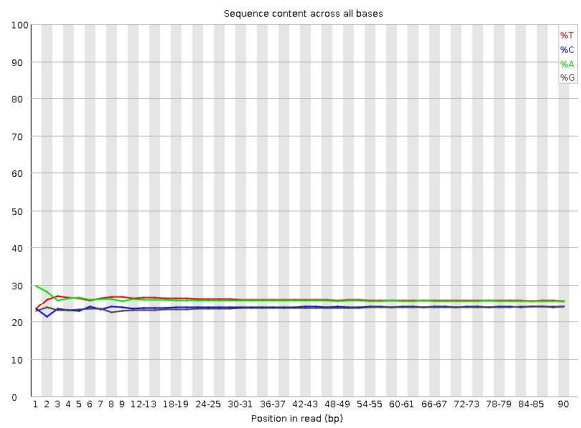
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

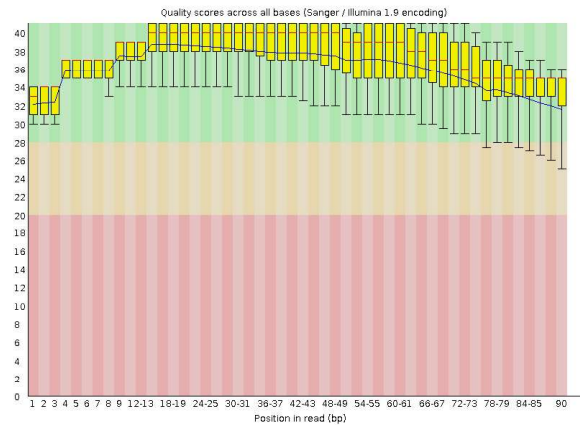


Sequence Content Across All Bases

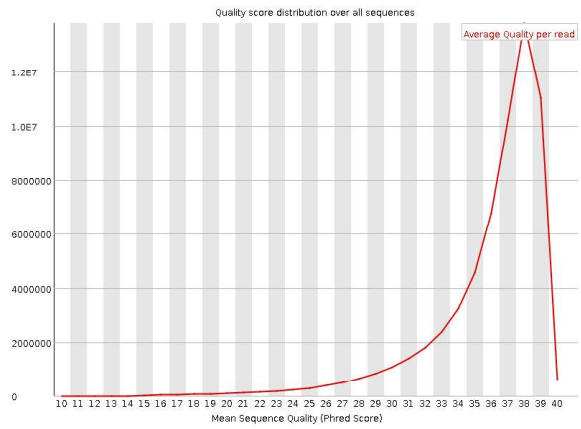


ODG7-TUMOUR

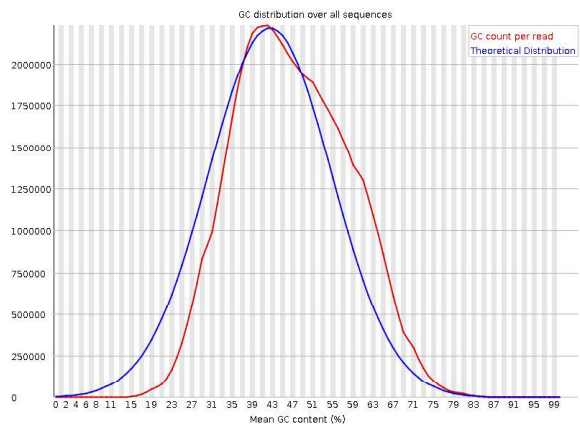
Quality Scores Across All Bases



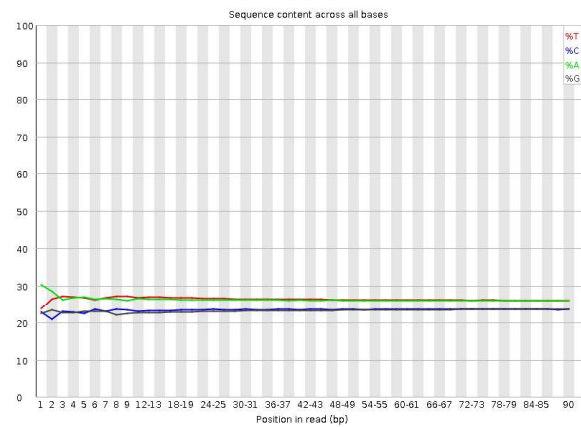
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

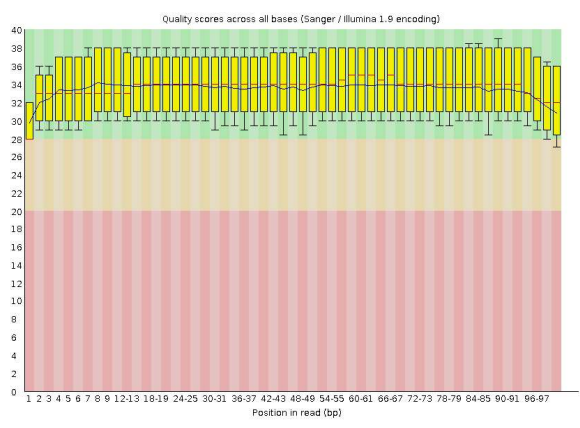


Sequence Content Across All Bases

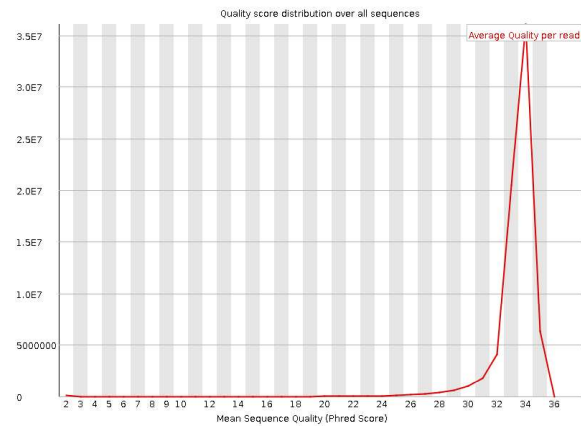


ODG8-BLOOD

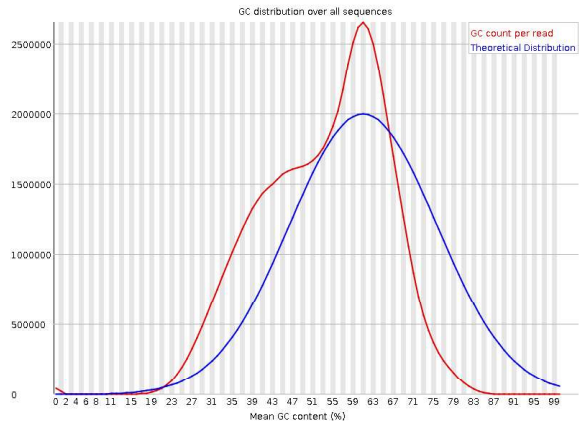
Quality Scores Across All Bases



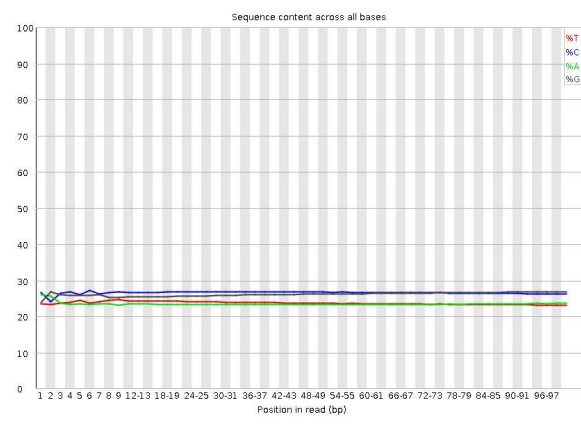
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

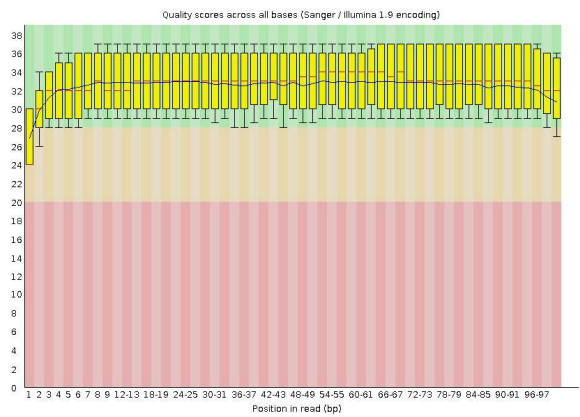


Sequence Content Across All Bases

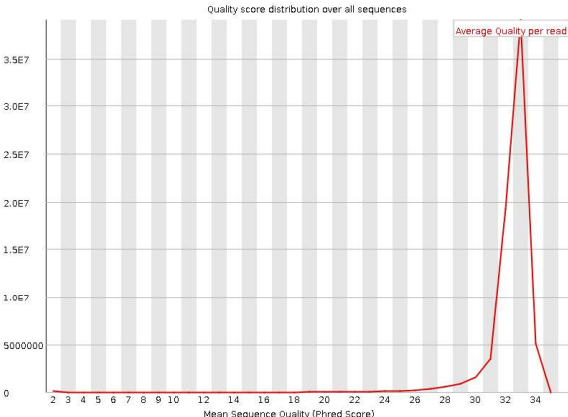


ODG8-TUMOUR

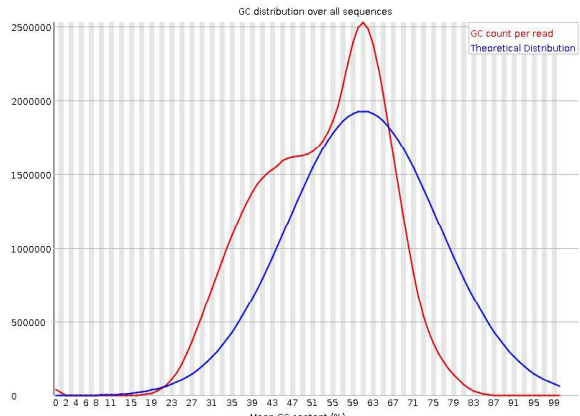
Quality Scores Across All Bases



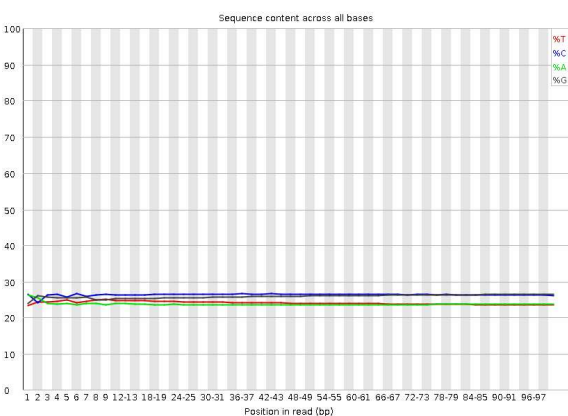
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

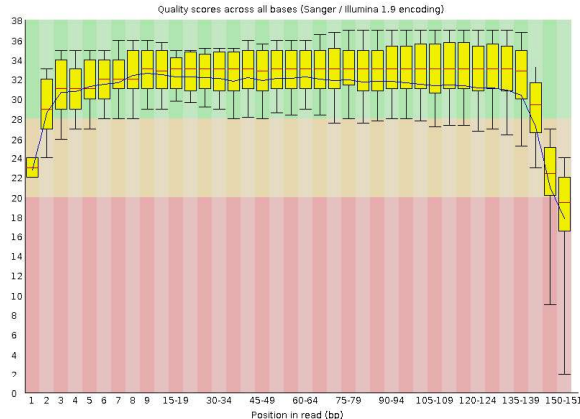


Sequence Content Across All Bases

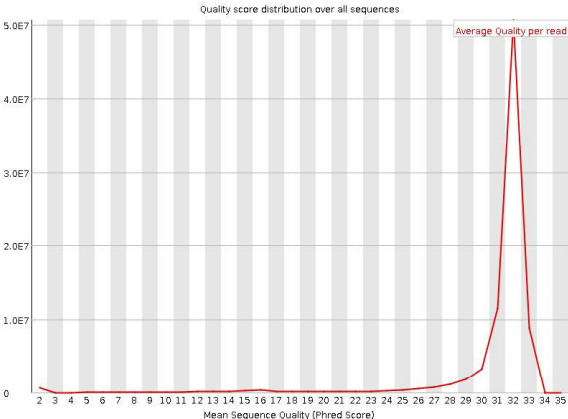


ODG9-BLOOD

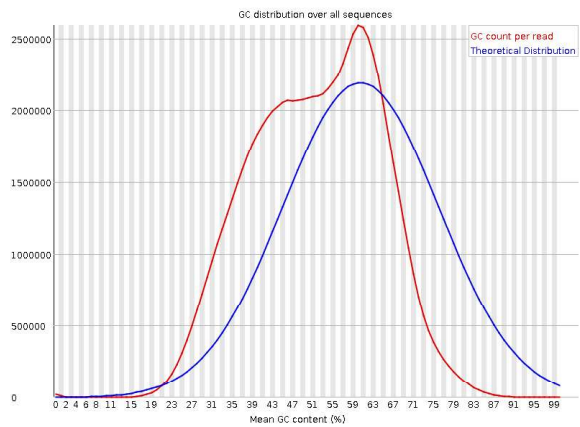
Quality Scores Across All Bases



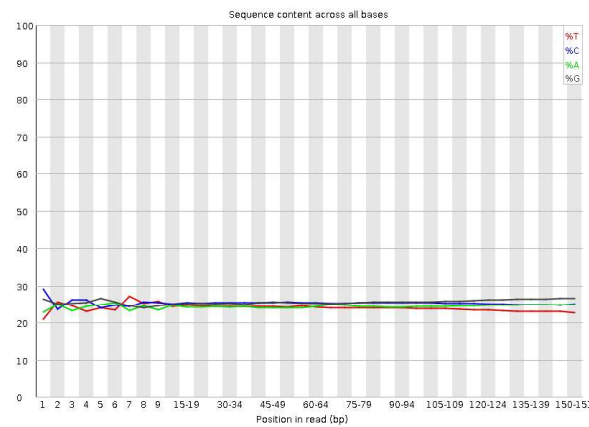
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

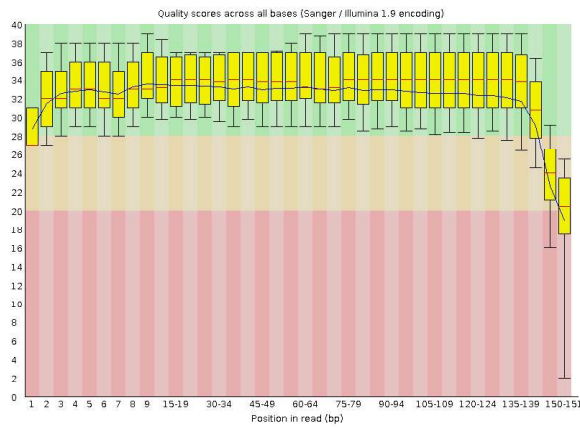


Sequence Content Across All Bases

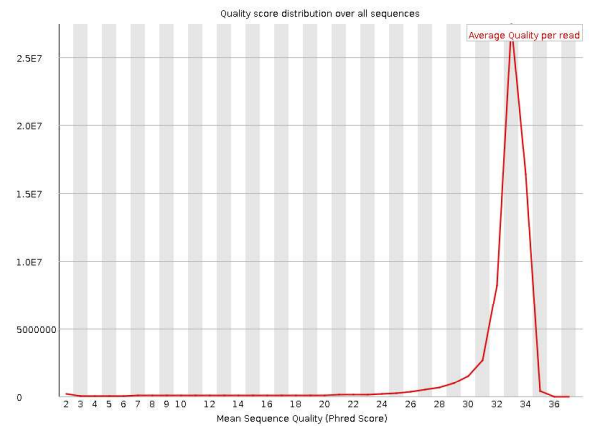


ODG9-TUMOUR

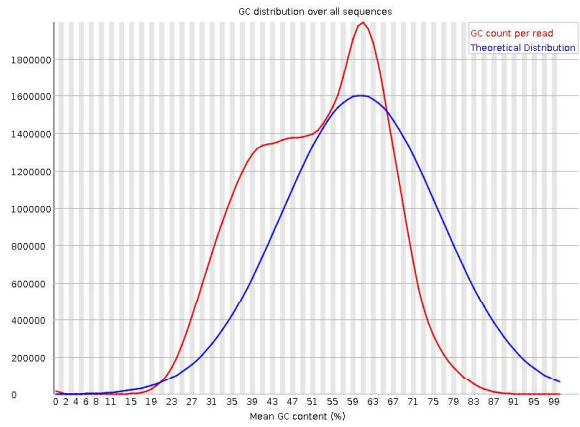
Quality Scores Across All Bases



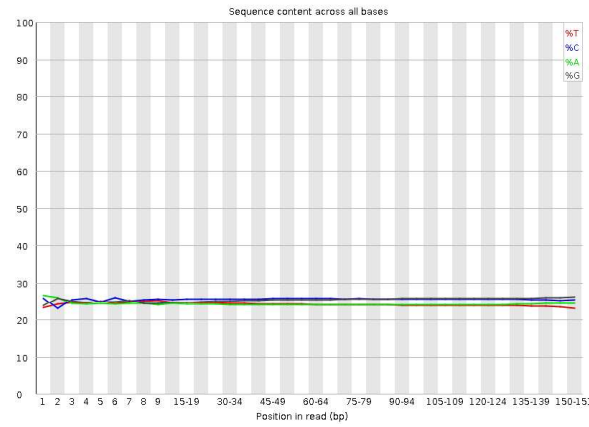
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

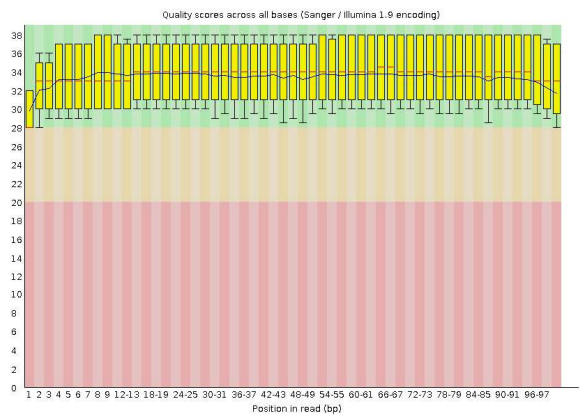


Sequence Content Across All Bases

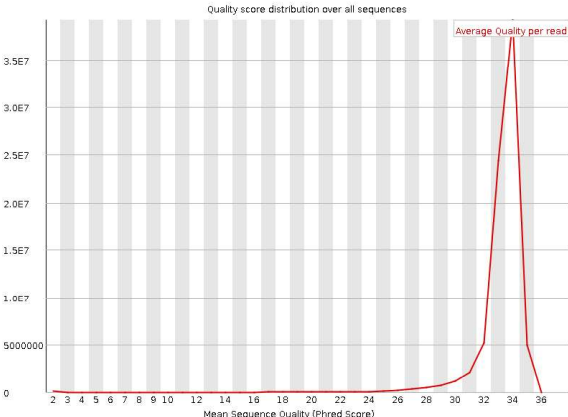


ODG10-BLOOD

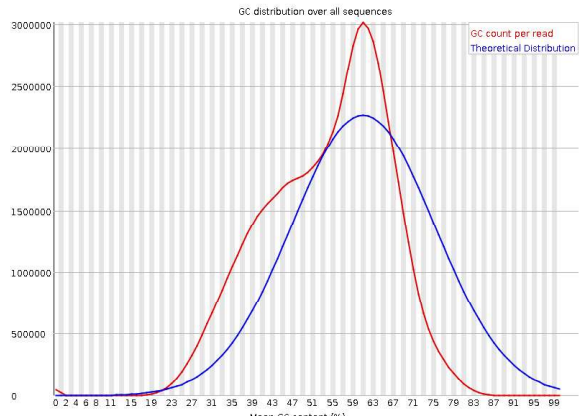
Quality Scores Across All Bases



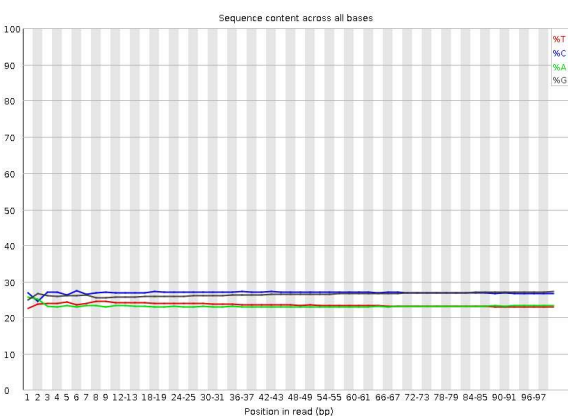
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

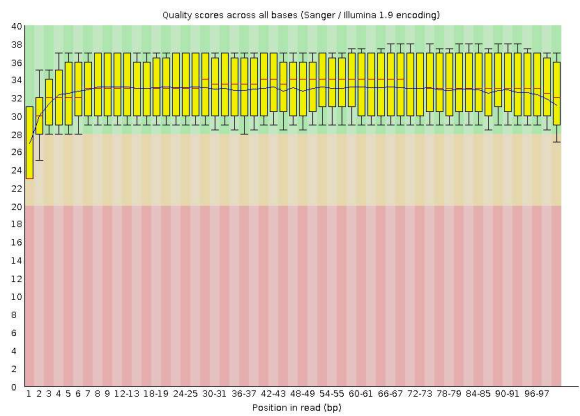


Sequence Content Across All Bases

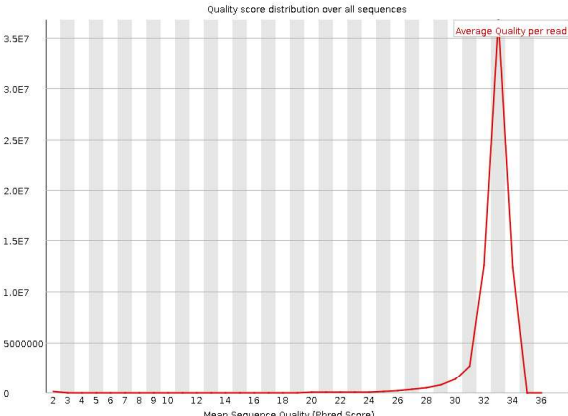


ODG10-TUMOUR

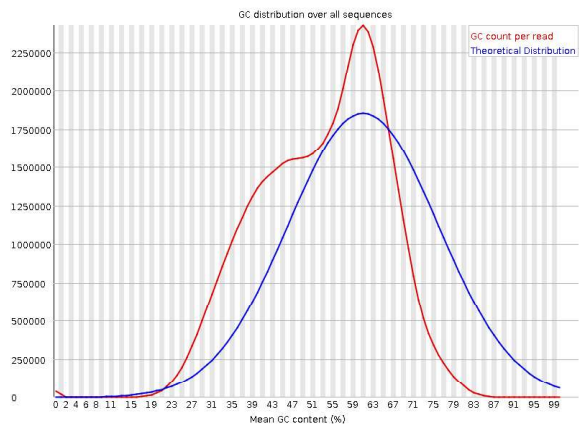
Quality Scores Across All Bases



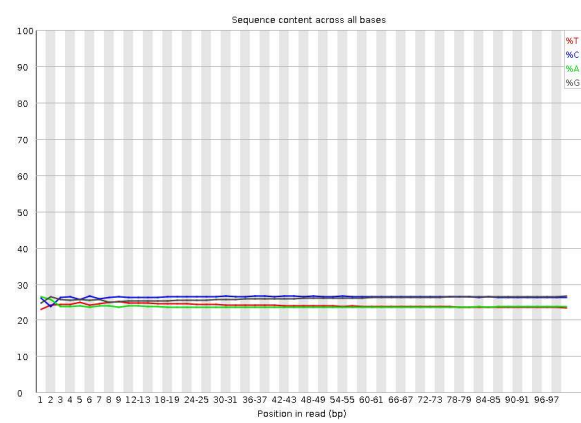
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

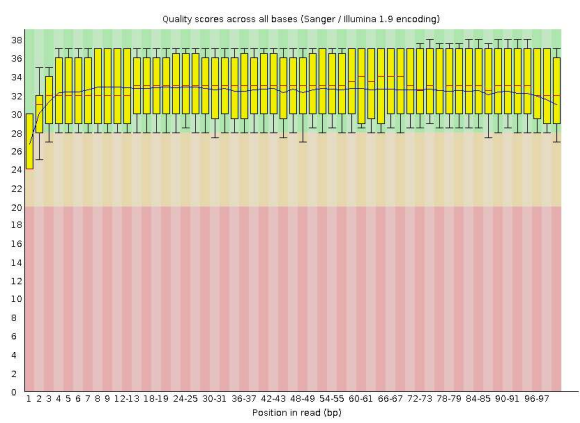


Sequence Content Across All Bases

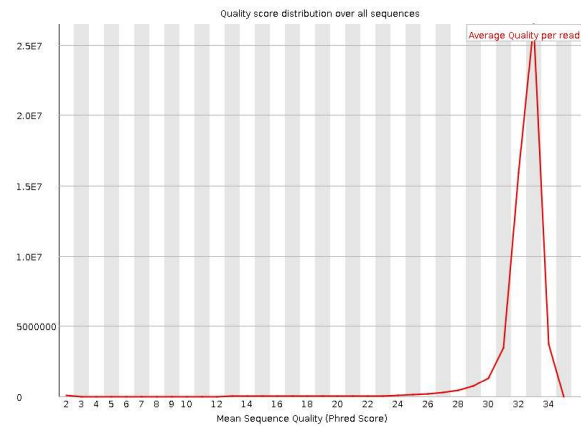


ODG11-BLOOD

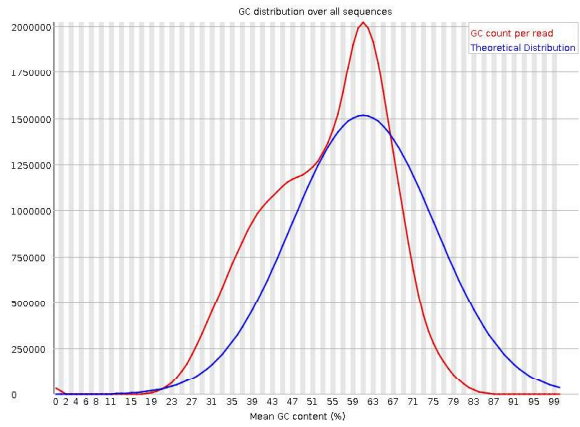
Quality Scores Across All Bases



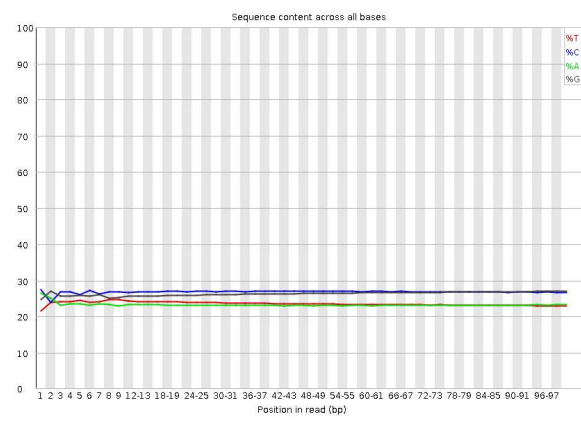
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

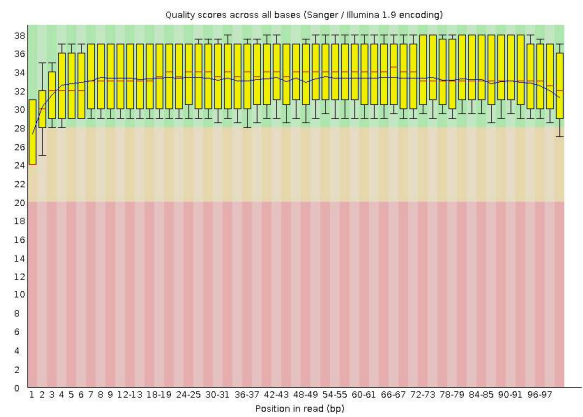


Sequence Content Across All Bases

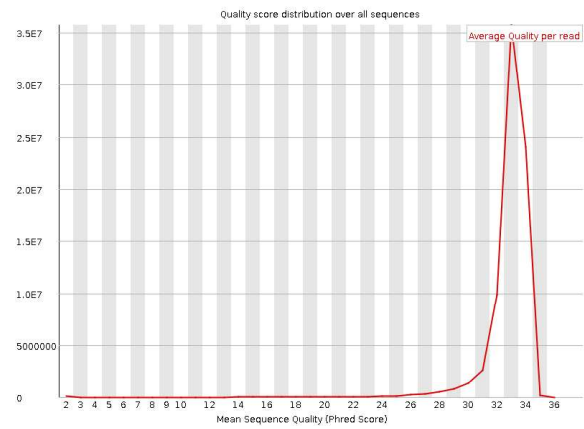


ODG11-TUMOUR

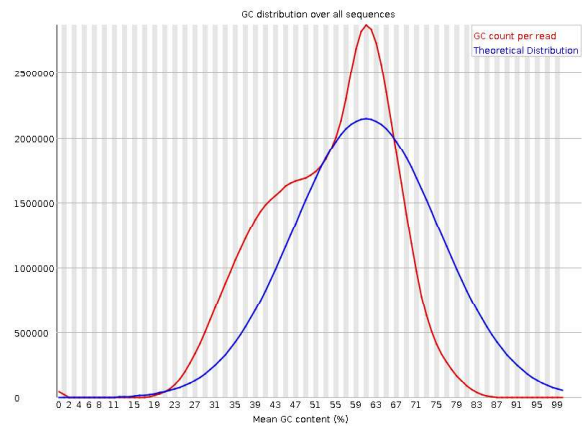
Quality Scores Across All Bases



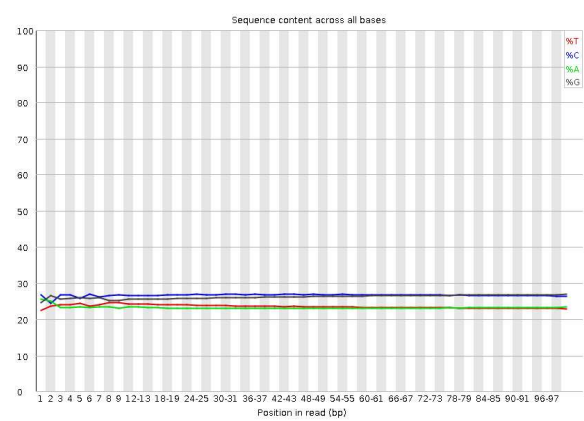
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences



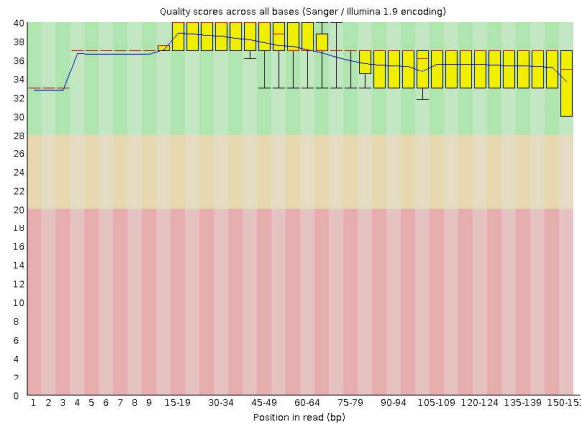
Sequence Content Across All Bases



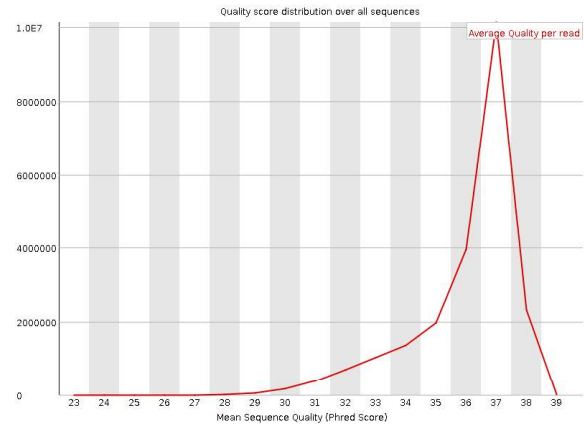
Appendix - III: RNAseq Data Quality

ODG1

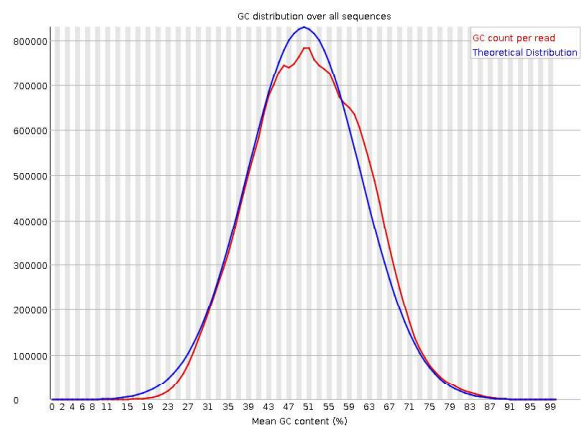
Quality Scores Across All Bases



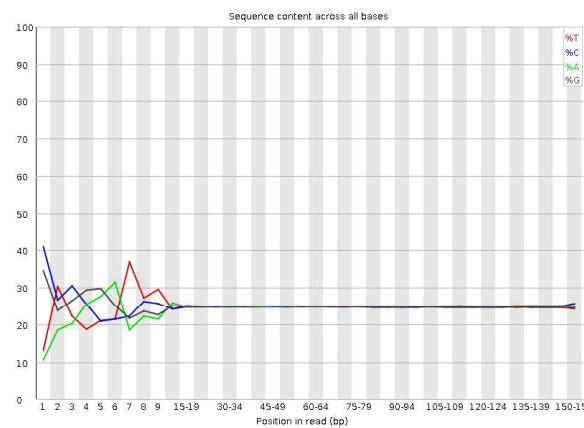
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

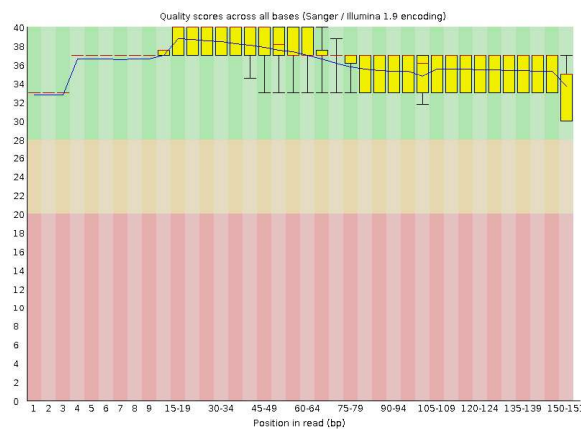


Sequence Content Across All Bases

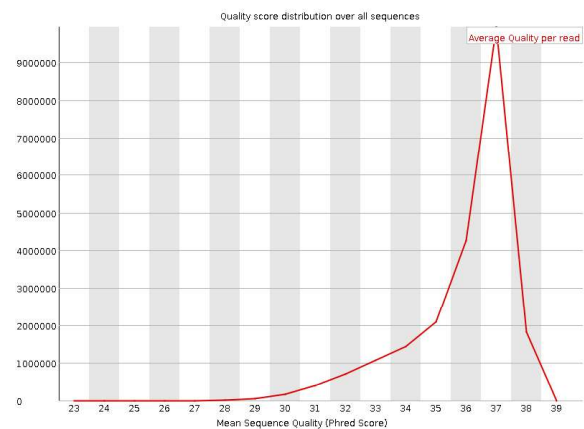


ODG2

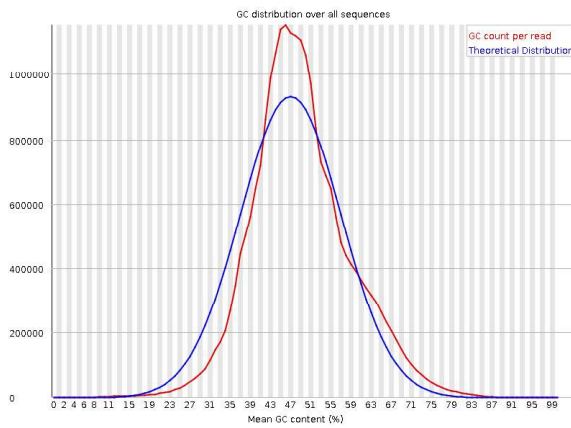
Quality Scores Across All Bases



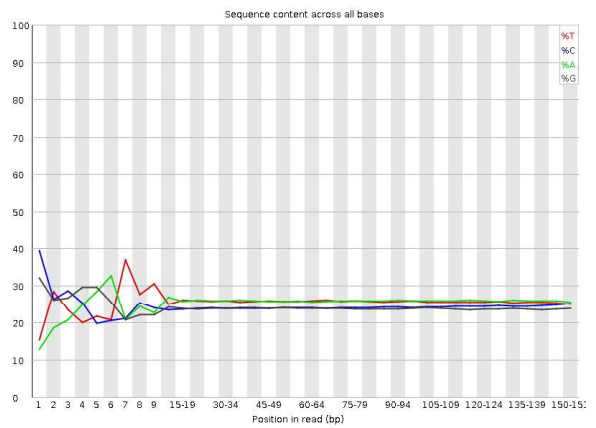
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

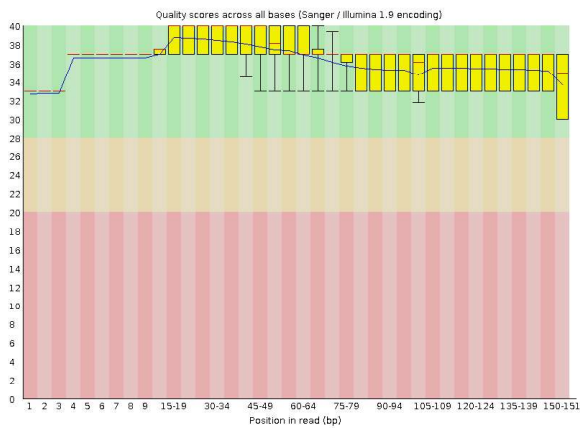


Sequence Content Across All Bases

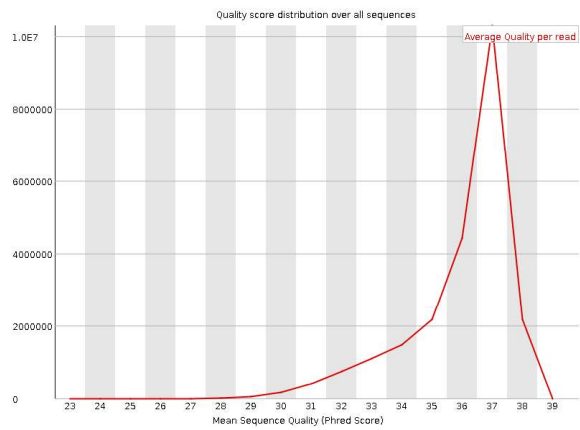


ODG3

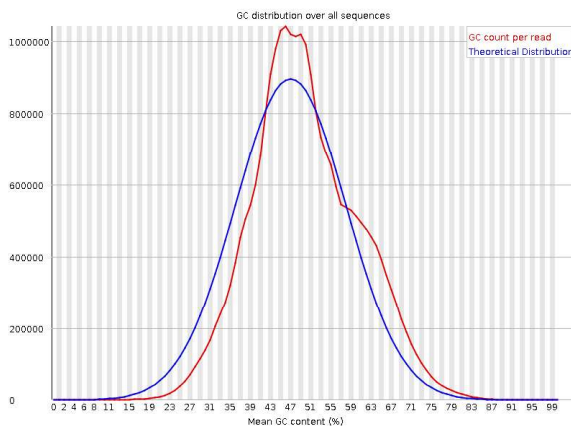
Quality Scores Across All Bases



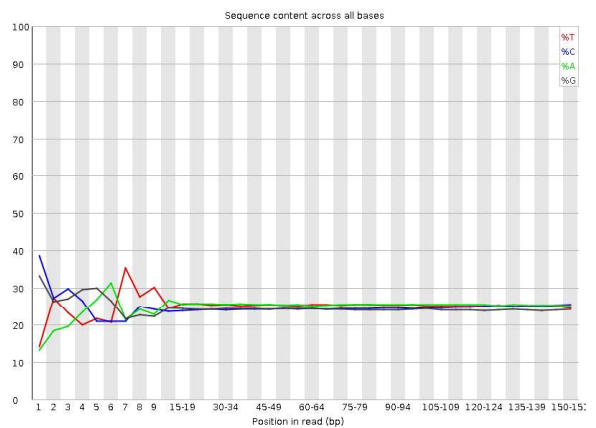
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

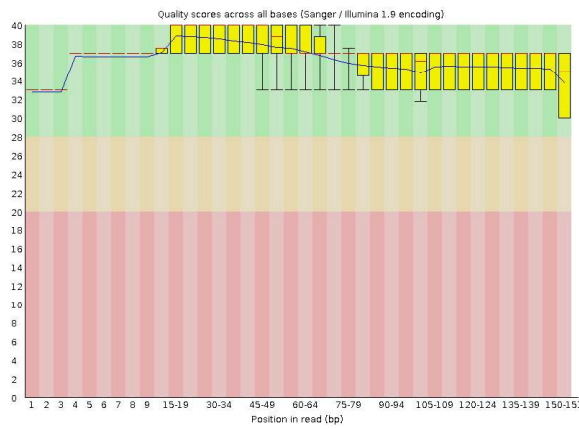


Sequence Content Across All Bases

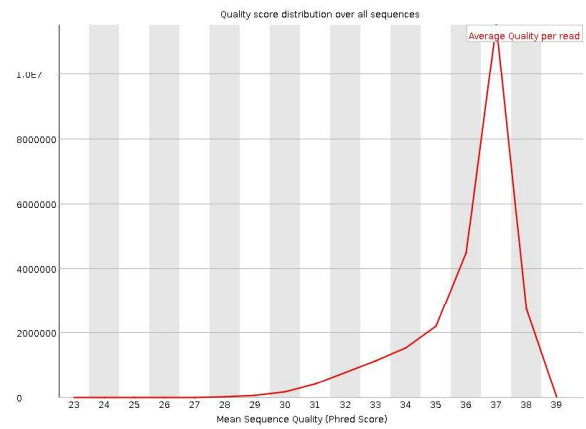


ODG4

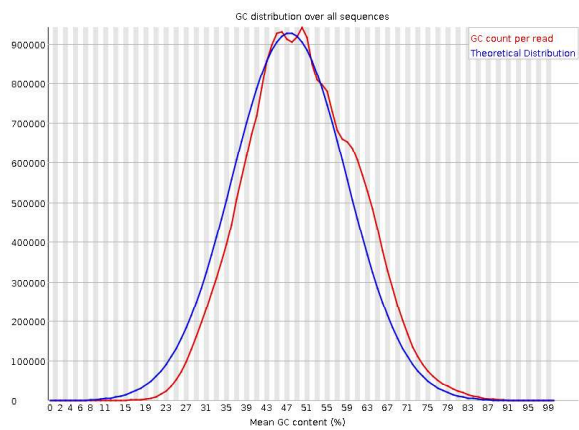
Quality Scores Across All Bases



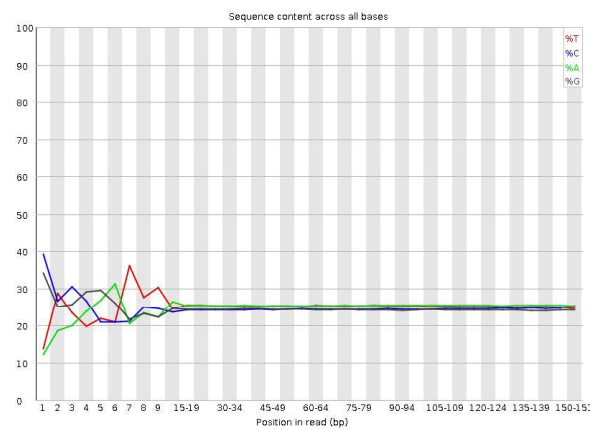
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

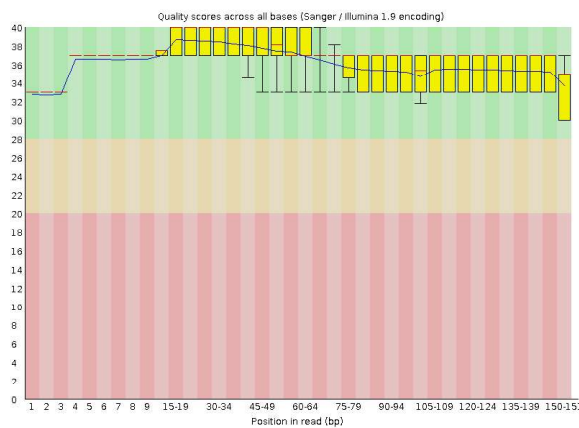


Sequence Content Across All Bases

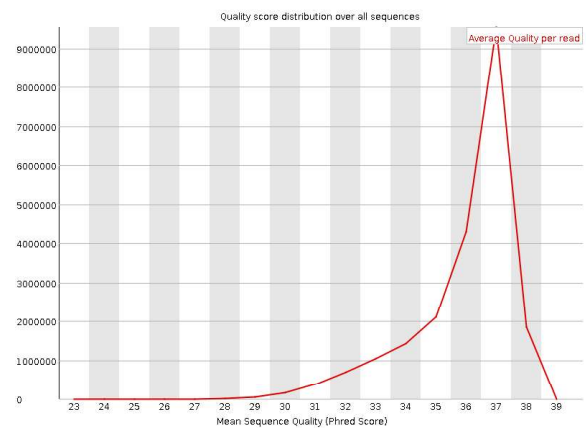


ODG5

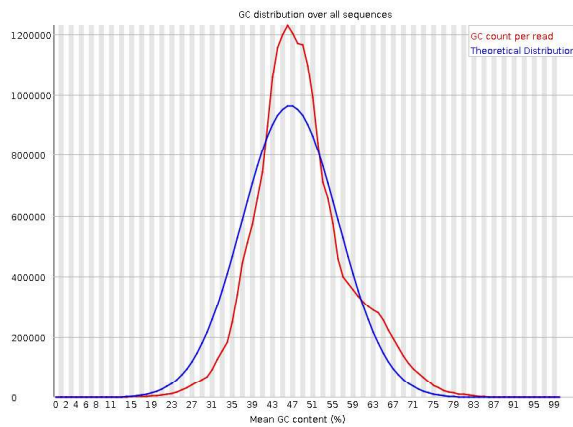
Quality Scores Across All Bases



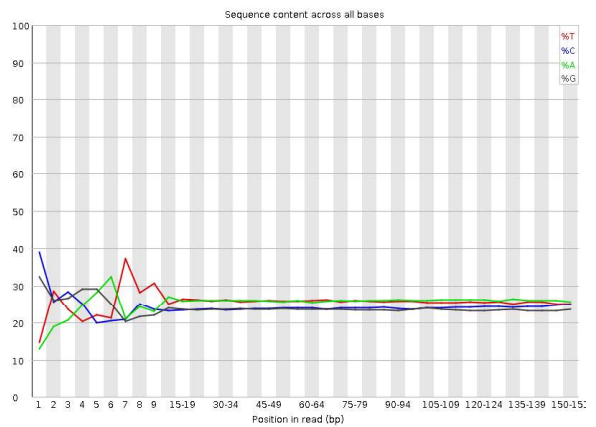
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

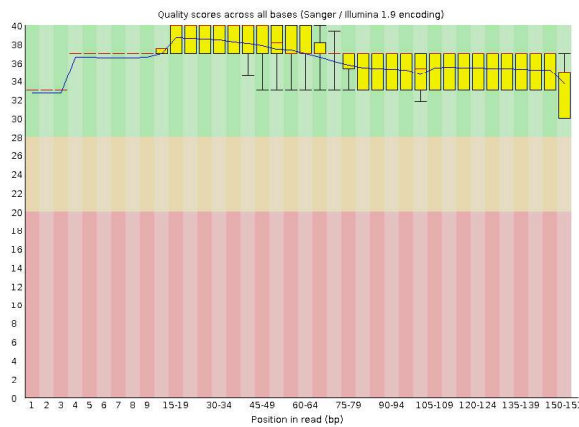


Sequence Content Across All Bases

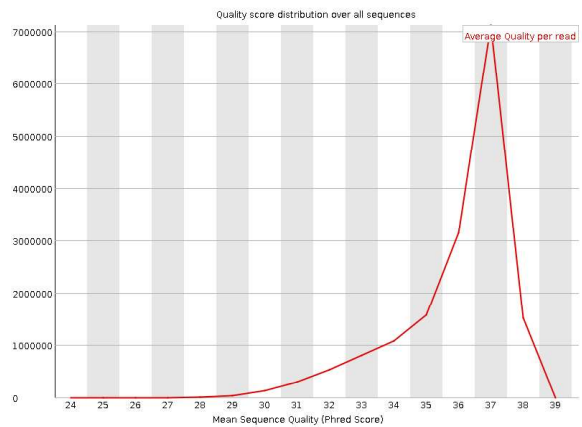


ODG6

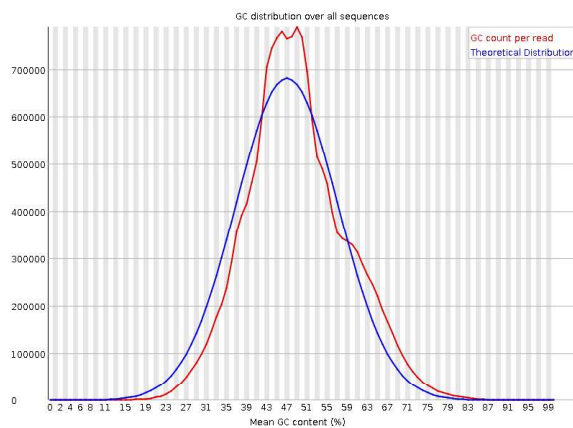
Quality Scores Across All Bases



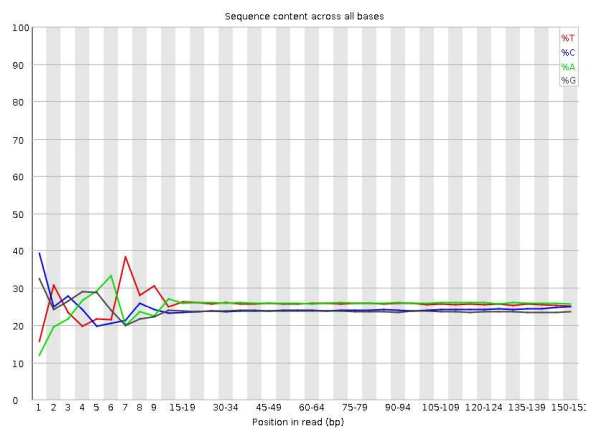
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

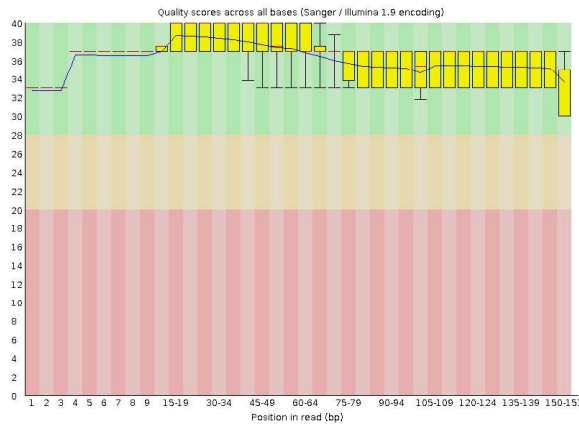


Sequence Content Across All Bases

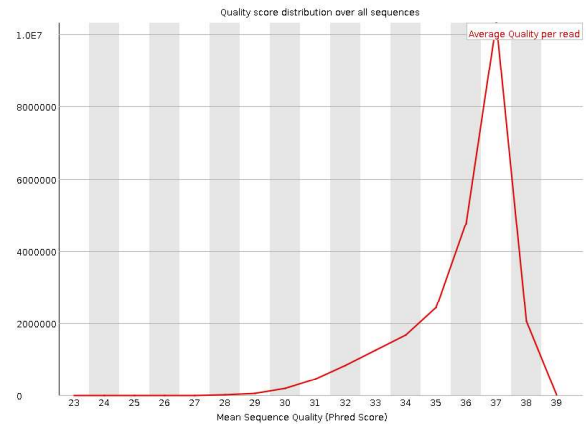


ODG7

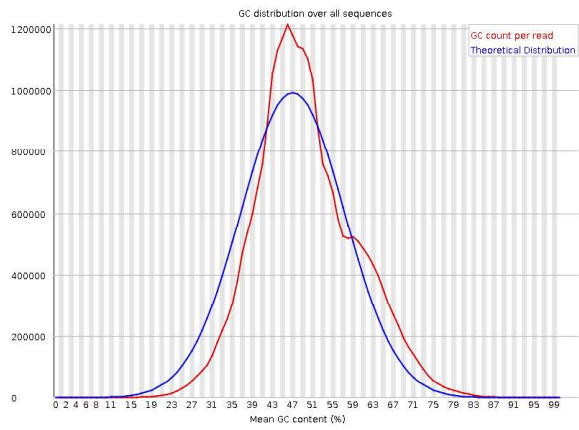
Quality Scores Across All Bases



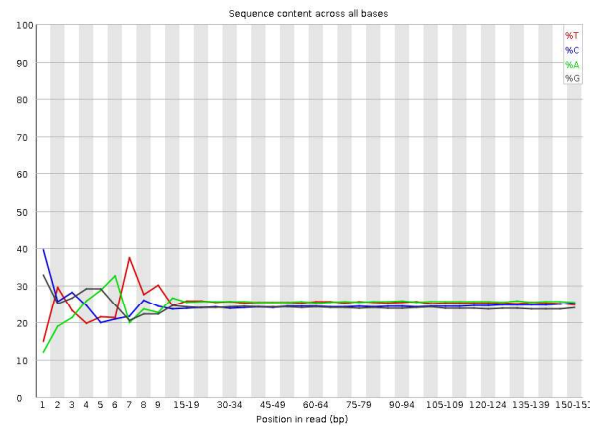
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

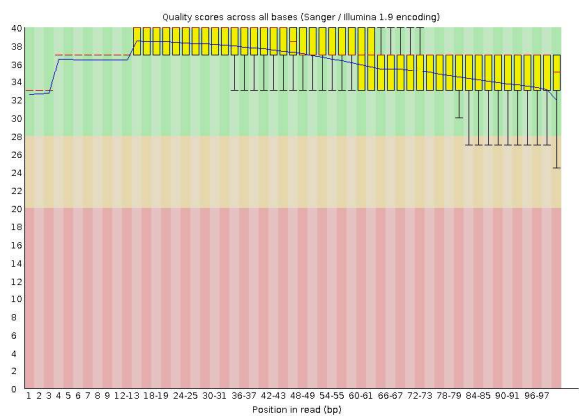


Sequence Content Across All Bases

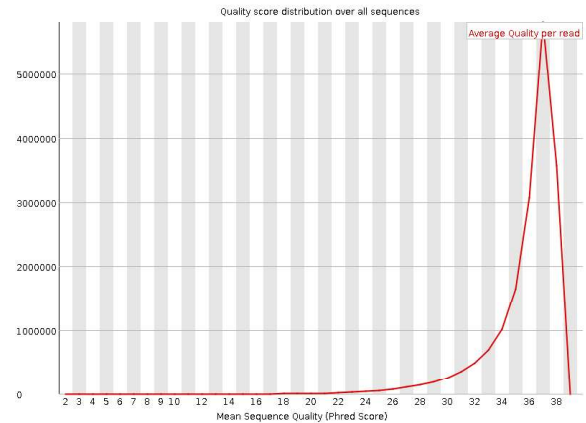


ODG8

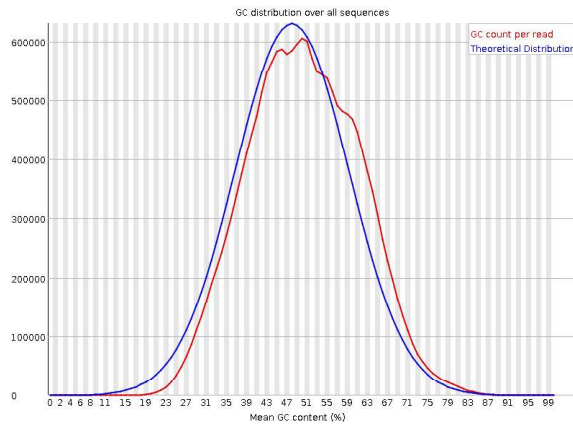
Quality Scores Across All Bases



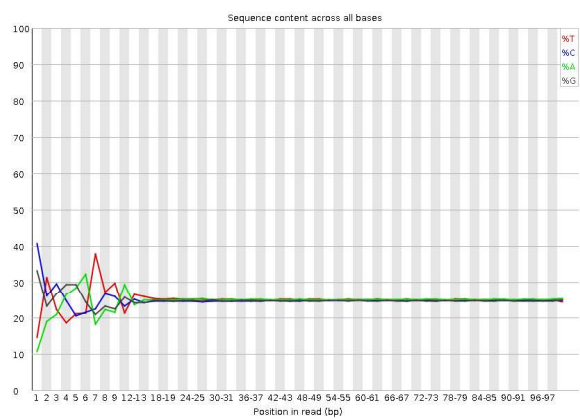
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences

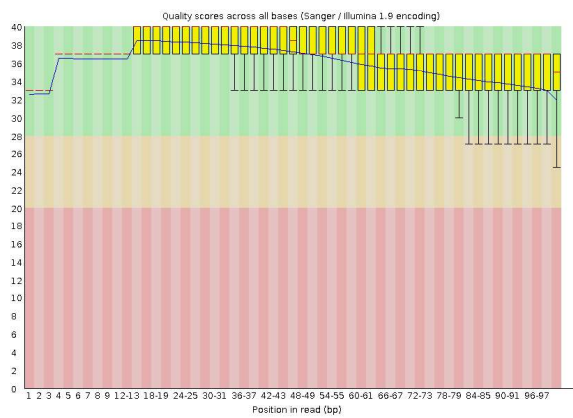


Sequence Content Across All Bases

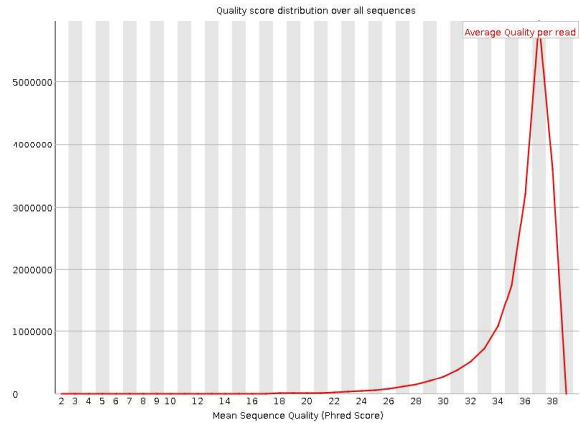


ODG9

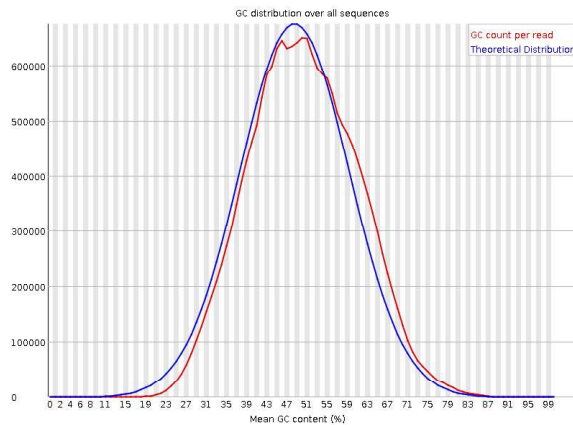
Quality Scores Across All Bases



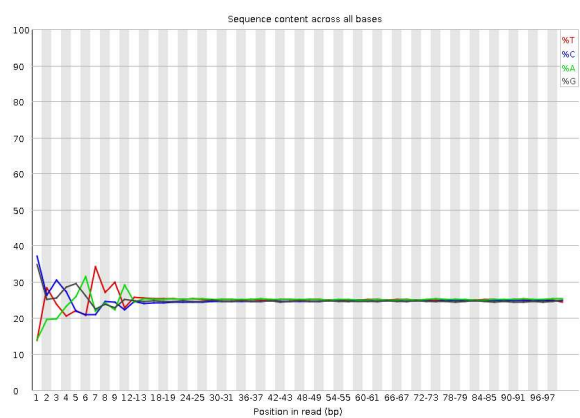
Quality Score Distribution Over All Sequences



GC Distribution Over All Sequences



Sequence Content Across All Bases



Appendix - IV: Somatic Mutations in 11 Tumours

The list of the somatic alterations in the 11 oligodendroglial tumours identified by the VarScan analysis of the exome sequence data. (Reference Assembly hg19)

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG11	<i>AICF</i>	nsSNV	AICF:NM_014576: exon4:c.C146T:p.A49V	chr10	52603836	52603836	G	A	18	0	25	13
ODG1	<i>AAR2</i>	nsSNV	AAR2:NM_001271874: exon3:c.C763T:p.L255F	chr20	34832624	34832624	C	T	60	0	27	13
ODG3	<i>AARD</i>	nsSNV	AARD:NM_001025357: exon1:c.T50G:p.L17R	chr8	117950532	117950532	T	G	38	0	25	24
ODG2	<i>ABCD1</i>	nsSNV	ABCD1:NM_000033: exon5:c.G1418T:p.G473V	chrX	153002635	153002635	G	T	19	0	25	13
ODG6	<i>ABHD14A</i>	sySNV	ABHD14A:NM_015407: exon2:c.C249T:p.Y83Y	chr3	52012066	52012066	C	T	46	0	34	19
ODG6	<i>ABLIM2</i>	sySNV	ABLIM2:NM_001130083: exon7:c.C678T:p.A226A	chr4	8062779	8062779	G	A	48	0	22	15
ODG11	<i>ADAM19</i>	nsSNV	ADAM19:NM_033274: exon16:c.G1886A:p.G629E	chr5	156920003	156920003	C	T	69	0	95	28
ODG6	<i>ADAMTS9</i>	nsSNV	ADAMTS9:NM_182920: exon13:c.G2035A:p.G679R	chr3	64619377	64619377	C	T	76	0	60	35
ODG4	<i>ADCY2</i>	sySNV	ADCY2:NM_020546: exon20:c.T2488C:p.L830L	chr5	7789773	7789773	T	C	27	0	8	11
ODG5	<i>AGO1</i>	nsSNV	AGO1:NM_012199: exon6:c.A715G:p.R239G	chr1	36359703	36359703	A	G	50	0	18	9
ODG1	<i>ALMS1</i>	sySNV	ALMS1:NM_015120: exon10:c.T8484C: p.N2828N	chr2	73717573	73717573	T	C	78	0	50	29
ODG6	<i>ALPK2</i>	nsSNV	ALPK2:NM_052947:	chr18	56203320	56203320	C	T	110	0	14	26

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon5:c.G4099A:p.E1367K									
ODG6	<i>ALS2CL</i>	nsSNV	ALS2CL:NM_001190707: exon12:c.G1264A:p.G422S	chr3	46723000	46723000	C	T	30	0	18	16
ODG1	<i>AMBP</i>	sySNV	AMBP:NM_001633: exon1:c.G75A:p.P25P	chr9	116840415	116840415	C	T	88	0	59	18
ODG9	<i>ANK3</i>	fsDel	ANK3:NM_001149: exon10:c.1383delC: p.S461fs	chr10	61844453	61844453	G	-	63	0	36	18
ODG10	<i>ANKRD12</i>	nsSNV	ANKRD12:NM_001083625: exon8:c.G2313C:p.E771D	chr18	9255647	9255647	G	C	33	0	20	14
ODG10	<i>ANKRD12</i>	nsSNV	ANKRD12:NM_001083625: exon8:c.G1468A:p.E490K	chr18	9254802	9254802	G	A	68	0	33	20
ODG3	<i>AOX1</i>	nsSNV	AOX1:NM_001159: exon12:c.G1123A:p.G375S	chr2	201474107	201474107	G	A	45	0	24	15
ODG11	<i>AP3B2</i>	nsSNV	AP3B2:NM_004644: exon4:c.G301T:p.A101S	chr15	83357547	83357547	C	A	34	0	32	19
ODG11	<i>ARHGAP20</i>	nsSNV	ARHGAP20: NM_001258415: exon15:c.G1826T:p.G609V	chr11	110451775	110451775	C	A	24	1	22	23
ODG3	<i>ARHGAP39</i>	nsSNV	ARHGAP39:NM_025251: exon6:c.C913T:p.R305C	chr8	145773557	145773557	G	A	52	0	28	17
ODG1	<i>ARHGAP9</i>	sySNV	ARHGAP9:NM_001080156: ;	chr12	57868455	57868455	G	A	133	0	68	30
ODG10	<i>ARID1A</i>	fsDel	ARID1A:NM_006015: exon20:c.5684delA:p.E1895 fs	chr1	27106073	27106073	A	-	74	0	39	11

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG6	<i>ARID1A</i>	sySNV	ARID1A:NM_006015: exon15:c.T3771G: p.G1257G	chr1	27099892	27099892	T	G	35	1	6	3
ODG5	<i>ARID1A</i>	fsDel	ARID1A:NM_006015: exon20:c.5577_5578del: p.1859_1860del	chr1	27105966	27105967	CG	-	44	1	13	5
ODG4	<i>ARID3C</i>	nsSNV	ARID3C:NM_001017363: exon4:c.C664T:p.R222C	chr9	34623623	34623623	G	A	134	0	74	42
ODG9	<i>ATP11C</i>	sySNV	ATP11C:NM_001010986: exon5:c.A408C:p.R136R	chrX	138897064	138897064	T	G	15	0	11	5
ODG6	<i>ATP1A2</i>	sySNV	ATP1A2:NM_000702: exon21:c.G2880A:p.A960A	chr1	160109469	160109469	G	A	63	0	32	33
ODG4	<i>ATP2A3</i>	nsSNV	ATP2A3:NM_005173: exon8:c.C743T:p.P248L	chr17	3851037	3851037	G	A	133	0	77	55
ODG10	<i>ATRX</i>	fsDel	ATRX:NM_138270: exon25:c.5953_5957del: p.1985_1986del	chrX	76849205	76849209	AAGAG	-	13	0	3	7
ODG3	<i>AVIL</i>	nsSNV	AVIL:NM_006576: exon6:c.G605A:p.R202H	chr12	58204288	58204288	C	T	39	0	21	20
ODG9	<i>BALAP2L1</i>	nsSNV	BALAP2L1:NM_018842: exon10:c.G1004A:p.G335E	chr7	97937160	97937160	C	T	53	0	30	29
ODG6	<i>BAX</i>	fsIns	BAX:NM_004324: exon3:c.113_114insG:p.M38 fs	chr19	49458970	49458970	-	G	58	0	9	12
ODG6	<i>BEND3</i>	sySNV	BEND3:NM_001080450: exon3:c.A24G:p.E8E	chr6	107420466	107420466	T	C	50	0	22	16
ODG11	<i>BICD2</i>	nsSNV	BICD2:NM_001003800:	chr9	95481707	95481707	T	C	89	0	75	54

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon5:c.A1220G:p.Q407R									
ODG3	<i>BODIL1</i>	sySNV	BODIL1:NM_148894: exon26:c.A9048G:p.T3016T	chr4	13571743	13571743	T	C	69	0	49	43
ODG9	<i>BRWD1</i>	sySNV	BRWD1:NM_018963: exon35:c.A4044C:p.P1348P	chr21	40582712	40582712	T	G	26	0	19	7
ODG11	<i>BSCL2</i>	nsSNV	BSCL2:NM_001122955: exon7:c.C943T:p.L315F	chr11	62458814	62458814	G	A	51	0	42	23
ODG9	<i>BTG2</i>	sySNV	BTG2:NM_006763: exon2:c.C382T:p.L128L	chr1	203276471	203276471	C	T	110	0	41	22
ODG4	<i>BZRAP1</i>	nsSNV	BZRAP1:NM_024418: exon16:c.T1975C:p.S659P	chr17	56390027	56390027	A	G	86	0	52	22
ODG10	<i>C14orf39</i>	nsSNV	C14orf39:NM_174978: exon15:c.T1315G:p.S439A	chr14	60923678	60923678	A	C	13	0	5	7
ODG2	<i>C15orf61</i>	nonfsDel	C15orf61:NM_001143936: exon1:c.150_152del: p.50_51del	chr15	67813736	67813738	CTT	-	51	0	46	17
ODG1	<i>C1orf65</i>	nsSNV	C1orf65:NM_152610: exon1:c.G1094A:p.R365H	chr1	223567911	223567911	G	A	48	0	24	10
ODG1	<i>C2CD3</i>	Splicing	NM_015531: exon7:c.1088+1G>A	chr11	73844469	73844469	C	T	26	0	12	16
ODG6	<i>C5orf34</i>	nsSNV	C5orf34:NM_198566: exon6:c.A1043G:p.N348S	chr5	43502583	43502583	T	C	45	0	28	9
ODG11	<i>CACNA1E</i>	nsSNV	CACNA1E:NM_001205294 ; exon19:c.C2755T:p.R919W	chr1	181702034	181702034	C	T	134	0	112	78
ODG11	<i>CALML5</i>	nsSNV	CALML5:NM_017422: exon1:c.C119T:p.T40M	chr10	5541283	5541283	G	A	79	0	109	72

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG1	<i>CAND2</i>	sySNV	CAND2:NM_001162499: exon1:c.C27T:p.S9S	chr3	12838246	12838246	C	T	16	0	5	8
ODG5	<i>CBLC</i>	sySNV	CBLC:NM_001130852: exon5:c.C83T:p.G277G	chr19	45293312	45293312	C	T	62	0	14	11
ODG6	<i>CCAR1</i>	nsSNV	CCAR1:NM_018237: exon3:c.C214A:p.Q72K	chr10	70496773	70496773	C	A	32	0	7	5
ODG4	<i>CCDC125</i>	nsSNV	CCDC125:NM_176816: exon9:c.G1060A:p.A354T	chr5	68588054	68588054	C	T	26	0	20	10
ODG3	<i>CCDC64</i>	nsSNV	CCDC64:NM_207311: exon4:c.A763G:p.I255V	chr12	120502537	120502537	A	G	59	0	28	19
ODG8	<i>CCNI2</i>	nsSNV	CCNI2:NM_001039780: exon2:c.C521T:p.A174V	chr5	132084130	132084130	C	T	54	0	33	18
ODG10	<i>CDI01</i>	nsSNV	CDI01:NM_001256106: exon9:c.G2975A:p.R992Q	chr1	117576632	117576632	G	A	76	2	51	13
ODG10	<i>CDC27</i>	sySNV	CDC27:NM_001114091: exon6:c.T594C:p.P198P	chr17	45234632	45234632	A	G	61	0	57	28
ODG10	<i>CDC42BPG</i>	nsSNV	CDC42BPG:NM_017525: exon9:c.C1139T:p.P380L	chr11	64605621	64605621	G	A	97	2	36	31
ODG1	<i>CDCA7</i>	sgSNV	CDCA7:NM_031942: exon3:c.G367T:p.E123X	chr2	174224202	174224202	G	T	62	0	33	31
ODG11	<i>CDKN2C</i>	sgSNV	CDKN2C:NM_078626: exon2:c.A163T:p.R55X	chr1	51439598	51439598	A	T	34	0	12	19
ODG11	<i>CEACAM16</i>	nsSNV	CEACAM16:NM_001039213: exon4:c.G421A:p.A141T	chr19	45207326	45207326	G	A	26	0	32	19
ODG8	<i>CENPJ</i>	fsIns	CENPJ:NM_018451: exon4:c.645_646insT:	chr13	25484148	25484148	-	A	34	0	25	9

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG10	CENPT	Splicing	p.T215fs	chr16	67866450	67866451	TG	-	109	0	38	30
ODG8	CFH	nsSNV	CFH:NM_000186: exon12:c.C1861T:p.P621S	chr1	196694415	196694415	C	T	33	0	24	19
ODG1	CIC	nsSNV	CIC:NM_015125: exon5:c.G692A:p.S231N	chr19	42791806	42791806	G	A	78	0	16	23
ODG2	CIC	fsDel	CIC:NM_015125: exon3:c.342delG:p.L114fs	chr19	42791282	42791282	G	-	33	0	18	6
ODG2	CIC	fsDel	CIC:NM_015125: exon2:c.150_151del: p.50_51del	chr19	42791005	42791006	CC	-	43	0	22	8
ODG3	CIC	nsSNV	CIC:NM_015125: exon5:c.C643T:p.R215W	chr19	42791757	42791757	C	T	77	0	1	36
ODG4	CIC	nsSNV	CIC:NM_015125: exon5:c.A614T:p.N205I	chr19	42791728	42791728	A	T	269	1	62	78
ODG10	CLCA4	nsSNV	CLCA4:NM_012128: exon14:c.C2696T:p.T899M	chr1	87045964	87045964	C	T	33	0	25	14
ODG3	CNOT3	nonfsDel	CNOT3:NM_014516: exon18:c.2241_2243del: p.747_748del	chr19	54659124	54659126	GGA	-	36	0	7	25
ODG6	CNTNAP5	nsSNV	CNTNAP5:NM_130773: exon13:c.C1927T:p.R643W	chr2	125405388	125405388	C	T	43	0	27	16
ODG11	COL5A2	sySNV	COL5A2:NM_000393: exon46:c.G3309A:p.P1103P	chr2	189910526	189910526	C	T	67	0	72	56
ODG6	CRYGB	nsSNV	CRYGB:NM_005210: exon3:c.C345G:p.D1115E	chr2	209007545	209007545	G	C	35	0	19	16
ODG3	CSMD3	nonfsDel	CSMD3:NM_052900:	chr8	113259317	113259319	AGA	-	13	0	13	6

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon62:c.9645_9647del: p.3215_3216del									
ODG4	<i>CSNK1D</i>	nsSNV	CSNK1D:NM_001893: exon5:c.G712A:p.E238K	chr17	80210334	80210334	C	T	28	0	25	14
ODG4	<i>CXorf27</i>	nsSNV	CXorf27:NM_012274: exon1:c.C235T:p.R79C	chrX	37850327	37850327	C	T	51	0	27	20
ODG6	<i>CYB56I</i>	nsSNV	CYB56I:NM_001017916: exon5:c.T413G:p.V138G	chr17	61512597	61512597	A	C	21	0	25	7
ODG10	<i>CYP2U1</i>	fsDel	CYP2U1:NM_183075: exon2:c.1080_1081del: p.360_361del	chr4	108866715	108866716	TT	-	49	0	12	24
ODG10	<i>CYTH4</i>	nsSNV	CYTH4:NM_013385: exon10:c.G833A:p.R278Q	chr22	37707053	37707053	G	A	186	0	124	33
ODG10	<i>DAB2</i>	fsDel	DAB2:NM_001244871: exon10:c.1425delG:p.G475fs	chr5	39381572	39381572	C	-	81	0	56	31
ODG10	<i>DCAF13</i>	nsSNV	DCAF13:NM_015420: exon10:c.C1662G:p.S554R	chr8	104453802	104453802	C	G	64	0	35	16
ODG9	<i>DCAF8</i>	nsSNV	DCAF8:NM_015726: exon11:c.A1328G:p.Y443C	chr1	160192553	160192553	T	C	37	1	25	10
ODG5	<i>DCTN6</i>	sySNV	DCTN6:NM_006571: exon3:c.C135T:p.A45A	chr8	30032647	30032647	C	T	79	0	25	26
ODG6	<i>DDX60</i>	nsSNV	DDX60:NM_017631: exon38:c.G5074A:p.D1692N	chr4	169138149	169138149	C	T	53	0	22	13
ODG9	<i>DGKZ</i>	nsSNV	DGKZ:NM_001199268: exon18:c.G1576A:p.D526N	chr11	46396530	46396530	G	A	58	0	31	33

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stop-gain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG7	<i>DISC1</i>	fsDel	DISC1:NM_001164540: exon7:c.1482_1483del: p.494_495del	chr1	231954130	231954131	AG	-	38	0	12	6
ODG6	<i>DLC1</i>	nsSNV	DLC1:NM_001164271: exon5:c.G629C:p.G210A	chr8	12957684	12957684	C	G	103	1	42	31
ODG1	<i>DNAH5</i>	sySNV	DNAH5:NM_001369: exon45:c.G7485A:p.A2495 A	chr5	13810292	13810292	C	T	35	0	13	5
ODG3	<i>DPP3</i>	nsSNV	DPP3:NM_001256670: exon2:c.G40A:p.V14M	chr11	66249711	66249711	G	A	31	0	23	17
ODG4	<i>DSC1</i>	sySNV	DSC1:NM_024421: exon16:c.G2547A:p.S849S	chr18	28710615	28710615	C	T	49	0	38	13
ODG3	<i>DSCAML1</i>	nsSNV	DSCAML1:NM_020693: exon22:c.A4115G:p.N1372S	chr11	117310580	117310580	T	C	33	0	14	14
ODG8	<i>DTX3L</i>	nsSNV	DTX3L:NM_138287: exon3:c.C1621A:p.P541T	chr3	122288557	122288557	C	A	21	0	12	13
ODG1	<i>ELMO3</i>	nsSNV	ELMO3:NM_024712: exon5:c.C526T:p.R176C	chr16	67234218	67234218	C	T	53	0	56	23
ODG2	<i>ENAM</i>	nsSNV	ENAM:NM_031889: exon9:c.C3310A:p.Q1104K	chr4	71510453	71510453	C	A	103	0	54	28
ODG6	<i>ENO3</i>	sySNV	ENO3:NM_001193503: exon5:c.T381C:p.I127I	chr17	4858435	4858435	T	C	26	0	19	5
ODG5	<i>EPRS</i>	nsSNV	EPRS:NM_004446: exon9:c.G1046T:p.G349V	chr1	220195758	220195758	C	A	45	0	18	8
ODG10	<i>ERCC4</i>	nsSNV	ERCC4:NM_005236: exon1:c.T80G:p.L27R	chr16	14014102	14014102	T	G	143	0	89	48
ODG10	<i>ERMN</i>	fsDel	ERMN:NM_020711:	chr2	158182079	158182082	TTGT	-	68	0	42	31

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon1:c.73_76del: p.25_26del									
ODG8	<i>F2</i>	nsSNV	F2:NM_000506: exon10:c.G1292A:p.R431H	chr11	46749707	46749707	G	A	171	0	73	65
ODG11	<i>FAM135B</i>	nsSNV	FAM135B:NM_015912: exon13:c.G2737T:p.V913L	chr8	139163981	139163981	C	A	25	0	27	19
ODG4	<i>FAM170A</i>	nsSNV	FAM170A:NM_001163991: exon2:c.G295A:p.E99K	chr5	118969879	118969879	G	A	56	0	34	29
ODG10	<i>FAM186B</i>	nonfsDel	FAM186B:NM_032130: exon4:c.1516_1518del: p.506_506del	chr12	49993905	49993907	CTT	-	281	0	116	73
ODG3	<i>FAM83H</i>	sySNV	FAM83H:NM_198488: exon5:c.C1449T:p.F483F	chr8	144810182	144810182	G	A	27	0	21	13
ODG6	<i>FBRS</i>	sgSNV	FBRS:NM_001105079: exon3:c.C40T:p.R14X	chr16	30676392	30676392	C	T	50	0	18	31
ODG11	<i>FBXL19</i>	nsSNV	FBXL19:NM_001099784: exon5:c.C473T:p.S158L	chr16	30939070	30939070	C	T	63	0	99	40
ODG11	<i>FBXL19</i>	nsSNV	FBXL19:NM_001099784: exon4:c.C410T:p.A137V	chr16	30938922	30938922	C	T	152	0	195	101
ODG8	<i>FBXO24</i>	nonfsDel	FBXO24:NM_001163499: exon10:c.1619_1621del: p.540_541del	chr7	100198434	100198436	AGA	-	143	0	85	33
ODG11	<i>FGFR1</i>	nsSNV	FGFR1:NM_001174066: exon13:c.A1699G:p.K567E	chr8	38272308	38272308	T	C	41	0	31	13
ODG11	<i>FGL1</i>	nsSNV	FGL1:NM_004467: exon7:c.G692T:p.R231I	chr8	17726144	17726144	C	A	36	0	38	27
ODG2	<i>FKBP8</i>	nsSNV	FKBP8:NM_012181:	chr19	18643604	18643604	G	A	52	1	36	29

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon8:c.C1025T;p.T342M									
ODG3	<i>FNBP1</i>	nsSNV	FNBP1:NM_015033: exon10:c.G1168T;p.G390W	chr9	132686125	132686125	C	A	64	0	24	13
ODG1	<i>FRMD4B</i>	sySNV	FRMD4B:NM_015123: exon22:c.T2859C;p.S953S	chr3	69225800	69225800	A	G	21	0	11	6
ODG3	<i>FSCB</i>	nsSNV	FSCB:NM_032135: exon1:c.G661C;p.E221Q	chr14	44975530	44975530	C	G	97	0	56	35
ODG1	<i>FUBP1</i>	sgSNV	FUBP1:NM_003902: exon16:c.G1530A;p.W510X	chr1	78425915	78425915	C	T	35	0	7	19
ODG5	<i>FUBP1</i>	nonfsDel	FUBP1:NM_003902: exon3:c.248_250del: p.83_84del	chr1	78433848	78433851	CAGT	-	46	0	17	9
ODG4	<i>GABPA</i>	nsSNV	GABPA:NM_001197297: exon3:c.A79G;p.L27V	chr21	27117522	27117522	A	G	56	0	32	32
ODG4	<i>GCOM1</i> , <i>MYZAP</i>	nsSNV	GCOM1:NM_001018090: exon4:c.G350A;p.R117Q	chr15	57913837	57913837	G	A	41	0	29	18
ODG11	<i>GIT1</i>	sySNV	GIT1:NM_001085454: exon6:c.C678T;p.L226L	chr17	27908800	27908800	G	A	77	0	27	88
ODG4	<i>GPR116</i>	sySNV	GPR116:NM_001098518: exon7:c.C627T;p.Y209Y	chr6	46849830	46849830	G	A	37	0	26	20
ODG8	<i>GRAMD4</i>	sySNV	GRAMD4:NM_015124: exon14:c.G1281A;p.P427P	chr22	47069608	47069608	G	A	63	0	52	13
ODG11	<i>GRM6</i>	sySNV	GRM6:NM_000843: exon8:c.C1905T;p.Y635Y	chr5	178413350	178413350	G	A	130	0	124	90
ODG6	<i>HDAC9</i>	nsSNV	HDAC9:NM_001204144: exon4:c.G115C;p.A39P	chr7	18535914	18535914	G	C	110	0	101	36
ODG6	<i>HDHD3</i>	nsSNV	HDHD3:NM_031219:	chr9	116136261	116136261	C	T	64	0	34	16

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon2:c.G374A:p.R125H									
			HIST1H1C:NM_005319: exon1:c.67_69del: p.23_23del	chr6	26056588	26056590	CTT	-	34	0	21	8
ODG2	HIST1H1C	nonfsDel										
ODG2	HPN	nsSNV	HPN:NM_182983: exon9:c.C662T:p.A221V	chr19	35551572	35551572	C	T	65	0	27	14
ODG11	HRG	nsSNV	HRG:NM_000412: exon7:c.G1517A:p.G506E	chr3	186395611	186395611	G	A	21	0	18	5
ODG5	HSP49	nonfsDel	HSPA9:NM_004134: exon16:c.1839_1841del: p.613_614del	chr5	137892262	137892264	TCT	-	35	0	21	8
ODG7	HTR34	nsSNV	HTR3A:NM_213621: exon8:c.C1493T:p.A498V	chr11	113860427	113860427	C	T	116	0	56	29
ODG9	HTR4	sySNV	HTR4:NM_001040169: exon6:c.C1113T:p.L371L	chr5	147830799	147830799	G	A	36	0	17	9
ODG11	HYOU1	nsSNV	HYOU1:NM_001130991: exon14:c.C1538T:p.S513F	chr11	118921874	118921874	G	A	31	0	24	17
ODG1	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	7	0	3	6
ODG10	IDH1	nsSNV	IDH1:NM_005896: exon4:c.C394T:p.R132C	chr2	209113113	209113113	G	A	11	0	5	5
ODG2	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	9	0	6	5
ODG3	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	12	0	3	10
ODG4	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	13	0	15	5

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG5	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	19	0	11	10
ODG6	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	9	0	4	8
ODG7	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	7	0	8	1
ODG8	IDH1	nsSNV	IDH1:NM_005896: exon4:c.G395A:p.R132H	chr2	209113112	209113112	C	T	8	0	9	5
ODG9	IDH2	nsSNV	IDH2:NM_002168: exon4:c.G515A:p.R172K	chr15	90631838	90631838	C	T	55	0	9	19
ODG5	IGFN1	nsSNV	IGFN1:NM_001164586: exon11:c.C1076T:p.S359F	chr1	201174369	201174369	C	T	54	0	24	7
ODG10	IGSF10	nsSNV	IGSF10:NM_178822: exon4:c.C1579T:p.R527W	chr3	151166190	151166190	G	A	73	0	41	29
ODG6	IGSF10	nsSNV	IGSF10:NM_178822: exon4:c.G1642A:p.V548I	chr3	151166127	151166127	C	T	114	0	69	71
ODG10	IRF2BPL	sySNV	IRF2BPL:NM_024496: exon1:c.G348A:p.Q116Q	chr14	77493788	77493788	C	T	15	0	8	8
ODG7	ITGAL	nsSNV	ITGAL:NM_001114380: exon22:c.C2501T:p.P834L	chr16	30522424	30522424	C	T	44	0	26	17
ODG6	ITGAX	sySNV	ITGAX:NM_000887: exon29:c.G3375A: p.A1125A	chr16	31392316	31392316	G	A	73	0	41	35
ODG4	ITGB8	sySNV	ITGB8:NM_002214: exon4:c.C504T:p.S168S	chr7	20418789	20418789	C	T	41	0	23	11
ODG11	ITIH3	nsSNV	ITIH3:NM_002217: exon5:c.C389T:p.A130V	chr3	52831123	52831123	C	T	26	1	28	18

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG10	<i>JAKMIP1</i>	nsSNV	JAKMIP1:NM_001099433: exon2:c.C47T:p.T16M	chr4	6114531	6114531	G	A	382	0	202	106
ODG6	<i>KCNIP2</i>	nsSNV	KCNIP2:NM_173197: exon5:c.C398T:p.A133V	chr10	103587692	103587692	G	A	75	0	56	40
ODG4	<i>KCNK15</i>	Splicing	NM_022358: exon2:c.284-2A>G	chr20	43378768	43378768	A	G	93	0	33	20
ODG4	<i>KCNS1</i>	nsSNV	KCNS1:NM_002251: exon4:c.C577T:p.R193C	chr20	43726836	43726836	G	A	50	0	34	17
ODG2	<i>KCNV2</i>	nsSNV	KCNV2:NM_133497: exon1:c.C833T:p.A278V	chr9	2718572	2718572	C	T	30	0	22	22
ODG1	<i>KDM6B</i>	sySNV	KDM6B:NM_001080424: exon20:c.C4500T: p.Y1500Y	chr17	7755844	7755844	C	T	75	0	49	27
ODG6	<i>KI4A0753</i>	nsSNV	KIAA0753:NM_014804: exon8:c.G1387A:p.D463N	chr17	6515397	6515397	C	T	69	0	39	44
ODG4	<i>KI4A1024</i>	sySNV	KIAA1024:NM_015206: exon2:c.C699A:p.I233I	chr15	79749188	79749188	C	A	87	0	39	39
ODG1	<i>KIF18B</i>	nsSNV	KIF18B:NM_001264573: exon3:c.C415T:p.R139C	chr17	43012683	43012683	G	A	47	0	25	14
ODG1	<i>KIF1A</i>	sySNV	KIF1A:NM_004321: exon43:c.C4684T:p.L1562L	chr2	241659225	241659225	G	A	35	0	20	13
ODG10	<i>KIF3A</i>	nsSNV	KIF3A:NM_007054: exon2:c.A40G:p.N14D	chr5	132070137	132070137	T	C	53	0	20	15
ODG6	<i>KIF4B</i>	nsSNV	KIF4B:NM_001099293: exon1:c.G2183A:p.R728Q	chr5	154395602	154395602	G	A	228	0	111	66
ODG6	<i>KLF4</i>	nsSNV	KLF4:NM_004235: exon5:c.A1427C:p.K476T	chr9	110248045	110248045	T	G	56	0	17	15

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG8	<i>KRAS</i>	nsSNV	KRAS:NM_004985: exon3:c.A182T:p.Q61L	chr12	25380276	25380276	T	A	87	1	58	30
ODG9	<i>KRAS</i>	nsSNV	KRAS:NM_004985: exon2:c.G35A:p.G12D	chr12	25398284	25398284	C	T	45	0	20	44
ODG1	<i>KRT81</i>	nsSNV	KRT81:NM_002281: exon9:c.G1453A:p.V485M	chr12	52680104	52680104	C	T	30	0	15	12
ODG3	<i>KRT9</i>	sgSNV	KRT9:NM_000226: exon7:c.C1560A:p.Y520X	chr17	39723837	39723837	G	T	61	0	50	40
ODG10	<i>KRTAP15-1</i>	nsSNV	KRTAP15-1:NM_181623: exon1:c.G17A:p.S6N	chr21	31812662	31812662	G	A	24	0	17	12
ODG4	<i>KXD1</i>	sySNV	KXD1:NM_001171949: exon5:c.G531A:p.X177X	chr19	18679441	18679441	G	A	62	0	50	19
ODG11	<i>LGR6</i>	sySNV	LGR6:NM_001017404: exon1:c.G9A:p.L3L	chr1	202183291	202183291	G	A	69	0	71	45
ODG9	<i>LMO7</i>	nsSNV	LMO7:NM_015842: exon10:c.G1433A:p.R478K	chr13	76393596	76393596	G	A	50	0	41	17
ODG4	<i>LMOD2</i>	nsSNV	LMOD2:NM_207163: exon2:c.C419A:p.T140K	chr7	123302059	123302059	C	A	74	0	64	35
ODG10	<i>LOXHD1</i>	nsSNV	LOXHD1:NM_144612: exon11:c.A1484G:p.K495R	chr18	44172495	44172495	T	C	58	0	35	10
ODG1	<i>LPIN1</i>	nsSNV	LPIN1:NM_145693: exon19:c.T2507G:p.I836S	chr2	11960634	11960634	T	G	70	0	31	25
ODG4	<i>LRP1B</i>	nsSNV	LRP1B:NM_018557: exon17:c.G2765A:p.C922Y	chr2	141747106	141747106	C	T	18	0	11	9
ODG3	<i>LRRC16A</i>	nsSNV	LRRC16A:NM_001173977: exon19:c.G1517A:p.S506N	chr6	25510774	25510774	G	A	50	1	32	28
ODG8	<i>MAML3</i>	fsDel	MAML3:NM_018717:	chr4	140811064	140811075	TGCTG	-	68	0	40	57

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon3:c.1513_1514del: p.505_505del				CTGCT GC					
ODG10	<i>MBD6</i>	fsDel	MBD6:NM_052897: exon10:c.2440delC; p.P814fs	chr12	57921963	57921963	C	-	29	0	22	20
ODG1	<i>MED23</i>	Splicing	NM_015979: exon4:c.72-1G>A	chr6	131948624	131948624	C	T	74	0	55	30
ODG8	<i>MESP2</i>	nonfsDel	MESP2:NM_001039958: exon1:c.547_558del: p.183_186del	chr15	90320135	90320146	GGGCA GGGGC AG	-	14	0	1	27
ODG11	<i>MMP17</i>	nsSNV	MMP17:NM_016155: exon8:c.G1090C:p.V364L	chr12	132329880	132329880	G	C	48	0	47	14
ODG4	<i>MMRN1</i>	Splicing	NM_007351: exon7:c.3265+2->A	chr4	90872904	90872904	-	A	51	2	24	15
ODG3	<i>MRPL39</i>	nsSNV	MRPL39:NM_017446: exon4:c.T443C:p.M148T	chr21	26973757	26973757	A	G	81	0	40	40
ODG4	<i>MTIF3</i>	nsSNV	MTIF3:NM_001166261: exon3:c.G232A:p.V78I	chr13	28014354	28014354	C	T	180	1	89	93
ODG11	<i>MUC5B</i>	nsSNV	MUC5B:NM_002458: exon31:c.C5723T:p.T1908I	chr11	1263833	1263833	C	T	297	0	287	155
ODG3	<i>MYCBPAP</i>	nsSNV	MYCBPAP:NM_032133: exon7:c.C985T:p.R329C	chr17	48597088	48597088	C	T	52	0	14	12
ODG2	<i>MYH10</i>	nonfsDel	MYH10:NM_005964: exon40:c.5675_5677del: p.1892_1893del	chr17	8380303	8380305	CTT	-	45	0	55	15
ODG11	<i>MYL7</i>	nsSNV	MYL7:NM_021223: exon4:c.G211C:p.E71Q	chr7	44180009	44180009	C	G	91	0	97	63

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG3	<i>MYO15A</i>	nsSNV	MYO15A:NM_016239: exon32:c.C6863T:p.S2288L	chr17	18052173	18052173	C	T	24	0	18	16
ODG4	<i>N4BP2</i>	nsSNV	N4BP2:NM_018177: exon8:c.G1724A:p.R575H	chr4	40119548	40119548	G	A	31	0	27	7
ODG6	<i>NBEA</i>	nsSNV	NBEA:NM_001204197: exon10:c.C1020G:p.N340K	chr13	36220419	36220419	C	G	92	0	50	29
ODG1	<i>NDST2</i>	sySNV	NDST2:NM_003635: exon5:c.C1176T:p.H392H	chr10	75566487	75566487	G	A	59	0	41	22
ODG10	<i>NEIL3</i>	nsSNV	NEIL3:NM_018248: exon6:c.A811G:p.R271G	chr4	178262738	178262738	A	G	40	0	21	17
ODG9	<i>NETO2</i>	nsSNV	NETO2:NM_001201477: exon4:c.A266G:p.D89G	chr16	47162451	47162451	T	C	39	0	31	11
ODG11	<i>NF1</i>	fsDel	NF1:NM_000267: exon8:c.797delT:p.V266fs	chr17	29509592	29509592	T	-	21	0	5	43
ODG1	<i>NFATC2IP</i>	nsSNV	NFATC2IP:NM_032815: exon1:c.G119A:p.R40Q	chr16	28962451	28962451	G	A	40	0	16	15
ODG4	<i>NHSL2</i>	nsSNV	NHSL2:NM_001013627: exon6:c.G1379A:p.R460H	chrX	71358777	71358777	G	A	86	0	79	40
ODG10	<i>NIPBL</i>	nonfsDel	NIPBL:NM_015384: exon38:c.6556_6558del: p.2186_2186del	chr5	37046268	37046270	GAA	-	39	0	25	22
ODG4	<i>NIPBL</i>	nsSNV	NIPBL:NM_015384: exon9:c.A1151G:p.N384S	chr5	36976160	36976160	A	G	45	0	19	24
ODG2	<i>NOTCH1</i>	nonfsDel	NOTCH1:NM_017617: exon6:c.1070_1072del: p.357_358del	chr9	139413070	139413072	AGA	-	32	0	28	15
ODG3	<i>NOTCH1</i>	nonfsDel	NOTCH1:NM_017617:	chr9	139413070	139413072	AGA	-	31	0	29	8

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon6:c.1070_1072del: p.357_358del									
ODG4	<i>NOTCH1</i>	nsSNV	NOTCH1:NM_017617: exon27:c.G5110A: p.G1704R	chr9	139397691	139397691	C	T	100	0	71	38
ODG6	<i>NOTCH1</i>	nsSNV	NOTCH1:NM_017617: exon7:c.A1163G:p.D388G	chr9	139412681	139412681	T	C	22	0	7	5
ODG4	<i>NRK</i>	sySNV	NRK:NM_198465: exon16:c.G2481A:p.S827S	chrX	105161633	105161633	G	A	32	0	17	22
			NSD1:NM_022455: exon7:c.A4012G: p.T1338A									
ODG6	<i>NSD1</i>	nsSNV		chr5	176665328	176665328	A	G	60	0	27	25
ODG1	<i>OCNL</i>	sySNV	OCNL:NM_001205255: exon2:c.G96A:p.K32K	chr5	68809894	68809894	G	A	54	0	38	12
ODG11	<i>OR1E2</i>	nsSNV	OR1E2:NM_003554: exon1:c.G7A:p.G3R	chr17	3337129	3337129	C	T	74	1	16	30
ODG11	<i>OR2Y1</i>	sgSNV	OR2Y1:NM_001001657: exon1:c.C604T:p.R202X	chr5	180166455	180166455	G	A	33	0	27	13
ODG9	<i>OR4S2</i>	nsSNV	OR4S2:NM_001004059: exon1:c.C382A:p.L128I	chr11	55418761	55418761	C	A	86	0	61	21
ODG4	<i>OR52E2</i>	nsSNV	OR52E2:NM_001005164: exon1:c.G683A:p.R228H	chr11	5080175	5080175	C	T	75	0	57	23
ODG8	<i>OR5D14</i>	sySNV	OR5D14:NM_001004735: exon1:c.C894T:p.D298D	chr11	55563925	55563925	C	T	67	0	39	26
ODG10	<i>OR5J2</i>	nsSNV	OR5J2:NM_001005492: exon1:c.A37G:p.I13V	chr11	55944130	55944130	A	G	25	0	19	9
ODG6	<i>PAX4</i>	nsSNV	PAX4:NM_006193:	chr7	127254980	127254980	C	T	44	0	43	17

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon2:c.G290A:p.R97H									
ODG2	<i>PCDHGA11</i>	sySNV	PCDHGA11:NM_018914: exon1:c.G1404A:p.R468R	chr5	140802198	140802198	G	A	152	0	148	42
ODG1	<i>PGR</i>	sySNV	PGR:NM_000926: exon1:c.G30A:p.R10R	chr11	100999772	100999772	C	T	15	0	15	7
ODG4	<i>PHF12</i>	sySNV	PHF12:NM_001033561: exon13:c.G2541A:p.E847E	chr17	27234608	27234608	C	T	28	0	11	12
ODG1	<i>PHLPP2</i>	sySNV	PHLPP2:NM_015020: exon4:c.A660G:p.G220G	chr16	71718454	71718454	T	C	55	0	27	7
ODG6	<i>PID1</i>	sySNV	PID1:NM_017933: exon2:c.C18T:p.G6G	chr2	230127504	230127504	G	A	81	0	49	39
ODG6	<i>PLA2G6</i>	sySNV	PLA2G6:NM_001004426: exon7:c.G957A:p.T319T	chr22	38528958	38528958	C	T	51	0	36	17
ODG3	<i>PNRC1</i>	nsSNV	PNRC1:NM_006813: exon1:c.A265G:p.T89A	chr6	89790878	89790878	A	G	24	0	16	17
ODG6	<i>POU2AF1</i>	nsSNV	POU2AF1:NM_006235: exon4:c.G261T:p.W87C	chr11	111228365	111228365	C	A	30	0	35	9
ODG9	<i>PPF1A1</i>	sgSNV	PPF1A1:NM_003626: exon24:c.C3163T: p.R1055X	chr11	70221047	70221047	C	T	36	1	11	10
ODG4	<i>PQBP1</i>	nsSNV	PQBP1:NM_001167992: exon6:c.A422C:p.Q141P	chrX	48760285	48760285	A	C	51	0	27	31
ODG11	<i>PREPL</i>	nsSNV	PREPL:NM_001042385: exon5:c.G688A:p.V230I	chr2	44569620	44569620	C	T	21	0	23	8
ODG11	<i>PRKCDDBP</i>	sySNV	PRKCDDBP:NM_145040: exon2:c.G573A:p.R191R	chr11	6340606	6340606	C	T	10	0	23	10
ODG10	<i>PROSER1</i>	nsSNV	PROSER1:NM_025138:	chr13	39586949	39586949	T	C	203	0	158	108

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon1:c.A2440G; p.T814A									
ODG6	<i>PRSS38</i>	sySNV	PRSS38:NM_183062: exon2:c.G198T:p.A66A	chr1	228003840	228003840	G	T	21	0	18	12
ODG11	<i>PRX</i>	sySNV	PRX:NM_181882: exon7:c.G498A:p.L166L	chr19	40903761	40903761	C	T	131	0	111	71
ODG9	<i>PTBP3</i>	nsSNV	PTBP3:NM_001244896: exon13:c.T1351C:p.S451P	chr9	114986154	114986154	A	G	28	0	19	7
ODG6	<i>PTCHD3</i>	sySNV	PTCHD3:NM_001034842: exon1:c.G288A:p.L96L	chr10	27702892	27702892	C	T	58	0	45	26
ODG10	<i>PTGDR2</i>	nsSNV	PTGDR2:NM_004778: exon2:c.C811T:p.R271W	chr11	60620385	60620385	G	A	65	0	40	18
ODG4	<i>PTPRG</i>	nsSNV	PTPRG:NM_002841: exon7:c.G785A:p.S262N	chr3	62142843	62142843	G	A	64	0	45	33
ODG4	<i>PTPRM</i>	nsSNV	PTPRM:NM_001105244: exon7:c.G847A:p.V283I	chr18	7955127	7955127	G	A	20	0	12	11
ODG4	<i>QPCTL</i>	sySNV	QPCTL:NM_001163377: exon6:c.G750A:p.T250T	chr19	46206190	46206190	G	A	19	1	2	9
			RAB11FIP1:NM_001002814: 4:									
ODG2	<i>RAB11FIP1</i>	nsSNV	exon3:c.C1415T:p.P472L	chr8	37732240	37732240	G	A	125	0	130	44
ODG9	<i>RAB3GAP2</i>	nsSNV	RAB3GAP2:NM_012414: exon12:c.G1043A:p.G348D	chr1	220368507	220368507	C	T	41	0	29	11
ODG3	<i>RASSF7</i>	nsSNV	RASSF7:NM_001143993: exon3:c.G391A:p.E131K	chr11	562345	562345	G	A	70	0	39	21
ODG3	<i>RBBP6</i>	fsDel	RBBP6:NM_018703: exon16:c.2545_2546del:	chr16	24580658	24580659	TA	-	113	0	80	40

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			p.849_849del									
ODG4	<i>RBMXL3</i>	nsSNV	RBMXL3:NM_001145346: exon1:c.G1192A:p.D398N	chrX	114425196	114425196	G	A	206	0	213	152
ODG4	<i>RBMXL3</i>	sySNV	RBMXL3:NM_001145346: exon1:c.G1206A:p.G402G	chrX	114425210	114425210	G	A	255	0	280	101
ODG7	<i>RCCD1</i>	nsSNV	RCCD1:NM_001017919: exon5:c.G740A:p.R247K	chr15	91503190	91503190	G	A	33	0	11	7
ODG11	<i>RCSD1</i>	sySNV	RCSD1:NM_052862: exon3:c.G141A:p.P47P	chr1	167654690	167654690	G	A	96	2	115	59
ODG3	<i>RELN</i>	nsSNV	RELN:NM_005045: exon20:c.G2684A:p.G895D	chr7	103270405	103270405	C	T	57	0	29	25
ODG6	<i>RGMB</i>	nsSNV	RGMB:NM_001012761: exon5:c.A802G:p.T268A	chr5	98128822	98128822	A	G	44	0	40	20
ODG2	<i>RIF1</i>	sySNV	RIF1:NM_001177665: exon11:c.G1248A:p.S416S	chr2	152293393	152293393	G	A	33	0	35	9
ODG11	<i>RNF111</i>	sySNV	RNF111:NM_001270528: exon7:c.C1792A:p.R598R	chr15	59368258	59368258	C	A	16	0	15	13
ODG10	<i>RPL22</i>	nsSNV	RPL22:NM_000983: exon3:c.C230G:p.P77R	chr1	6253002	6253002	G	C	29	0	15	18
ODG3	<i>RPL6</i>	sySNV	RPL6:NM_000970: exon4:c.C438A:p.P146P	chr12	112844593	112844593	G	T	22	0	21	13
ODG4	<i>RYR2</i>	sySNV	RYR2:NM_001035: exon90:c.G12120A: p.K4040K	chr1	237947132	237947132	G	A	45	0	30	15
ODG10	<i>SEPT9</i>	sySNV	SEPT9:NM_001113493: exon2:c.C111T:p.P37P	chr17	75398196	75398196	C	T	68	0	46	18
ODG4	<i>SERBP1</i>	nsSNV	SERBP1:NM_001018067:	chr1	67880955	67880955	A	G	67	0	18	38

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon7:c.T1064C:p.L355P									
ODG5	<i>SERPINEB13</i>	nsSNV	SERPINEB13:NM_012397: exon6:c.C493T:p.P165S	chr18	61261609	61261609	C	T	34	0	29	8
ODG2	<i>SERPINE3</i>	nsSNV	SERPINE3:NM_001101320: exon7:c.C1163T:p.P388L	chr13	51936021	51936021	C	T	67	0	40	28
ODG6	<i>SF3B1</i>	nsSNV	SF3B1:NM_012433: exon16:c.G2252T:p.G751V	chr2	198266584	198266584	C	A	22	0	16	6
ODG11	<i>SF3B3</i>	sySNV	SF3B3:NM_012426: exon15:c.G1977A:p.S659S	chr16	70590899	70590899	G	A	33	0	47	14
ODG11	<i>SGSM1</i>	nsSNV	SGSM1:NM_001039948: exon8:c.G691T:p.G231C	chr22	25251537	25251537	G	T	99	3	39	81
ODG11	<i>SHKBP1</i>	nsSNV	SHKBP1:NM_138392: exon9:c.C821T:p.A274V	chr19	41086819	41086819	C	T	47	0	36	28
ODG6	<i>SI</i>	nsSNV	SI:NM_001041: exon13:c.T1482A:p.H494Q	chr3	164773012	164773012	A	T	32	0	28	16
ODG11	<i>SIN3A</i>	nsSNV	SIN3A:NM_001145357: exon7:c.T1127G:p.F376C	chr15	75702509	75702509	A	C	24	0	13	11
ODG4	<i>SKIDA1</i>	nsSNV	SKIDA1:NM_207371: exon4:c.G1745A:p.S582N	chr10	21805007	21805007	C	T	29	0	16	5
ODG2	<i>SLC12A7</i>	sySNV	SLC12A7:NM_006598: exon19:c.C2562T:p.D854D	chr5	1064243	1064243	G	A	66	0	51	47
ODG11	<i>SLC16A3</i>	sySNV	SLC16A3:NM_001042422: exon5:c.G1383A:p.P461P	chr17	80196837	80196837	G	A	39	1	10	36
ODG4	<i>SLC22A9</i>	nsSNV	SLC22A9:NM_080866: exon6:c.C1006T:p.P336S	chr11	63149682	63149682	C	T	28	0	19	18
ODG10	<i>SLC25A2</i>	nsSNV	SLC25A2:NM_031947: exon1:c.A199G:p.T67A	chr5	140683234	140683234	T	C	22	0	12	8

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
ODG10	<i>SLC25A47</i>	sySNV	SLC25A47:NM_207117: exon6:c.C846T:p.C282C	chr14	100795901	100795901	C	T	105	0	51	31
ODG6	<i>SLC34A1</i>	Splicing	NM_001167579: exon3:c.110-2A>T	chr5	176812986	176812986	A	T	63	0	37	26
ODG11	<i>SLC7A10</i>	sySNV	SLC7A10:NM_019849: exon2:c.C252T:p.G84G	chr19	33706779	33706779	G	A	48	1	45	29
ODG11	<i>SLC9A3</i>	nsSNV	SLC9A3:NM_004174: exon5:c.C827T:p.T276M	chr5	484740	484740	G	A	125	2	99	61
ODG9	<i>SLC9C1</i>	Splicing	NM_183061: exon9:c.776-1G>A	chr3	111985188	111985188	C	T	31	0	23	7
ODG3	<i>SLITRK2</i>	fsDel	SLITRK2:NM_001144009: exon2:c.76_77del: p.26_26del	chrX	144904019	144904020	AA	-	53	0	7	21
ODG1	<i>SNRPA1</i>	sySNV	SNRPA1:NM_003090: exon2:c.T135C:p.D45D	chr15	101833325	101833325	A	G	15	0	9	6
ODG11	<i>SPATS2L</i>	sySNV	SPATS2L:NM_001100424: exon12:c.G1113A: p.Q371Q	chr2	201342397	201342397	G	A	68	0	66	35
ODG9	<i>SPTA1</i>	nsSNV	SPTA1:NM_003126: exon12:c.C1490A: p.A497D	chr1	158641242	158641242	G	T	44	0	26	17
ODG1	<i>SSH2</i>	fsDel	SSH2:NM_033389: exon14:c.1576_1579del: p.526_527del	chr17	27963588	27963591	TTCT	-	105	0	52	27
ODG9	<i>ST6GAL2</i>	nsSNV	ST6GAL2:NM_001142351: exon2:c.C272T:p.A91V	chr2	107460162	107460162	G	A	48	0	31	29
ODG6	<i>STK39</i>	nsSNV	STK39:NM_013233:	chr2	168869200	168869200	C	A	60	0	32	31

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon16:c.G1442T:p.G481V									
ODG5	<i>STOM</i>	sySNV	STOM:NM_001270526: exon5:c.C414T:p.N138N	chr9	124111509	124111509	G	A	40	0	31	9
ODG3	<i>STX2</i>	nsSNV	STX2:NM_001980: exon4:c.G253A:p.A85T	chr12	131297529	131297529	C	T	22	0	19	9
ODG8	<i>STXBP3</i>	Splicing		chr1	109301211	109301214	GTA A	-	38	0	6	11
ODG9	<i>SYAP1</i>	nsSNV	SYAP1:NM_032796: exon9:c.A983G:p.E328G	chrX	16778406	16778406	A	G	57	0	16	15
ODG6	<i>TCHHL1</i>	sySNV	TCHHL1:NM_001008536: exon3:c.C1416A:p.T472T	chr1	152058742	152058742	G	T	139	0	95	32
ODG2	<i>TDRD5</i>	nsSNV	TDRD5:NM_001199085: exon2:c.C193T:p.R65C	chr1	179561943	179561943	C	T	63	0	78	21
ODG2	<i>TGFBR1</i>	nonfsDel	TGFBR1:NM_001130916: exon3:c.471_473del: p.157_158del				CTC	-	54	0	53	14
ODG4	<i>TIE1</i>	sySNV	TIE1:NM_001253357: exon13:c.C1896T:p.Y632Y	chr1	43778909	43778909	C	T	103	0	12	55
ODG7	<i>TLL1</i>	nsSNV	TLL1:NM_012464: exon13:c.G1604A:p.R535H	chr4	166976307	166976307	G	A	83	0	73	27
ODG10	<i>TLR9</i>	sySNV	TLR9:NM_017442: exon2:c.C1092T:p.V364V	chr3	52257240	52257240	G	A	176	0	88	51
ODG3	<i>TMEM2</i>	sySNV	TMEM2:NM_001135820: exon17:c.C2922T:p.V974V	chr9	74319594	74319594	G	A	87	0	36	27
ODG3	<i>TMF1</i>	sySNV	TMF1:NM_007114: exon13:c.G2718A:p.R906R	chr3	69077090	69077090	C	T	26	0	19	7
ODG6	<i>TNN</i>	fsDel	TNN:NM_022093:	chr1	175067610	175067610	G	-	61	0	36	30

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon9:c.1998delG;p.V666fs									
ODG6	<i>TNNI2</i>	sySNV	TNNI2:NM_001145841: exon5:c.C375G;p.A125A	chr11	1862359	1862359	C	G	39	0	43	11
ODG11	<i>TNXB</i>	nsSNV	TNXB:NM_019105: exon31:c.C10427T;p.S3476 F	chr6	32014125	32014125	G	A	127	2	138	55
ODG10	<i>TP53</i>	fsDel	TP53:NM_001126118: exon3:c.248_249del: p.83_83del	chr17	7579321	7579322	CA	-	16	0	8	5
ODG10	<i>TP53</i>	sgSNV	TP53:NM_001126115: exon5:c.A562T;p.K188X	chr17	7576888	7576888	T	A	56	0	19	13
ODG10	<i>TPD52L2</i>	nsSNV	TPD52L2:NM_001243891: exon1:c.A1G;p.M1V	chr20	62496719	62496719	A	G	88	0	22	9
ODG11	<i>TREX2</i>	nsSNV	TREX2:NM_080701: exon2:c.C326A;p.A109E	chrX	152710563	152710563	G	T	131	1	107	58
ODG10	<i>TRIM13</i>	fsDel	TRIM13:NM_213590: exon2:c.798delT;p.P266fs	chr13	50586874	50586874	T	-	56	0	24	12
ODG2	<i>TRIM4</i>	fsIns	TRIM4:NM_033017: exon1:c.310_311insG: p.P104fs	chr7	99516715	99516715	-	C	11	0	10	5
ODG9	<i>TTN</i>	sySNV	TTN:NM_003319: exon154:c.G54957A: p.A18319A	chr2	179428707	179428707	C	T	56	0	26	30
ODG4	<i>UBA7</i>	nsSNV	UBA7:NM_003335: exon22:c.G2747A: p.R916H	chr3	49843567	49843567	C	T	142	0	81	57
ODG5	<i>UBE4A</i>	sySNV	UBE4A:NM_004788:	chr11	118243386	118243386	G	A	24	0	23	9

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			exon6:c.G726A:p.E242E									
			UBL3:NM_007106: exon3:c.147_149del: p.49_50del	chr13	30346323	30346325	TCT	-	52	0	28	9
ODG10	<i>UBL3</i>	nonfsDel										
ODG7	<i>URB2</i>	sySNV	URB2:NM_014777: exon4:c.C1737T:p.L579L	chr1	229772097	229772097	C	T	51	0	31	21
ODG2	<i>USHBP1</i>	sySNV	USHBP1:NM_031941: exon9:c.G1407A:p.Q469Q	chr19	17367343	17367343	C	T	22	0	13	18
ODG3	<i>USP34</i>	fsDel	USP34:NM_014709: exon68:c.8222_8223del: p.2741_2741del	chr2	61441654	61441655	AT	-	76	0	39	27
ODG11	<i>VARS2</i>	nsSNV	VARS2:NM_001167733: exon12:c.C830T:p.A277V	chr6	30887950	30887950	C	T	92	0	95	66
ODG7	<i>VPS13B</i>	nsSNV	VPS13B:NM_017890: exon29:c.G438T:p.R1462L	chr8	100523417	100523417	G	T	117	0	78	48
ODG11	<i>WDR36</i>	nsSNV	WDR36:NM_139281: exon9:c.A1151G:p.N384S	chr5	110440472	110440472	A	G	28	0	37	12
ODG6	<i>WDR6</i>	sySNV	WDR6:NM_018031: exon2:c.G1404C:p.V468V	chr3	49050281	49050281	G	C	50	0	30	8
ODG1	<i>YEATS4</i>	nsSNV	YEATS4:NM_006530: exon7:c.A619T:p.T207S	chr12	69784031	69784031	A	T	55	0	35	20
ODG4	<i>ZBTB20</i>	nsSNV	ZBTB20:NM_001164342: exon4:c.G788T:p.S263I	chr3	114070137	114070137	C	A	95	0	47	41
ODG7	<i>ZBTB20</i>	nsSNV	ZBTB20:NM_001164342: exon5:c.G1951T:p.V651L	chr3	114058127	114058127	C	A	87	0	85	25
ODG8	<i>ZFHX3</i>	nonfsIns	ZFHX3:NM_001164766: exon9:c.7812_7813insTGG:	chr16	72821621	72821621	-	CC	11	0	14	10

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Sample	Gene	Exonic Function	AAChange	Chr	Start	End	Ref	Obs	R1	V1	R2	V2
			p.G2604delinsGG									
ODG8	ZFH3	sySNV	ZFH3:NM_001164766: exon9:c.T7815C:p.G2605G	chr16	72821618	72821618	A	G	12	0	15	7
ODG10	ZIC1	nsSNV	ZIC1:NM_003412: exon1:c.G911A:p.C304Y	chr3	147128810	147128810	G	A	192	1	175	98
ODG6	ZMIZ1	nsSNV	ZMIZ1:NM_020338: exon11:c.G859A:p.A287T	chr10	81052015	81052015	G	A	26	0	11	15
ODG3	ZMYM4	sySNV	ZMYM4:NM_005095: exon15:c.G2499A: p.G833G	chr1	35855611	35855611	G	A	53	0	10	15
ODG8	ZNF182	nsSNV	ZNF182:NM_001007088: exon6:c.C1012A:p.Q338K	chrX	47836417	47836417	G	T	117	0	60	45
ODG8	ZNF257	sySNV	ZNF257:NM_033468: exon4:c.T864C:p.D288D	chr19	22271416	22271416	T	C	177	0	125	56
ODG4	ZNF280A	nsSNV	ZNF280A:NM_080740: exon2:c.G1522A:p.V508I	chr22	22868433	22868433	C	T	51	0	40	22
ODG1	ZNF608	nsSNV	ZNF608:NM_020747: exon4:c.C2525G:p.S842C	chr5	123983552	123983552	G	C	159	0	110	69
ODG3	ZNF730	nsSNV	ZNF730:NM_001277403: exon4:c.C887T:p.S296F	chr19	23328733	23328733	C	T	11	0	4	6
ODG7	ZNF804B	nsSNV	ZNF804B:NM_181646: exon4:c.A1646C:p.Q549P	chr7	88963942	88963942	A	C	140	1	104	59

Appendix - V: TCGA Sample Information

List of 65 grade II and grade III gliomas with 1p/19q codeletion from the TCGA cohort used for the transcriptome analysis along with the indicated mutation/s in the indicated genes

Case_ID	Histologic diagnosis	Grade	Gender	Group	IDH1 / IDH2	CIC [NM_015125]	NF1 [NM_000267] / EGFR [NM_005228] / NRAS [NM_002524]
TCGA-HT-7695	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon19:c.G4412A; p.R1471Q	
TCGA-DU-8168	Oligodendroglioma	G3	F	CIC-Mutant	R132H	exon10:c.2673_2674del; p.891_892fsdel	
TCGA-HT-7608	Oligoastrocytoma	G2	M	CIC-Mutant	R132H	exon6:c.824_825del; p.275_275fsdel	NF1:exon28:c.C3719G; p.A1240G
TCGA-CS-5394	Astrocytoma	G3	M	CIC-Mutant	R132H	exon19:c.4358_4359insA; p.L1453fsins	
TCGA-HT-7694	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon8:c.1245_1246insC; p.R415fsins	
TCGA-DB-5274	Oligoastrocytoma	G3	F	CIC-Mutant	R132H	exon10:c.2673_2674del; p.891_892fsdel	
TCGA-E1-5319	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon14:c.3363_3364del; p.1121_1122fsdel	
TCGA-EZ-7264	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon19:c.T4450G; p.F1484V	
TCGA-HT-7471	Oligodendroglioma	G3	F	CIC-Mutant	R132H	exon10:c.1491delT; p.P497fsdel	
TCGA-HT-7616	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon10:c.C2470T; p.Q824X	
TCGA-HT-7677	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon20:c.G4535T; p.R1512L	
TCGA-HW-7487	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon5:c.G644A; p.R215Q	

Case_ID	Histologic diagnosis	Grade	Gender	Group	IDH1 / IDH2	CIC [NM_015125]	NF1 [NM_000267] / EGFR [NM_005228] / NRAS [NM_002524]
TCGA-FG-5962	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon20:c.C4543T; p.R1515C	
TCGA-DU-7018	Oligodendroglioma	G3	F	CIC-Mutant	R132H	exon3:c.432_433del; p.144_145fsdel	
TCGA-P5-A5F0	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon5:c.G714A; p.W238X exon10:c.2695delA; p.K899fsdel exon14:c.C3337T; p.Q1113X	
TCGA-DU-7302	Oligodendroglioma	G3	F	CIC-Mutant	R132H	exon17:c.4053_4054del; p.1351_1352fsdel	
TCGA-DU-6400	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon5:c.C607T;p.P203S exon8:c.1183delT; p.Y395fsdel	
TCGA-E1-5311	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon5:c.C604T;p.R202W	
TCGA-HT-7875	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon12:c.C2980T;p.Q994X exon14:c.3363_3364del; p.1121_1122fsdel	
TCGA-HT-8012	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon3:c.267_268del; p.89_90fsdel exon7:c.1107_1110del; p.369_370fsdel exon13:c.3111_3112del; p.1037_1038fsdel	
TCGA-IK-8125	Oligoastrocytoma	G3	M	CIC-Mutant	R132H	exon16:c.3802delA;p.K1268 fsdel	

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Case_ID	Histologic diagnosis	Grade	Gender	Group	IDH1 / IDH2	CIC [NM_015125]	NF1 [NM_000267] / EGFR [NM_005228] / NRAS [NM_002524]
TCGA-HT-7480	Oligodendroglioma	G2	M	CIC-Mutant	R132H	CIC(NM_015125:exon13:c.3009-1G>T)splicing	
TCGA-FG-7634	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon19:c.G4412A:p.R1471Q	
TCGA-HT-7468	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon5:c.C643T:p.R215W	
TCGA-DU-5870	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon5:c.C643T:p.R215W	
TCGA-CS-6668	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon10:c.2154_2155insG:p.p718fsins	
TCGA-DB-5278	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon5:c.T629G:p.F210C exon20:c.G4544A:p.R1515H	
TCGA-DB-A4XA	Oligoastrocytoma	G2	M	CIC-Mutant	R132H	exon5:c.708_710del:p.236_237nfsdel	
TCGA-DU-6393	Oligodendroglioma	G3	M	CIC-Mutant	R132H	exon12:c.2970_2971del:p.990_991fsdel	
TCGA-DU-7294	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon5:c.T701A:p.L234Q	
TCGA-FG-5964	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon16:c.4032_4033del:p.1344_1345fsdel	
TCGA-FG-7641	Oligodendroglioma	G2	M	CIC-Mutant	R132H	exon5:c.A614G:p.N205S exon5:c.C682T:p.R228W exon19:c.G4420C:p.V1474L	
TCGA-HT-7877	Oligodendroglioma	G2	F	CIC-Mutant	R132H	exon5:c.C601T:p.R201W	
TCGA-HT-A4DV	Oligodendroglioma	G3	F	CIC-Mutant	R132H	exon5:c.C758T:p.A253V	
TCGA-E1-5318	Oligodendroglioma	G2	F	CIC-Mutant	R172M	exon10:c.2321_2322insGCC CCCT :p.A774fsins	

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Case_ID	Histologic diagnosis	Grade	Gender	Group	IDH1 / IDH2	CIC [NM_015125]	NF1 [NM_000267] / EGFR [NM_005228] / NR4A1 [NM_002524]
TCGA-HT-A5R9	Oligodendroglioma	G3	F	CIC-Mutant	R172G	exon20:c.4551_4553del; p.L1517_1518nfsdel	
TCGA-HW-7491	Oligodendroglioma	G2	M	CIC-Mutant	R172K	exon5:c.C643T;p.R215W	
TCGA-DH-5141	Oligodendroglioma	G3	M	CIC-Mutant	R172K	exon15:c.3661_3662del; p.L221_1221fsdel	
TCGA-HT-7681	Oligoastrocytoma	G2	F	CIC-Mutant	R172K	exon5:c.C643T;p.R215W	
TCGA-FG-8186	Oligoastrocytoma	G3	F	CIC- wildType	R132H		
TCGA-HW-8322	Oligodendroglioma	G2	M	CIC- wildType	R132H		
TCGA-HW-A5KJ	Oligodendroglioma	G3	M	CIC- wildType	R132H		
TCGA-P5-A5ET	Oligodendroglioma	G2	M	CIC- wildType	R132H		
TCGA-DU-6410	Oligodendroglioma	G3	M	CIC- wildType	R132H		EGFR:exon7:c.C866T; p.A289V
TCGA-HT-7605	Oligodendroglioma	G2	M	CIC- wildType	R132H		
TCGA-FG-8187	Oligoastrocytoma	G2	M	CIC- wildType	R132H		
TCGA-HT-7692	Oligoastrocytoma	G2	M	CIC- wildType	R132H		
TCGA-HT-8010	Oligodendroglioma	G2	F	CIC- wildType	R132H		NF1:exon22:c.2983delC; p.L995fsdel
TCGA-DB-5279	Oligodendroglioma	G2	M	CIC- wildType	R132H		
TCGA-DU-7009	Oligodendroglioma	G2	F	CIC-	R132H		

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Case_ID	Histologic diagnosis	Grade	Gender	Group	IDH1 / IDH2	CIC [NM_015125]	NF1 [NM_000267] / EGFR [NM_005228] / NR4A1 [NM_002524]
				wildType			
TCGA-DB-A4XH	Oligoastrocytoma	G2	F	CIC-wildType	R132H		
TCGA-DU-5874	Oligodendroglioma	G2	F	CIC-wildType	R132H		
TCGA-HT-7687	Oligodendroglioma	G3	M	CIC-wildType	R132H		
TCGA-HT-8109	Oligodendroglioma	G3	M	CIC-wildType	R132H		
TCGA-HW-7486	Oligodendroglioma	G2	M	CIC-wildType	R132H		
TCGA-HW-7495	Oligodendroglioma	G2	F	CIC-wildType	R132H		
TCGA-DB-A4XG	Oligodendroglioma	G3	M	CIC-wildType	R172K		NF1:exon45:c.6787_6790fsdel;p.2263_2264del
TCGA-HT-7481	Oligodendroglioma	G2	M	CIC-wildType	R132H		
TCGA-DU-6394	Oligodendroglioma	G3	M	CIC-wildType	R132H		NF1:exon14:c.1541_1542del;p.514_514del
TCGA-CS-5390	Oligodendroglioma	G2	F	CIC-wildType	R132H		NR4A1:exon3:c.C181A;p.Q61K
TCGA-DU-8164	Oligodendroglioma	G2	M	CIC-wildType	R132H		
TCGA-HT-7620	Oligodendroglioma	G3	M	CIC-wildType	R132H		
TCGA-DU-6397	Oligodendroglioma	G3	M	CIC-wildType	R132H		NF1:exon40:c.5944-1G>C

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=non-synonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Case_ID	Histologic diagnosis	Grade	Gender	Group	IDH1 / IDH2	CIC [NM_015125]	NF1 [NM_000267] / EGFR [NM_005228] / NRAS [NM_002524]
TCGA-HT-7467	Oligodendroglioma	G2	M	CIC- wildType	R132H		
TCGA-DU-5849	Oligodendroglioma	G2	M	CIC- wildType	R172K		

AAChange=(Gene:Transcript:Exon:NucleotideChange:AminoAcidChange); nsSNV=nonsynonymous SNV; sySNV=synonymous SNV; fsDel=frameshift deletion; fsIns=frameshift insertion; sgSNV=stopgain SNV; R1=Reference-supporting reads in normal; R2=Reference-supporting reads in tumour; V1=Variant-supporting reads in normal; V2=Variant-supporting reads in tumour

Appendix - VI: Tables of Differentially Expressed Genes

Table 4.7: List of 106 genes significantly differentially expressed at 5% false discovery rate from the differential expression analysis between *CIC* mutated and *CIC* wild type samples from the study cohort using EdgeR.

Gene	log FC	log CPM	P Value	FDR
<i>RGS9</i>	4.009370014	5.838731771	5.45E-09	3.92E-05
<i>CERCAM</i>	5.530465343	10.77473686	5.63E-09	3.92E-05
<i>ATHL1</i>	2.986231339	6.323405376	2.58E-08	1.10E-04
<i>USH1C</i>	4.240192914	6.760915158	4.35E-08	1.10E-04
<i>BCAS1</i>	4.756071361	9.558639979	4.41E-08	1.10E-04
<i>CDK18</i>	2.863592827	7.041432863	4.74E-08	1.10E-04
<i>C2orf82</i>	4.966592074	6.359412241	6.78E-08	1.35E-04
<i>PHYHD1</i>	3.004737707	6.206692863	8.86E-08	1.38E-04
<i>SOX10</i>	4.184716651	6.653438045	8.93E-08	1.38E-04
<i>HLA-DQB1</i>	4.785607334	4.106854928	2.19E-07	3.04E-04
<i>CDH23</i>	3.131076321	3.672429203	4.97E-07	6.29E-04
<i>ABLIM3</i>	2.518070579	7.069704549	6.96E-07	8.08E-04
<i>ATAD3B</i>	2.659298056	5.024628055	2.13E-06	2.13E-03
<i>PTGDS</i>	3.42719371	9.978341837	2.14E-06	2.13E-03
<i>MIR6723</i>	7.550911765	7.833706758	2.93E-06	2.62E-03
<i>XIST</i>	-7.797826108	6.204889307	3.01E-06	2.62E-03
<i>RHBDL3</i>	2.448988017	7.702092199	3.29E-06	2.69E-03
<i>KIF6</i>	3.166175022	4.118962664	3.89E-06	2.89E-03
<i>SEMA3B</i>	2.362498212	7.852564017	3.95E-06	2.89E-03
<i>SLC14A1</i>	4.101245121	5.231225768	7.65E-06	5.23E-03
<i>TMC6</i>	2.830366096	6.173789595	8.06E-06	5.23E-03
<i>GPAT2</i>	3.339514069	4.410200578	8.51E-06	5.23E-03
<i>RPGR</i>	2.731870023	4.713253236	8.64E-06	5.23E-03
<i>IL32</i>	3.754292917	6.597186573	1.12E-05	6.50E-03
<i>DAPL1</i>	4.565316192	4.936356069	1.20E-05	6.67E-03
<i>CHIT1</i>	4.61638472	3.082131422	1.28E-05	6.84E-03
<i>RHPN1</i>	2.217945255	6.506626421	1.43E-05	7.37E-03
<i>VWA3A</i>	3.011022883	3.380891436	1.49E-05	7.39E-03
<i>CCDC159</i>	2.527481653	6.568782859	1.74E-05	8.33E-03
<i>FAHD2CP</i>	2.720497783	3.603913269	2.20E-05	1.02E-02
<i>GFAP</i>	2.510933668	14.51561341	2.28E-05	1.02E-02
<i>DLEC1</i>	2.839822244	2.505586521	2.40E-05	1.02E-02
<i>LINC01094</i>	2.971939282	4.965927236	2.41E-05	1.02E-02
<i>NKX6-2</i>	2.420498954	5.36542084	2.58E-05	1.06E-02
<i>CFAP70</i>	2.24928042	4.252489153	2.80E-05	1.12E-02
<i>HAUS7</i>	3.482640387	8.489813114	3.02E-05	1.17E-02
<i>FAM107A</i>	2.426767644	11.08237769	3.77E-05	1.42E-02
<i>NR6A1</i>	3.040379183	3.245448916	3.97E-05	1.46E-02
<i>GLIS1</i>	2.795252639	2.541745622	4.49E-05	1.60E-02

Appendix-VI: Tables of Differentially Expressed Genes

Gene	log FC	log CPM	P Value	FDR
<i>TENC1</i>	2.127179374	8.357171221	4.72E-05	1.64E-02
<i>MBP</i>	2.523126306	11.74372956	5.08E-05	1.72E-02
<i>LINC00844</i>	2.005523234	6.092010693	5.29E-05	1.72E-02
<i>STMN2</i>	-4.852142433	6.949511685	5.37E-05	1.72E-02
<i>TMPRSS5</i>	3.279139113	7.284848698	5.45E-05	1.72E-02
<i>ARMC12</i>	2.424333155	4.344234916	6.02E-05	1.84E-02
<i>TSHR</i>	3.864277741	3.75175813	6.38E-05	1.84E-02
<i>C1orf228</i>	3.080227215	4.976822954	6.39E-05	1.84E-02
<i>GPIHBP1</i>	2.407834328	4.611391933	6.49E-05	1.84E-02
<i>PGF</i>	2.884938591	6.213413398	6.82E-05	1.84E-02
<i>SIL1</i>	1.863086151	5.814933157	6.90E-05	1.84E-02
<i>IL17RC</i>	1.869391738	6.078394548	6.99E-05	1.84E-02
<i>MTIF</i>	2.868385495	4.891386904	7.06E-05	1.84E-02
<i>HHATL</i>	2.349951565	5.359405297	7.07E-05	1.84E-02
<i>PDGFRA</i>	-2.729123852	7.747769264	7.12E-05	1.84E-02
<i>IL18</i>	3.164369275	3.185452588	7.76E-05	1.94E-02
<i>SLC8B1</i>	2.409453975	4.680219395	7.81E-05	1.94E-02
<i>SIGLEC1</i>	2.679973041	3.570277526	8.59E-05	2.10E-02
<i>NNMT</i>	3.783832872	3.039729707	8.79E-05	2.11E-02
<i>HP</i>	3.257397078	3.100176463	9.25E-05	2.16E-02
<i>ESRRG</i>	-3.008696486	3.388927857	9.34E-05	2.16E-02
<i>VEGFA</i>	-4.130574787	7.027911301	9.48E-05	2.16E-02
<i>NPTX1</i>	-3.617718557	6.057319563	9.93E-05	2.23E-02
<i>RHBDF1</i>	2.589641832	5.21623785	1.05E-04	2.30E-02
<i>FLJ16779</i>	2.911282763	10.48031652	1.06E-04	2.30E-02
<i>ITGB4</i>	2.668622839	5.751025302	1.08E-04	2.30E-02
<i>CYP2D7</i>	2.93184853	4.530442388	1.09E-04	2.30E-02
<i>GADD45G</i>	2.667634812	8.493051758	1.16E-04	2.41E-02
<i>CAPN9</i>	2.558847594	2.959115533	1.27E-04	2.60E-02
<i>PCDHGA11</i>	-5.935428705	4.087522169	1.29E-04	2.61E-02
<i>S100A13</i>	1.815811265	5.606245836	1.36E-04	2.70E-02
<i>SDS</i>	2.526944302	4.802954142	1.39E-04	2.73E-02
<i>RYR1</i>	2.179058652	4.851728599	1.46E-04	2.83E-02
<i>KREMEN2</i>	2.209519094	4.10659598	1.52E-04	2.90E-02
<i>LINC01268</i>	3.067471592	5.35442122	1.58E-04	2.96E-02
<i>LINC00689</i>	3.624031864	5.164937401	1.59E-04	2.96E-02
<i>ENTPD2</i>	2.65660222	5.229371255	1.64E-04	3.00E-02
<i>CRACR2B</i>	2.490133684	3.736689591	1.68E-04	3.05E-02
<i>LOC389332</i>	3.711506927	2.946709481	1.71E-04	3.05E-02
<i>F5</i>	2.840176836	4.167018514	1.76E-04	3.07E-02
<i>TMEM189</i>	2.062103714	5.812004112	1.76E-04	3.07E-02
<i>SH3TC1</i>	3.032216692	4.890705517	1.89E-04	3.26E-02
<i>PTPRVP</i>	2.958886594	2.296315338	2.04E-04	3.46E-02

Appendix-VI: Tables of Differentially Expressed Genes

Gene	log FC	log CPM	P Value	FDR
<i>FADS3</i>	1.797645041	6.163121559	2.07E-04	3.47E-02
<i>TRAF1</i>	2.400384789	4.079697151	2.14E-04	3.54E-02
<i>CALN1</i>	2.476503642	6.052488905	2.17E-04	3.55E-02
<i>IQCK</i>	1.825332156	6.934190116	2.28E-04	3.68E-02
<i>COLCA2</i>	2.557176001	2.505935884	2.30E-04	3.68E-02
<i>SFTPC</i>	2.53301931	2.532793606	2.35E-04	3.72E-02
<i>ETV1</i>	-2.401391061	8.590480901	2.43E-04	3.81E-02
<i>NEAT1</i>	2.156986562	7.079883654	2.53E-04	3.87E-02
<i>MGP</i>	2.794449861	4.907811789	2.53E-04	3.87E-02
<i>SPRED2</i>	-1.993761919	5.664144357	2.71E-04	4.05E-02
<i>KCNQ3</i>	-2.979986627	4.72051889	2.71E-04	4.05E-02
<i>SRPX2</i>	2.754551327	2.233271764	2.77E-04	4.10E-02
<i>SPRED1</i>	-1.87231971	6.30691559	2.83E-04	4.14E-02
<i>IGFN1</i>	3.411478247	5.759231863	2.88E-04	4.18E-02
<i>PCDH20</i>	-3.57222355	4.331468934	2.96E-04	4.24E-02
<i>ABHD2</i>	-2.520348688	7.025061231	3.33E-04	4.74E-02
<i>THBS3</i>	1.86431969	6.733853368	3.46E-04	4.83E-02
<i>ACCS</i>	2.804068095	1.879419684	3.47E-04	4.83E-02
<i>TNFRSF14</i>	2.170150192	4.534497086	3.54E-04	4.88E-02
<i>HLA-DQA1</i>	3.849146803	2.780483425	3.64E-04	4.92E-02
<i>SDC4</i>	3.001914095	5.629844312	3.70E-04	4.92E-02
<i>SLC35F1</i>	-1.817183564	6.425793074	3.72E-04	4.92E-02
<i>MICALL2</i>	1.852420485	6.264122207	3.73E-04	4.92E-02
<i>KCNG1</i>	2.909618522	3.441693912	3.75E-04	4.92E-02

Table 4.8: List of 158 genes significantly differentially expressed at 5% false discovery rate from the differential expression analysis between *CIC* mutated (n=39) and *CIC* wild type (n=26) samples from TCGA cohort using EdgeR.

Gene	logFC	logCPM	P Value	FDR
<i>SPRY4</i>	-1.32875	4.914182	2.89E-09	4.55E-05
<i>MGC12982</i>	-1.80552	-0.00647	1.53E-08	8.86E-05
<i>LOC154822</i>	2.681276	6.807209	2.18E-08	8.86E-05
<i>SCARA5</i>	-2.3743	2.430202	2.75E-08	8.86E-05
<i>SLC26A7</i>	2.059799	-0.28358	3.23E-08	8.86E-05
<i>PCDH20</i>	-1.55091	4.730227	3.38E-08	8.86E-05
<i>CREB3L1</i>	-1.58863	4.026043	4.15E-08	9.15E-05
<i>ADAMTS16</i>	2.641477	1.663475	4.66E-08	9.15E-05
<i>METTL7B</i>	2.289396	3.750983	5.84E-08	1.00E-04
<i>PLN</i>	2.755691	3.411297	6.39E-08	1.00E-04
<i>ACTC1</i>	2.82838	2.203654	7.07E-08	1.01E-04
<i>ZBTB8B</i>	-1.40926	2.053342	8.70E-08	1.14E-04
<i>IGJ</i>	-4.69221	2.578908	1.67E-07	2.00E-04
<i>SLC35F1</i>	-0.82595	7.013459	1.93E-07	2.00E-04
<i>NPPA</i>	-2.07933	4.767187	1.97E-07	2.00E-04
<i>ETV5</i>	-1.38861	7.531032	2.04E-07	2.00E-04
<i>EPN2</i>	-0.73487	8.509935	2.34E-07	2.17E-04
<i>ETV1</i>	-1.34139	8.895254	2.51E-07	2.19E-04
<i>GSG1L</i>	1.314638	4.502424	3.40E-07	2.81E-04
<i>ADAM6</i>	-4.6283	5.460899	4.16E-07	3.17E-04
<i>NKD1</i>	0.99413	5.185948	4.23E-07	3.17E-04
<i>MSTN</i>	1.684685	3.427894	6.55E-07	4.69E-04
<i>TMEM158</i>	-1.2519	4.549119	9.79E-07	6.70E-04
<i>DUSP6</i>	-0.99789	5.884152	1.41E-06	9.16E-04
<i>GCNT2</i>	-0.96594	4.560459	1.46E-06	9.16E-04
<i>TMEM139</i>	1.110238	0.619905	2.09E-06	1.26E-03
<i>ARL4A</i>	-0.98174	5.758728	2.25E-06	1.31E-03
<i>KIF26B</i>	-1.26209	4.344027	2.45E-06	1.33E-03
<i>ZSWIM6</i>	0.602392	5.719098	2.51E-06	1.33E-03
<i>SPRED2</i>	-0.75923	6.223939	2.53E-06	1.33E-03
<i>WSCD1</i>	-0.85494	8.018907	3.93E-06	1.99E-03
<i>DUSP4</i>	-1.50918	3.071095	4.05E-06	1.99E-03
<i>DUSP13</i>	2.011129	-1.5979	4.99E-06	2.38E-03
<i>ELFN2</i>	-0.88146	7.182117	5.32E-06	2.46E-03
<i>SLC29A1</i>	-0.55751	4.856609	7.57E-06	3.37E-03
<i>DES</i>	-2.76716	0.785065	7.72E-06	3.37E-03
<i>DLL3</i>	-1.07685	8.154101	8.39E-06	3.54E-03
<i>ETV4</i>	-2.02452	4.939026	8.54E-06	3.54E-03
<i>BMPER</i>	-1.06923	4.130735	9.46E-06	3.82E-03
<i>NXPH3</i>	-0.75088	5.198053	1.00E-05	3.95E-03
<i>FGD3</i>	-0.94988	3.425837	1.23E-05	4.72E-03

Appendix-VI: Tables of Differentially Expressed Genes

Gene	logFC	logCPM	P Value	FDR
<i>CRYBG3</i>	0.869692	3.836604	1.37E-05	4.88E-03
<i>CHRM5</i>	1.961471	0.360373	1.38E-05	4.88E-03
<i>PDZD2</i>	-0.82582	7.302615	1.42E-05	4.88E-03
<i>FOXD2</i>	-1.33892	-1.27896	1.44E-05	4.88E-03
<i>TMC3</i>	-2.22314	-0.07828	1.46E-05	4.88E-03
<i>ELFN1</i>	-0.93264	3.645059	1.46E-05	4.88E-03
<i>UHRF1</i>	-0.69629	5.137706	1.54E-05	5.05E-03
<i>FGFR1</i>	-0.53868	6.777867	1.79E-05	5.74E-03
<i>SCEL</i>	-1.40677	-1.11407	1.84E-05	5.77E-03
<i>CHRNA4</i>	-0.73769	5.159052	2.16E-05	6.67E-03
<i>PDE4B</i>	-0.69382	7.613269	2.21E-05	6.67E-03
<i>EHF</i>	1.448794	-0.91227	2.25E-05	6.67E-03
<i>STAMBPL1</i>	-0.50578	3.557292	2.29E-05	6.67E-03
<i>KEL</i>	1.263575	-0.70521	2.47E-05	7.02E-03
<i>EN2</i>	1.29121	-0.9489	2.50E-05	7.02E-03
<i>CMKLR1</i>	-1.09527	5.021605	2.76E-05	7.63E-03
<i>COL1A1</i>	-1.77287	4.15372	2.99E-05	8.10E-03
<i>KCNIP1</i>	-0.73156	6.217127	3.69E-05	9.84E-03
<i>PTX3</i>	-0.98965	0.662583	4.19E-05	1.10E-02
<i>SPSB4</i>	-0.74995	3.991896	4.38E-05	1.11E-02
<i>AIF1L</i>	0.608886	7.718841	4.39E-05	1.11E-02
<i>SHOX2</i>	-2.67852	-1.5889	4.56E-05	1.14E-02
<i>KCNIP3</i>	-0.58597	7.180679	4.97E-05	1.22E-02
<i>TNFSF13B</i>	-1.74405	1.725282	5.20E-05	1.26E-02
<i>TACC2</i>	-0.52707	5.917189	5.44E-05	1.30E-02
<i>PAQR4</i>	0.559702	6.70232	5.60E-05	1.31E-02
<i>DUOX2</i>	-2.54854	0.323641	5.65E-05	1.31E-02
<i>FGFBP3</i>	-0.89374	5.682398	5.89E-05	1.34E-02
<i>KCTD4</i>	0.802654	2.915325	6.02E-05	1.35E-02
<i>COL4A4</i>	-1.3485	3.721539	6.42E-05	1.42E-02
<i>NXPH1</i>	-0.98491	6.418292	6.65E-05	1.45E-02
<i>GLDN</i>	1.252828	4.945355	7.03E-05	1.52E-02
<i>NRG1</i>	-1.11699	2.005897	7.24E-05	1.54E-02
<i>NLGN3</i>	-0.45739	8.252162	7.89E-05	1.65E-02
<i>C8orf56</i>	-0.98824	1.864533	8.27E-05	1.68E-02
<i>CXorf64</i>	1.797504	-1.04312	8.28E-05	1.68E-02
<i>ACTG2</i>	-1.93927	0.159896	8.33E-05	1.68E-02
<i>CPT1A</i>	-0.63	6.273725	8.52E-05	1.70E-02
<i>STARD5</i>	-0.75847	0.192114	8.99E-05	1.74E-02
<i>ALK</i>	-1.00131	3.545734	9.08E-05	1.74E-02
<i>SHC3</i>	-1.1737	5.430138	9.13E-05	1.74E-02
<i>GFRA1</i>	-0.73319	7.198355	9.18E-05	1.74E-02
<i>PDE1C</i>	1.370729	2.980975	9.38E-05	1.74E-02

Appendix-VI: Tables of Differentially Expressed Genes

Gene	logFC	logCPM	P Value	FDR
<i>LOC92659</i>	-0.55838	2.50875	9.41E-05	1.74E-02
<i>SGK3</i>	0.620884	4.210901	9.88E-05	1.81E-02
<i>TP53INP2</i>	0.817993	7.522343	1.09E-04	1.96E-02
<i>C21orf62</i>	1.126386	3.744125	1.10E-04	1.96E-02
<i>DNAH2</i>	-1.54703	0.289859	1.13E-04	2.00E-02
<i>PRKG2</i>	-1.55919	1.46445	1.20E-04	2.09E-02
<i>TMTC2</i>	0.987499	4.488665	1.31E-04	2.25E-02
<i>H19</i>	-1.73372	0.813574	1.31E-04	2.25E-02
<i>TMOD1</i>	-0.54107	6.702946	1.35E-04	2.29E-02
<i>WNT2</i>	1.825554	0.009643	1.45E-04	2.43E-02
<i>LPPR5</i>	-0.57424	4.316456	1.52E-04	2.52E-02
<i>C6orf118</i>	-1.0304	-0.6436	1.58E-04	2.60E-02
<i>TF</i>	1.102519	9.245702	1.63E-04	2.65E-02
<i>ISM1</i>	-0.87954	2.674393	1.65E-04	2.65E-02
<i>PLEKHH1</i>	0.839233	7.007612	1.74E-04	2.76E-02
<i>THSD4</i>	-0.90171	5.305801	1.77E-04	2.79E-02
<i>RHBDL2</i>	1.345035	1.463947	1.84E-04	2.87E-02
<i>SPRED1</i>	-0.55917	6.845477	1.87E-04	2.89E-02
<i>FLRT3</i>	-0.76302	3.760729	1.90E-04	2.90E-02
<i>ADAMTSL3</i>	0.894595	1.352462	2.00E-04	3.03E-02
<i>LOC283392</i>	-0.96175	0.218699	2.05E-04	3.08E-02
<i>FGFRL1</i>	0.789989	3.884099	2.13E-04	3.10E-02
<i>FAM84B</i>	0.690918	7.734316	2.15E-04	3.10E-02
<i>ADAMTS9</i>	0.860277	4.905188	2.15E-04	3.10E-02
<i>TRIB2</i>	-0.60114	7.154406	2.16E-04	3.10E-02
<i>GCOM1</i>	0.964945	0.370461	2.17E-04	3.10E-02
<i>CACNG1</i>	-0.9492	0.474512	2.26E-04	3.20E-02
<i>THBS1</i>	-1.21894	3.354929	2.32E-04	3.25E-02
<i>C13orf31</i>	1.104671	3.061934	2.34E-04	3.25E-02
<i>DACH1</i>	-1.05032	3.253596	2.38E-04	3.25E-02
<i>GPR123</i>	-0.48457	6.327445	2.40E-04	3.25E-02
<i>KCNK3</i>	-0.72093	5.709383	2.41E-04	3.25E-02
<i>COL3A1</i>	-1.62596	4.514925	2.42E-04	3.25E-02
<i>HN1L</i>	0.581753	4.785674	2.47E-04	3.28E-02
<i>HMP19</i>	-0.55352	8.693961	2.48E-04	3.28E-02
<i>CDH3</i>	-0.79751	2.590028	2.57E-04	3.35E-02
<i>COL4A3</i>	-1.31947	2.478813	2.60E-04	3.35E-02
<i>ST3GAL5</i>	-0.50325	6.865255	2.62E-04	3.35E-02
<i>C21orf125</i>	-1.12837	0.390559	2.62E-04	3.35E-02
<i>TRAK2</i>	0.425279	6.458399	2.65E-04	3.36E-02
<i>HSF2BP</i>	-0.72323	2.87476	2.73E-04	3.43E-02
<i>FOXP4</i>	-0.54363	5.316299	2.75E-04	3.44E-02
<i>RHOU</i>	0.606368	7.380455	2.80E-04	3.45E-02

Appendix-VI: Tables of Differentially Expressed Genes

Gene	logFC	logCPM	P Value	FDR
<i>PLP1</i>	1.054026	11.72707	2.81E-04	3.45E-02
<i>INHBB</i>	-0.70091	3.905635	2.95E-04	3.59E-02
<i>DIAPH2</i>	-0.65383	3.643558	2.97E-04	3.59E-02
<i>SH3TC2</i>	1.129603	4.011236	3.02E-04	3.62E-02
<i>IGFBP4</i>	-0.78379	5.56667	3.06E-04	3.62E-02
<i>SEMA5B</i>	0.686184	5.152634	3.06E-04	3.62E-02
<i>ZSWIM4</i>	-0.55781	4.61568	3.15E-04	3.70E-02
<i>ALMS1P</i>	-0.84524	-0.76692	3.20E-04	3.70E-02
<i>TMEM132C</i>	-0.62663	4.213564	3.21E-04	3.70E-02
<i>CPVL</i>	-0.75213	4.25295	3.22E-04	3.70E-02
<i>PRCP</i>	0.537973	7.47439	3.41E-04	3.88E-02
<i>TRAF4</i>	-0.53655	6.259476	3.43E-04	3.88E-02
<i>HS6ST2</i>	-0.68405	4.340794	3.47E-04	3.90E-02
<i>EVI2A</i>	1.074793	5.714988	3.56E-04	3.97E-02
<i>NRBP2</i>	0.580091	6.447105	3.64E-04	4.01E-02
<i>ICOSLG</i>	1.043803	3.796001	3.69E-04	4.01E-02
<i>TMTC4</i>	0.609854	5.524642	3.71E-04	4.01E-02
<i>ACTA2</i>	-0.86046	5.18325	3.72E-04	4.01E-02
<i>PLEKHG4</i>	0.935971	0.985213	3.72E-04	4.01E-02
<i>LRRC8D</i>	0.423313	5.544459	3.93E-04	4.20E-02
<i>TBC1D10A</i>	-0.68073	5.770416	4.08E-04	4.30E-02
<i>PLXNB3</i>	0.46399	7.431745	4.10E-04	4.30E-02
<i>SYNJ2</i>	1.061992	5.431599	4.10E-04	4.30E-02
<i>SEMA4D</i>	0.898789	6.449906	4.13E-04	4.31E-02
<i>LECT1</i>	1.486456	-0.91221	4.16E-04	4.31E-02
<i>DDX11L2</i>	1.205605	0.096721	4.54E-04	4.66E-02
<i>PAIP2B</i>	0.775665	5.73611	4.65E-04	4.75E-02
<i>HTRA3</i>	-1.23314	0.210424	4.75E-04	4.82E-02
<i>RMST</i>	1.271312	-0.5322	4.87E-04	4.88E-02
<i>ANKRD55</i>	-1.16893	1.914486	4.87E-04	4.88E-02
<i>CDKN2B</i>	1.020296	4.327832	5.01E-04	4.99E-02

Table 4.9: List of 76 genes significantly differentially expressed at 10% false discovery rate from the differential expression analysis between *CIC* mutated and *CIC* wild type samples from TCGA cohort using DEseq.

Id	Base MeanA	Base MeanB	Fold Change	log2 Fold Change	Pval	Padj
<i>CPA3</i>	9.09	118.63	13.05	3.71	8.47E-40	1.68E-35
<i>TPSAB1</i>	16.24	186.56	11.49	3.52	1.56E-30	1.54E-26
<i>MS4A2</i>	2.54	25.91	10.22	3.35	7.06E-17	4.66E-13
<i>HOXA2</i>	0.33	10.72	32.26	5.01	3.82E-13	1.90E-09
<i>SILV</i>	9.43	50.46	5.35	2.42	5.76E-13	2.28E-09
<i>SLC26A7</i>	17.81	71.32	4.00	2.00	1.57E-10	5.19E-07
<i>DIRAS3</i>	115.67	314.53	2.72	1.44	2.90E-10	8.22E-07
<i>TMEM158</i>	1571.04	648.83	0.41	-1.28	3.86E-09	9.58E-06
<i>ABCC3</i>	85.65	387.49	4.52	2.18	1.32E-08	2.90E-05
<i>CREB3L1</i>	1176.49	376.18	0.32	-1.65	3.13E-08	6.20E-05
<i>ETV1</i>	31972.40	12172.59	0.38	-1.39	1.43E-07	2.58E-04
<i>KCNG1</i>	162.65	801.51	4.93	2.30	1.69E-07	2.72E-04
<i>HOXA4</i>	0.77	9.40	12.24	3.61	1.78E-07	2.72E-04
<i>SPRY4</i>	2022.73	911.11	0.45	-1.15	6.16E-07	8.73E-04
<i>ELFN2</i>	9197.60	4772.01	0.52	-0.95	7.26E-07	9.35E-04
<i>HSPA7</i>	29.64	71.43	2.41	1.27	7.55E-07	9.35E-04
<i>ETV5</i>	12489.34	5360.50	0.43	-1.22	1.11E-06	1.29E-03
<i>GCNT2</i>	1498.19	774.78	0.52	-0.95	1.26E-06	1.38E-03
<i>HOXA3</i>	0.86	8.85	10.27	3.36	1.82E-06	1.89E-03
<i>ZBTB8B</i>	278.19	103.61	0.37	-1.42	2.25E-06	2.13E-03
<i>C6</i>	1.77	11.06	6.24	2.64	2.34E-06	2.13E-03
<i>FGFRL1</i>	592.47	1358.14	2.29	1.20	2.36E-06	2.13E-03
<i>KLRC1</i>	23.06	6.93	0.30	-1.73	2.65E-06	2.28E-03
<i>DUSP4</i>	573.72	210.41	0.37	-1.45	3.70E-06	3.06E-03
<i>CCL7</i>	0.50	5.68	11.44	3.52	4.24E-06	3.36E-03
<i>SLN</i>	51.03	144.36	2.83	1.50	4.78E-06	3.65E-03
<i>CT45A2</i>	7.16	1.10	0.15	-2.70	6.03E-06	4.43E-03
<i>COL23A1</i>	160.68	502.57	3.13	1.65	1.20E-05	8.41E-03
<i>SCEL</i>	28.73	10.05	0.35	-1.52	1.23E-05	8.41E-03
<i>GFPT2</i>	1030.82	1939.54	1.88	0.91	1.72E-05	1.14E-02
<i>SLC35F1</i>	8113.32	4598.76	0.57	-0.82	1.87E-05	1.20E-02
<i>C8orf56</i>	235.92	114.46	0.49	-1.04	2.73E-05	1.69E-02
<i>CHRNA4</i>	2221.77	1273.89	0.57	-0.80	2.91E-05	1.75E-02
<i>FLRT3</i>	815.54	454.77	0.56	-0.84	3.91E-05	2.28E-02
<i>LRRN4CL</i>	87.45	287.98	3.29	1.72	4.32E-05	2.45E-02
<i>BMPER</i>	1112.67	518.40	0.47	-1.10	4.89E-05	2.61E-02
<i>PCDH20</i>	1806.82	593.43	0.33	-1.61	4.98E-05	2.61E-02
<i>PLN</i>	162.12	1072.97	6.62	2.73	5.01E-05	2.61E-02
<i>BFSP2</i>	10.81	2.05	0.19	-2.40	5.14E-05	2.61E-02

Appendix-VI: Tables of Differentially Expressed Genes

Id	Base MeanA	Base MeanB	Fold Change	log2 Fold Change	Pval	Padj
<i>CACNG1</i>	88.79	42.53	0.48	-1.06	5.38E-05	2.67E-02
<i>WSCD1</i>	16245.93	8809.11	0.54	-0.88	5.71E-05	2.76E-02
<i>EPN2</i>	22342.28	13521.84	0.61	-0.72	6.60E-05	3.11E-02
<i>CD274</i>	25.95	56.36	2.17	1.12	7.18E-05	3.14E-02
<i>TYR</i>	0.24	9.11	38.35	5.26	7.20E-05	3.14E-02
<i>GFRA1</i>	8837.07	5203.30	0.59	-0.76	7.36E-05	3.14E-02
<i>SPSB4</i>	970.91	566.95	0.58	-0.78	7.51E-05	3.14E-02
<i>ARL4A</i>	3543.55	1650.19	0.47	-1.10	7.62E-05	3.14E-02
<i>KCNIP1</i>	4637.73	2599.51	0.56	-0.84	7.74E-05	3.14E-02
<i>UHRF1</i>	2118.68	1265.99	0.60	-0.74	7.76E-05	3.14E-02
<i>C6orf118</i>	38.83	19.01	0.49	-1.03	8.03E-05	3.18E-02
<i>ETV4</i>	2224.20	622.72	0.28	-1.84	8.65E-05	3.36E-02
<i>PDZD2</i>	9733.68	5499.84	0.57	-0.82	8.99E-05	3.43E-02
<i>FGFBP3</i>	3165.78	1653.53	0.52	-0.94	9.39E-05	3.47E-02
<i>ZFY</i>	329.73	554.76	1.68	0.75	9.97E-05	3.59E-02
<i>CHI3L1</i>	1515.11	4853.82	3.20	1.68	1.09E-04	3.86E-02
<i>SHC3</i>	2828.70	1229.63	0.43	-1.20	1.14E-04	3.95E-02
<i>COL16A1</i>	2058.11	3299.74	1.60	0.68	1.16E-04	3.97E-02
<i>KDM5D</i>	1089.65	1833.98	1.68	0.75	1.41E-04	4.75E-02
<i>FGD3</i>	709.81	364.15	0.51	-0.96	1.45E-04	4.78E-02
<i>TMEM132C</i>	1109.80	671.66	0.61	-0.72	1.64E-04	5.32E-02
<i>NKD1</i>	1329.07	2632.49	1.98	0.99	1.76E-04	5.63E-02
<i>NXPH1</i>	5397.20	2562.81	0.47	-1.07	1.86E-04	5.87E-02
<i>SPRED2</i>	4573.96	2797.18	0.61	-0.71	1.92E-04	5.90E-02
<i>LOC283392</i>	70.64	37.55	0.53	-0.91	1.93E-04	5.90E-02
<i>CYorf15A</i>	210.87	362.79	1.72	0.78	2.01E-04	5.99E-02
<i>ELFN1</i>	831.76	412.97	0.50	-1.01	2.02E-04	5.99E-02
<i>DUSP6</i>	3765.73	2154.20	0.57	-0.81	2.08E-04	6.07E-02
<i>TMEM132D</i>	933.32	459.42	0.49	-1.02	2.14E-04	6.16E-02
<i>NXPH3</i>	2242.87	1377.42	0.61	-0.70	2.41E-04	6.83E-02
<i>ALK</i>	753.35	382.54	0.51	-0.98	2.54E-04	6.99E-02
<i>KRT15</i>	1.41	8.38	5.92	2.57	2.54E-04	6.99E-02
<i>GALNT5</i>	2.22	10.87	4.91	2.29	3.33E-04	9.06E-02
<i>CRYBG3</i>	544.86	1056.83	1.94	0.96	3.41E-04	9.15E-02
<i>FAM81B</i>	3.64	13.40	3.68	1.88	3.47E-04	9.19E-02
<i>HS6ST2</i>	1228.43	763.51	0.62	-0.69	3.65E-04	9.46E-02
<i>PCOLCE2</i>	332.85	193.95	0.58	-0.78	3.67E-04	9.46E-02

PUBLICATION

ETV/Pea3 Family Transcription Factor-Encoding Genes are Overexpressed in *CIC*-Mutant Oligodendrogliomas

Vijay Padul,¹ Sridhar Epari,² Aliasgar Moiyadi,³ Prakash Shetty,³ and Neelam Vishwanath Shirsat^{1,*}

¹Shirsat Laboratory, Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, India

²Department of Pathology, Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, India

³Neurosurgery Services, Department of Surgical Oncology, Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai, India

Oligodendrogliomas with combined loss of chromosome arms 1p and 19q are known to be particularly sensitive to chemotherapy, and the *CIC* gene located on 19q is known to be mutated in over 50% of the 1p/19q codeleted oligodendrogliomas. However, the role of *CIC* in the oligodendroglioma pathogenesis is not known. Exome sequencing of 11 oligodendroglial tumors identified 9 tumors with combined loss of 1p and 19q. Somatic mutations were found in the *CIC* and *FUBP1* genes. Recurrent somatic mutations were also identified in the Notch signaling pathway genes *NOTCH1* and *MAML3*, the chromatin modifying gene *ARID1A* and in *KRAS*. Comparison of the transcriptome profiles of *CIC*-mutant and *CIC*-wild type oligodendrogliomas from the study cohort as well as 65 1p/19q codeleted oligodendrogliomas from the TCGA cohort identified genes encoding the ETV transcription factor family to be significantly upregulated in the *CIC*-mutant tumors. Upregulation of a number of negative regulators of the receptor tyrosine kinase signaling pathway like Sprouty and SPRED family members in the *CIC*-mutant oligodendrogliomas is likely due to the constitutive activation of the pathway resulting from inactive *CIC* protein. Higher expression of the oncogenic ETV transcription factors in the *CIC*-mutant oligodendrogliomas may make these tumors more aggressive than the *CIC*-wild type tumors. © 2015 Wiley Periodicals, Inc.

INTRODUCTION

Brain tumors are the second leading cause of cancer-related deaths. Gliomas, which account for almost 80% of the primary malignant brain tumors in adults (Ostrom et al., 2013), are classified as astrocytic, oligodendroglial, or ependymal depending on the resemblance of the tumor cell morphology to a specific glial cell type (Collins, 2004). Furthermore, each glioma type is graded from grade I to III/IV based on the histological grade of malignancy.

Glioblastoma (GBM), the grade IV glioma, accounts for ~45% of all malignant brain tumors in adults and has been studied most extensively of all the brain tumors. Astrocytomas and oligodendrogliomas account for most of the low grade (grade II and III) gliomas in adults. Oligodendrogliomas generally have slower growth rates and have better prognosis than astrocytomas of similar grade (Bromberg and van den Bent, 2009). Recurrent mutations in the *IDH1* or *IDH2* gene have been reported to occur at high frequency

(70–90%) in astrocytomas, oligodendrogliomas, and secondary GBMs, while rare or absent in primary GBMs (Ichimura et al., 2009). Oligodendrogliomas with combined loss of chromosome arms 1p and 19q are known to be particularly sensitive to chemotherapy with much longer progression free survival (Jenkins et al., 2006). Bettgowda et al. (2011) first reported mutations in the *CIC* gene and *FUBP1* gene located on 19q and 1p

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Dr. Neelam Vishwanath Shirsat, Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai 410210, India. E-mail: nshirsat@actrec.gov.in

Supported by: The Department of Biotechnology, and a Research Fellowship from the Council of Scientific and Industrial Research, New Delhi, India

Received 20 February 2015; Revised 11 June 2015; Accepted 15 June 2015

DOI 10.1002/gcc.22283

Published online 10 September 2015 in Wiley Online Library (wileyonlinelibrary.com).

respectively, in oligodendrogliomas having 1p/19q codeletion.

Integrated genomic analysis of a large number of adult brain low grade glioma (LGG) tissues is being carried out by the Cancer Genome Atlas (TCGA) consortium (www.cancergenome.nih.gov). In the present study, we performed exome sequencing of 11 tumors with classic oligodendroglial morphology and their paired blood samples. We compared transcriptome profiles of the *CIC*-mutant with those of the *CIC*-wild type, 1p/19q co-deleted oligodendrogliomas from our cohort as well as 65 tumors from the TCGA cohort in order to understand role of the *CIC* gene in oligodendroglioma pathogenesis.

MATERIALS AND METHODS

Tumor Tissues and Paired Blood Samples

The study was approved by the Tata Memorial Centre's Institutional Ethics Committee III. Oligodendroglioma tumor tissues and paired blood samples were obtained after acquiring informed consent from the patients. The tumor tissues were snap frozen in liquid nitrogen immediately after the surgical resection and stored at -70°C . The histopathological diagnosis and grading of the tumor tissues was done as per the WHO 2007 classification of tumors of the Central Nervous System (Louis et al., 2007) and only tumors diagnosed as oligodendrogliomas were included in the study.

DNA and RNA Isolation

Genomic DNA and RNA were extracted from the tumor tissues after ensuring at least 80% tumor cell content. Genomic DNA was extracted from the tumor tissues and paired whole blood using QIAamp DNA mini Kit as per the manufacturer's protocol (Qiagen, GmbH, Hilden, Germany). Total RNA was isolated from the tumor tissues using RNeasy plus mini kit as per the manufacturer's protocol (Qiagen). The RNA and DNA was quantified in a Qubit 2.0 fluorimeter (Life Technologies, Carlsbad, CA) and the quality was checked by agarose gel electrophoresis.

Mutation Detection by Sanger Sequencing

The Mutational status of the R132 and R172 amino acid codons of the *IDH1* and *IDH2* genes respectively, as well as mutations identified in the *CIC* gene by exome sequencing were validated by Sanger sequencing using the ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA).

Deep Sequencing of Tumor and Paired Blood Exome and Tumor RNA

DNA libraries for the paired end multiplex deep sequencing were prepared from the genomic DNA isolated from the tumor tissues and their paired blood using the DNA library preparation kit for Illumina from Kapa Biosystems (Wilmington, MA) as per the manufacturer's protocol. Two microgram genomic DNA was sheared using a Covaris M220 focused ultrasonicator (Covaris, Woburn, MA), end repaired, and ligated to the single indexed DNA adapters, followed by the separation of the fragmented DNA by agarose gel electrophoresis and purification of ~ 300 bp DNA fragments using QIAquick gel extraction kit (Qiagen). For the targeted exome sequencing, exome capture was performed using the Truseq Exome enrichment kit (Catalogue No.-FC-121-1008, Illumina, San Diego, CA) that captures 62 Mb exomic region corresponding to the 20,794 genes. Single indexed RNA libraries were prepared using the Truseq RNA sample prep kit V2 (Catalogue No. RS-122-2001, Illumina) using 4 μg of total RNA as per the manufacturer's protocol. The multiplexed exome and RNA libraries were subjected to 100 and 150 bp paired end deep sequencing, respectively, using the HiSeq 1500 ultra-high-throughput sequencing system (Illumina). The exome sequencing was done to get at least 50 X average depth of coverage, while the RNA sequencing was done to obtain a minimum of 20 million reads per sample.

Bioinformatic Analyses

The sequencing data generated as BCL basecall files were de-multiplexed using the Illumina Bcl2FastQ version 1.8.4. FastQC (bioinformatics.babraham.ac.uk/projects/fastqc/) was used for quality analysis of raw FASTQ reads.

The FASTQ reads of the exome sequence data were aligned to the human reference genome hg19 using the Burrows-Wheeler Aligner version 0.7.9 (www.bio-bwa.sourceforge.net) with default parameters. Duplicate reads were removed using the Picard Tools version 1.80 (<http://broadinstitute.github.io/picard>). The alignment files were refined by local realignment of the reads and base quality recalibration by the Genome Analysis Toolkit (GATK) version 2.1.3 (<https://www.broadinstitute.org/gatk>). Exome enrichment analysis of the binary reads alignment (BAM) files was done using the NGSrich Version 0.7.8 (<http://sourceforge.net/projects/ngsrich>). Somatic single nucleotide variants (SNVs) and insertions and deletions

(indels) were identified using the VarScan variant detection tool version 2.3.5 (<http://varscan.sourceforge.net>) using the filtering criteria of a minimum coverage 10 and at least 5 somatic variants. Functional annotation of the somatic variant list was done using the ANNOVAR software (www.open-bioinformatics.org/annovar) (Wang et al., 2010). From the ANNOVAR-annotated list, variants located in segmental duplications were excluded. The remaining variants were manually verified in IGV (www.broadinstitute.org/igv). Ambiguous variants (variants represented in reads with low mapping quality, variants present near indels, and variants surrounded by mismatched bases) were discarded. The copy number variations in the tumor genome were identified from the paired exome sequence data using the FishingCNV software version 1.5.2 (<http://sourceforge.net/projects/fishingcnv>). Segmentation means of less than -0.3 and more than 0.3 were considered as deletion and amplification, respectively. The copy number variations in the tumor genome were also analyzed using the Control-FREEC software (<http://bioinfo-out.curie.fr/projects/freec/>), which uses input aligned reads in samtools mpileup format to construct and normalize the copy number profile and the B-allele frequency (BAF) profile. By performing segmentation of both profiles, it ascribes the genotype status and annotates genomic alterations using both copy number and allelic frequency information.

The reads of the RNA sequencing data containing adapter overlaps were cleaned using the reads trimming tool Trimmomatic version 0.32 (<http://www.usadellab.org>). The cleaned reads were aligned to the reference human genome hg19 using TopHat version 2.0.13 (<http://ccb.jhu.edu/software/tophat>) with default parameters. Raw counts for the reads aligned to the gene intervals were produced by the python package HTSeq version 0.6.1 (www-huber.embl.de/users/anders/HTSeq) using the default union-counting mode. The read count based gene level differential expression analysis comparing the transcriptome profiles of the *CIC*-mutant and *CIC*-wild type oligodendrogliomas was carried out using the EdgeR package of R bioconductor (www.bioconductor.org).

Analysis of the TCGA Data on Brain Low Grade Gliomas

A total of 65 *IDH1/IDH2*-mutant, 1p/19q codeleted oligodendroglioma tumors for which the RNAseq V2 data were available were used for differential gene expression analysis comparing the

transcriptome profiles of the *CIC*-mutant and *CIC*-wild type tumor tissues. The gene level RSEM (<http://deweylab.biostat.wisc.edu/rsem/>) raw counts from the TCGA RNAseq V2 data were rounded to the nearest integer for each gene in each sample. The data were normalized by variance stabilizing transformation using the Bioconductor package DESeq that takes into account the RNA-seq data size of each sample (<http://bioconductor.org/packages/release/bioc/html/DESeq.html>). The differential gene expression in the *CIC*-mutant *vs.* *CIC*-wild type oligodendrogliomas was analyzed using the significance analysis of microarrays (SAM) tool in the MeV version 4.9.0 (www.TM4.org). The pathway enrichment analysis of the gene set significantly differentially expressed in the *CIC*-mutant *vs.* *CIC*-wild type was carried out using the Web based Gene Set Analysis Toolkit (<http://bioinfo.vanderbilt.edu/webgestalt>). The gene set enrichment analysis was carried out using the SeqGSEA package (1.8.0 version) of the R bioconductor (www.bioconductor.org).

RESULTS

Copy number variation analysis of the exome sequence data of 11 tumors diagnosed histopathologically as oligodendrogliomas was done using two algorithms, Fishing CNV and Control-FREEC, taking into account both the coverage and B allele frequency. The analysis identified concurrent loss of a copy of 1p and 19q, a known characteristic of oligodendrogliomas, in 9 out of the 11 tumor tissues (Fig. 1, Supporting Information Fig. 1). Other than the 1p/19q codeletion, recurrent chromosome 14 deletions were found in three 1p/19q codeleted tumors. The number of non-synonymous somatic variants per tumor genome ranged from 10 to 46 (Supporting Information Table 1). R132H/R132C or R172K mutation in the *IDH1* or *IDH2* gene, respectively, was identified in 10 tumors while one tumor lacked mutation in the *IDH1* as well as *IDH2* gene (Supporting Information Fig. 2). Four of the 9 tumors with 1p/19q codeletion were found to carry a missense or a frame-shift deletion mutation in the *CIC* gene, located on 19q while two tumors carried mutations in *FUBP1* gene located on 1p. Two tumors with 1p/19q codeletion but no somatic alteration in the *CIC* gene were found to carry an activating mutation (Q61L, G12D) in the *KRAS* gene. Recurrent mutations were identified in the Notch signaling pathway genes, including four tumors with mutation in *NOTCH1* and one tumor

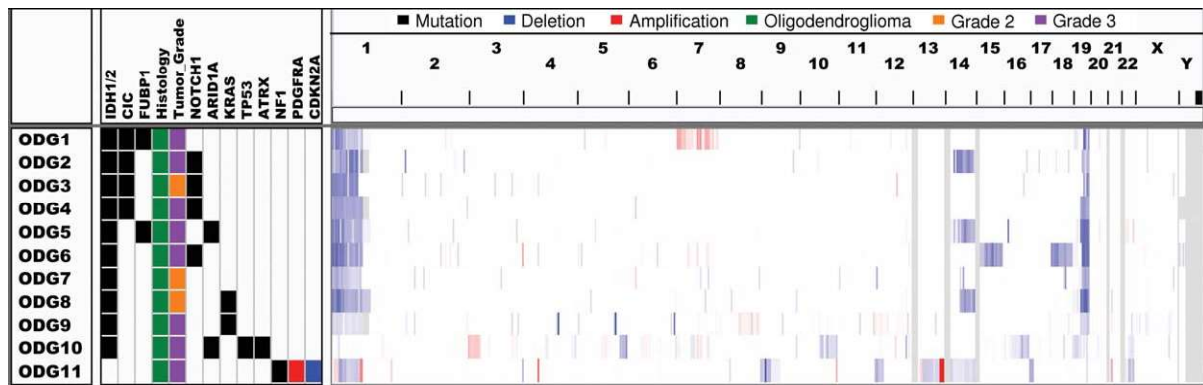


Figure 1. Integrated genomic view of the copy number variation and somatic mutations of the 11 oligodendrogliomas as analyzed using the FishingCNV software. The tumors are numbered sequentially (ODG1 to ODG11). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

with a mutation in *MAML3*. Two tumors were found to carry a mutation in the chromatin modifier *ARID1A* gene (Supporting Information Table 1, Supporting Information Fig. 3).

The tumor ODG10 lacking 1p/19q codeletion carried mutations in the *ATRX*, *TP53*, and *IDH1* genes. ODG11 lacked 1p/19q codeletion as well as a mutation in *IDH1/IDH2*. This tumor was found to carry a frame-shift deletion in the *NF1* gene, amplification of the *PDGFRA* gene, and deletion of 9p including the *CDKN2A* gene locus.

RNA-seq data on 65 oligodendrogliomas with 1p/19q codeletion from the TCGA cohort was analyzed for differential gene expression. SAM analysis identified 148 genes to be significantly differentially expressed in the 39 *CIC*-mutant oligodendrogliomas as compared with the 26 *CIC*-wild type oligodendrogliomas from the TCGA cohort at a False Discovery Rate of <5% (Fig. 2A, Supporting Information Tables 2 and 3). The differential gene expression comparing the *CIC*-mutant and *CIC*-wild type oligodendrogliomas from our cohort as well as the TCGA cohort was also done using EdgeR analysis (Supporting Information Table 3). The genes identified to be significantly differentially expressed in the TCGA cohort showed differential expression in our cohort as well, although some genes did not reach statistical significance due to the small sample size (Fig. 2B; Supporting Information Table 3). *ETV1*, *ETV4*, and *ETV5*, the three genes belonging to the ETS/PEA3 family of transcription factors, were found to be upregulated in the *CIC*-mutant tumors. The gene set enrichment analysis of the TCGA data comparing the expression profiles of the *CIC*-mutant and *CIC*-wild type tumors identified a number of genes involved in the negative regulation of the MAP kinase (MAPK) signaling pathway

and those upregulated by the *KRAS* oncogene to be significantly enriched (Fig. 3). The KEGG pathway analysis of the gene set significantly differentially expressed between *CIC*-mutant and *CIC*-wild type tumors also identified enrichment of a number of genes in the MAPK signaling pathway ($P = 0.0019$ and $FDR = 0.0199$). These MAPK pathway genes included the dual specificity phosphatase genes *DUSP4*, *DUSP6*, and *DUSP19*, the Sprouty family members *SPRY4*, *SPRED1*, and *SPRED2*, and the receptor tyrosine kinase encoding genes *ALK*, *PDGFRA*, *FGFR1*, and *EPHB*. *CREB3L1*, a member of the CREB/ATF family transcription factors that modulates unfolded protein response signaling, was also found to be upregulated in the *CIC*-mutant oligodendrogliomas. Two oligodendrogliomas (ODG8 and ODG9) carried activating mutation in *KRAS* gene. These two cases had higher expression of some of the negative regulators of tyrosine kinase receptor signaling pathway such as *SPRY4*, *SPRED2*, and *DUSP6* (Fig. 2B). The TCGA data, however, contains only one tumor with an activating mutation in the *NRAS* gene out of the 65 oligodendrogliomas.

DISCUSSION

Molecular Markers Essential for Accurate Diagnosis of Adult Gliomas

Nine out of 11 oligodendrogliomas were found to carry 1p/19q codeletion and *IDH1* mutation, genetic alterations characteristic of oligodendrogliomas. However, one tumor (ODG10) lacking 1p/19q codeletion was found to carry mutations in the *IDH1*, *ATRX*, and *TP53* gene. Mutations in the *ATRX* gene have been shown to be restricted to *IDH1/IDH2*-mutated gliomas, mutually

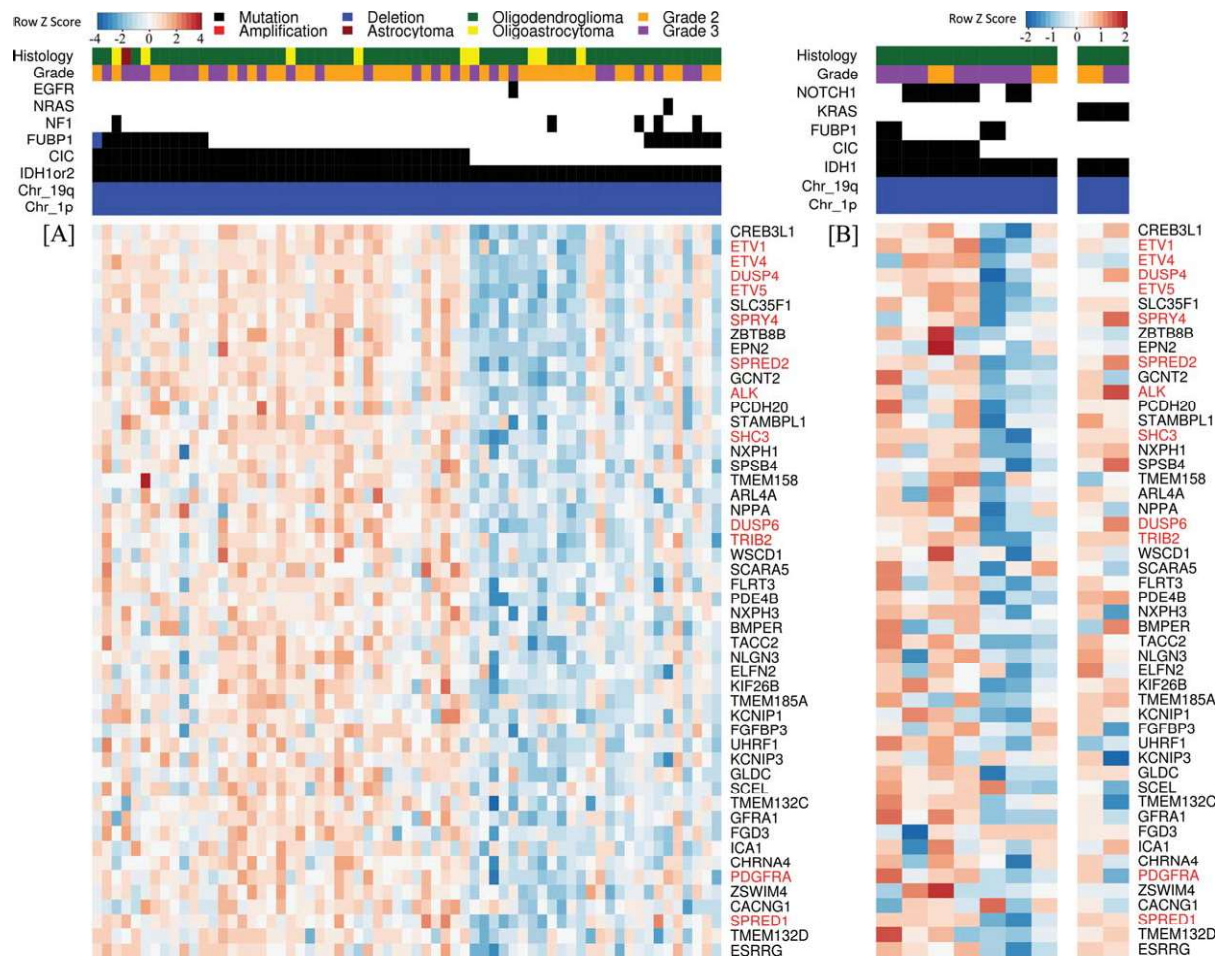


Figure 2. Heat map showing the top 50 genes most significantly differentially expressed in the 39 *CIC*-mutant vs. 26 *CIC*-wild type oligodendrogliomas from the TCGA cohort (A) and from the present study cohort of 9 oligodendrogliomas (B). The genes belonging to the MAPK signaling pathway are highlighted in red. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

exclusive with 1p/19q codeletion, and to correlate with astrocytic morphology (Jiao et al., 2012; Kannan et al., 2012). Therefore, although ODG10 was diagnosed as oligodendroglioma based on the characteristic histological appearance, it is identified as an astrocytoma based on the genetic alterations. One out of the 11 tumors without 1p/19q codeletion as well as *IDH1/IDH2* mutation lacked mutation in *ATRX* and *TP53* as well and hence cannot be classified as an oligodendroglioma or astrocytoma. This tumor was found to carry some of the genetic alterations known to occur in GBMs, such as *CDKN2A* deletion, *PDGFRA* amplification, and mutation in the *NF1* gene. Thus, based on the genetic alterations, this tumor is closer to GBMs than low grade gliomas. An integrated DNA methylation and copy-number profil-

ing study on a cohort of 228 anaplastic gliomas has also identified three similar molecular types (Wiesler et al., 2014). The three subtypes consisted of a group of *IDH1/2*-mutated CpG island methylator phenotype (CIMP)-positive tumors with 1p/19q codeletion, having the best prognosis, CIMP-positive tumors lacking 1p/19q codeletion having intermediate prognosis (likely to correspond to astrocytomas), and GBM-like CIMP-negative tumors having copy number alterations similar to those in GBMs, having the worst prognosis (Wiesler et al., 2014). Thus, in addition to the histopathological characterization, molecular characterization of adult gliomas is necessary for accurate diagnosis. Integrated diagnosis based on histopathology and molecular markers has been recently recommended for inclusion in the WHO

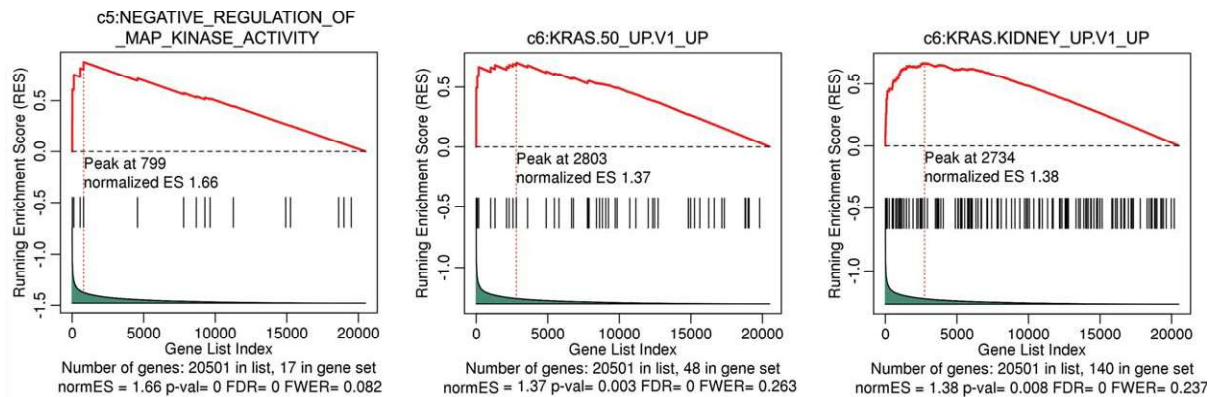


Figure 3. The gene sets significantly enriched in the differential expression analysis comparing the *CIC*-mutant and *CIC*-wild type tumors in the TCGA dataset done using the c5 GO gene sets and c6 oncogenic signature gene sets (Molecular Signatures Database v5.0). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

guidelines for diagnosis of central nervous system tumors by the International society of Neuropathologists (Louis et al., 2014).

Oncogenic Potential of the ETS/ETV/Pea3 Subfamily Transcription Factors Upregulated in the *CIC*-Mutant Oligodendrogliomas

CIC appears to act as a tumor suppressor gene in oligodendrogliomas, with one of the two copies of the gene deleted and the other copy carrying either protein truncating mutation or potentially deleterious missense mutation, as indicated by the present study as well as other reports (Bettegowda et al., 2011; Sahm et al., 2012). The missense mutations in the *CIC* gene in the TCGA study as well in our study were predominantly localized in the exons 5 and 20 of the *CIC* gene, as has been reported before (Yip et al., 2012). These exons of the *CIC* gene are known to be involved in the HMG box DNA binding domain and protein–protein interaction domain of the *CIC* protein, respectively. A number of studies indicate an oncogenic potential of the ETS/PEA3 family transcription factors upregulated in *CIC*-mutant oligodendrogliomas (Sharrocks, 2001; Oikawa and Yamada, 2003). The first *ETS* (E26 Transformation Specific) transcription factor-encoding gene was identified as a transforming gene in the avian E26 erythroblastosis virus. Twenty-eight human ETS family members are known that share ~85 amino acid long DNA binding ETS domain. The PEA3 (Polyoma virus enhancer activator 3) subfamily of the ETS family transcription factors includes ETV1, ETV4, and ETV5 (Oh et al., 2012), all of which were found to be upregulated

in *CIC*-mutant oligodendrogliomas. ETS family genes including *ETV1* and *ETV4* are known to be involved in chromosomal translocations in Ewing's sarcoma and in peripheral primitive neuroectodermal tumors resulting in EWS-ETS fusion protein (Jeon et al., 1995; Janknecht, 2005). The EWS-ETS fusion protein contains a N-terminal transactivation domain of the EWS gene and a C-terminal DNA binding domain of the *ETS* gene, suggesting a role for *ETS*-regulated genes in the pathogenesis of these tumors (Janknecht, 2005). The majority of the prostate cancers carry a chromosomal translocation involving the ETS family genes *ERG* or *ETV1/4/5* and the prostate organ specific, androgen-inducible *TMPRSS2* gene resulting in TMPRSS2-ETS fusion protein that is overexpressed in an androgen-inducible manner in prostate tumors (Tomlins et al., 2005, 2006; Helgeson et al., 2008).

Role of *CIC* in the RTK-RAS-MAPK Signaling Pathway and in the Expression of the ETV Transcription Factors

Cic, Capicua meaning head-and-tail in Catalan, was identified in developmental studies of *Drosophila* (Jimenez et al., 2000). *Drosophila Cic* plays an essential role downstream of the TORISO and the epidermal growth factor receptors, two tyrosine kinases that transmit the signaling via the RAS-RAF-MAP kinase pathway (Roch et al., 2002; Tseng et al., 2007; Jimenez et al., 2012). Apart from *Cic*'s role in the cell fate determination downstream of RTK pathways in *Drosophila*, *Cic* is also known to play a role in regulating growth of imaginal discs downstream of the RTK/RAS-

MAPK pathway. In human cells as well, *CIC* appears to play a role downstream of the RTK-MAPK signaling pathway (Dissanayake et al., 2011). EGF stimulation of melanoma cells results in the phosphorylation of *CIC* at multiple sites and upregulation of the *ETV4/ETV5* transcription factors. siRNA mediated knock-down of *CIC* in a melanoma cell line with constitutively activated ERK signaling was found to result in upregulation of *ETV1*, *ETV4*, and *ETV5* indicating *CIC*'s role in the suppression of PEA3 family transcription factors (Dissanayake et al., 2011). Some Ewing's sarcomas contain a *CIC-DUX4* translocation resulting in a fusion protein that converts the *CIC* protein from transcriptional repressor to an activator (Kawamura-Saito et al., 2006). Expression profile of Ewing's sarcomas containing the *CIC-DUX4* translocation showed upregulation of the *ETV1/ETV5* transcription factors. Furthermore, the *CIC-DUX4* fusion protein binds to the promoter region of *ETV5*, indicating upregulation of *ETV5* gene as a result of direct binding of the *CIC* repressor turned into activator. Ewing's sarcomas carrying *CIC-DUX4* translocation have also been reported to have distinct transcription profiles, with overexpression of ETV family members as compared to the *EWSR1-FLI1* fusion (Specht et al., 2014). Upregulation of the ETS/PEA3 transcription factors in the *CIC*-mutant oligodendrogliomas is thus consistent with their upregulation in Ewing's sarcomas carrying the *CIC-DUX4* translocation and upregulation in melanoma cells on siRNA mediated knock-down of the *CIC* gene expression.

DUSP4, *DUSP6*, and *DUSP19*, the dual specificity phosphatase encoding genes known to inactivate MAP kinases like ERK1/ERK2/JNK (Patterson et al., 2009), and the Sprouty family members *SPRY4*, *SPRED1*, and *SPRED2*, known to be inhibitors of the RTK-MAPK signaling pathway (Mason et al., 2006), were also found to be significantly upregulated in the *CIC*-mutant oligodendrogliomas. Upregulation of these negative regulators of the RTK signaling pathway is likely to be due to the constitutive activation of the pathway resulting from the inactive *CIC* mutant protein. Whether downregulation of *CIC* alone is sufficient for the upregulation of the RTK/MAPK signaling target genes like PEA3 transcription factors in oligodendrogloma tumor tissues needs to be investigated further.

In some of the oligodendrogliomas lacking mutations in the *CIC* gene, other components of the RTK/RAS/MAPK signaling pathway appear to

be activated by inactivating mutations in the negative regulators of the pathway like *NF1* and by activating mutations in *KRAS*, *NRAS*, and *EGFR* (Fig. 2A). PDGF expression in neural progenitor cells or overexpression of mutant EGFR under S100 beta promoter has been found to induce oligodendrogliomas in mouse models (Dai et al., 2001; Weiss et al., 2003; Appolloni et al., 2009). Thus, activation of the RTK/RAS/MAPK signaling pathway appears to be a major driver of the oligodendrogloma pathogenesis.

Genetic Alterations in Oligodendrogliomas and Their Possible Impact on Future Treatment Strategies

While oligodendrogliomas are known to be more sensitive to chemotherapy than astrocytomas, the tumors recur despite multimodal treatment including surgery, radiation and conventional chemotherapy (Engelhard et al., 2003). Development of novel treatment strategies based on the knowledge of the underlying genetic alterations is necessary for effective treatment of these tumors. The presence of mutations in chromatin modifier genes such as *ARID1A* and Notch signaling pathway genes, like *NOTCH1*, and *MAML3* may make these tumors amenable to treatment using chromatin modifying drugs and Notch signaling modulators (Bhalla, 2005). Higher expression of the oncogenic ETV transcription factors in the *CIC*-mutant oligodendrogliomas may make these tumors more aggressive than the *CIC*-wild type oligodendrogliomas. Loss of *CIC* expression has been reported to correlate with shorter progression free survival in a study done of 55 oligodendroglial tumors (Chan et al., 2014). Doxorubicin has been found to inhibit cancer cell proliferation by stimulating proteolytic cleavage of *CREB3L1* (Denard et al., 2012), a gene upregulated in the *CIC*-mutant oligodendrogliomas (Fig. 2). *CREB3L1* overexpression was shown to render human breast cancer and hepatoma cell lines sensitive to doxorubicin. While doxorubicin due to its poor ability to cross the blood brain barrier is not used for the treatment of oligodendrogliomas, its derivatives that can cross blood brain barrier may be effective in treatment of *CIC*-mutant oligodendrogliomas. Further validation of these findings using established oligodendrogloma cell lines and/or *in vivo* human tumor xenografts is necessary to translate the genetic alterations identified in oligodendrogliomas into the development of

effective targeted treatment strategies for this presently incurable malignant brain tumor.

ACKNOWLEDGMENTS

The authors thank Mr. Anand Sawant for technical assistance and Department of Biotechnology for bioinformatics facility at ACTREC.

REFERENCES

- Appolloni I, Calzolari F, Tutucci E, Caviglia S, Terrile M, Corte G, Malatesta P. 2009. PDGF-B induces a homogeneous class of oligodendrogliomas from embryonic neural progenitors. *Int J Cancer* 124:2251–2259.
- Bettgowda C, Agrawal N, Jiao Y, Sausen M, Wood LD, Hruban RH, Rodriguez FJ, Cahill DP, McLendon R, Riggins G, Velculescu VE, Oba-Shinjo SM, Marie SK, Vogelstein B, Bigner D, Yan H, Papadopoulos N, Kinzler KW. 2011. Mutations in CIC and FUBP1 contribute to human oligodendroglioma. *Science* 333:1453–1455.
- Bhalla KN. 2005. Epigenetic and chromatin modifiers as targeted therapy of hematologic malignancies. *J Clin Oncol* 23:3971–3993.
- Bromberg JE, van den Bent MJ. 2009. Oligodendrogliomas: Molecular biology and treatment. *Oncologist* 14:155–163.
- Chan AK, Pang JC, Chung NY, Li KK, Poon WS, Chan DT, Shi Z, Chen L, Zhou L, Ng HK. 2014. Loss of CIC and FUBP1 expressions are potential markers of shorter time to recurrence in oligodendroglial tumors. *Mod Pathol* 27:332–342.
- Collins VP. 2004. Brain tumours: Classification and genes. *J Neurol Neurosurg Psychiatry* 75 (Suppl 2):ii2–ii11.
- Dai C, Celestino JC, Okada Y, Louis DN, Fuller GN, Holland EC. 2001. PDGF autocrine stimulation dedifferentiates cultured astrocytes and induces oligodendrogliomas and oligoastrocytomas from neural progenitors and astrocytes in vivo. *Genes Dev* 15:1913–1925.
- Denard B, Lee C, Ye J. 2012. Doxorubicin blocks proliferation of cancer cells through proteolytic activation of CREB3L1. *Elife* 1:e00090.
- Dissanayake K, Toth R, Blakey J, Olsson O, Campbell DG, Prescott AR, MacKintosh C. 2011. ERK/p90(RSK)/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicua. *Biochem J* 433: 515–525.
- Engelhard HH, Stelea A, Mundt A. 2003. Oligodendroglioma and anaplastic oligodendroglioma: Clinical features, treatment, and prognosis. *Surg Neurol* 60:443–456.
- Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR, Cao X, Singla N, Montie JE, Varambally S, Mehra R, Chinnaiyan AM. 2008. Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. *Cancer Res* 68:73–80.
- Ichimura K, Pearson DM, Kocalkowski S, Backlund LM, Chan R, Jones DT, Collins VP. 2009. IDH1 mutations are present in the majority of common adult gliomas but rare in primary glioblastomas. *Neuro Oncol* 11:341–347.
- Janknecht R. 2005. EWS-ETS oncoproteins: The linchpins of Ewing tumors. *Gene* 363:1–14.
- Jenkins RB, Blair H, Ballman KV, Giannini C, Arusell RM, Law M, Flynn H, Passe S, Felten S, Brown PD, Shaw EG, Buckner JC. 2006. A t(1;19)(q10;p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma. *Cancer Res* 66:9852–9861.
- Jeon IS, Davis JN, Braun BS, Sublett JE, Roussel MF, Denny CT, Shapiro DN. 1995. A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. *Oncogene* 10:1229–1234.
- Jiao Y, Killela PJ, Reitman ZJ, Rasheed AB, Heaphy CM, de Wilde RF, Rodriguez FJ, Rosenberg S, Oba-Shinjo SM, Nagahashi Marie SK, Bettgowda C, Agrawal N, Lipp E, Pirozzi C, Lopez G, He Y, Friedman H, Friedman AH, Riggins GJ, Holdhoff M, Burger P, McLendon R, Bigner DD, Vogelstein B, Meeker AK, Kinzler KW, Papadopoulos N, Diaz LA, Yan H. 2012. Frequent ATRX, CIC, FUBP1 and IDH1 mutations refine the classification of malignant gliomas. *Oncotarget* 3:709–722.
- Jimenez G, Guichet A, Ephrussi A, Casanova J. 2000. Relief of gene repression by torso RTK signaling: Role of capicua in Drosophila terminal and dorsoventral patterning. *Genes Dev* 14:224–231.
- Jimenez G, Shvartsman SY, Paroush Z. 2012. The Capicua repressor—A general sensor of RTK signaling in development and disease. *J Cell Sci* 125:1383–1391.
- Kannan K, Inagaki A, Silber J, Gorovets D, Zhang J, Kastenhuber ER, Heguy A, Petrini JH, Chan TA, Huse JT. 2012. Whole-exome sequencing identifies ATRX mutation as a key molecular determinant in lower-grade glioma. *Oncotarget* 3:1194–1203.
- Kawamura-Saito M, Yamazaki Y, Kaneko K, Kawaguchi N, Kanda H, Mukai H, Gotoh T, Motoi T, Fukayama M, Aburatani H, Takizawa T, Nakamura T. 2006. Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum Mol Genet* 15:2125–2137.
- Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, Scheithauer BW, Kleihues P. 2007. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 114:97–109.
- Louis DN, Perry A, Burger P, Ellison DW, Reifenberger G, von Deimling A, Aldape K, Brat D, Collins VP, Eberhart C, Figarella-Branger D, Fuller GN, Giangaspero F, Giannini C, Hawkins C, Kleihues P, Korshunov A, Kros JM, Beatriz Lopes M, Ng HK, Ohgaki H, Paulus W, Pietsch T, Rosenblum M, Rushing E, Soylemezoglu F, Wiestler O, Wesseling P. 2014. International society of neuropathology—Haarlem consensus guidelines for nervous system tumor classification and grading. *Brain Pathol* 24:429–435.
- Mason JM, Morrison DJ, Basson MA, Licht JD. 2006. Sprouty proteins: Multifaceted negative-feedback regulators of receptor tyrosine kinase signaling. *Trends Cell Biol* 16:45–54.
- Oh S, Shin S, Janknecht R. 2012. ETV1, 4 and 5: An oncogenic subfamily of ETS transcription factors. *Biochim Biophys Acta* 1826:1–12.
- Oikawa T, Yamada T. 2003. Molecular biology of the Ets family of transcription factors. *Gene* 303:11–34.
- Ostrom QT, Gittleman H, Farah P, Ondracek A, Chen Y, Wolinsky Y, Stroup NE, Kruchko C, Barnholtz-Sloan JS. 2013. CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006–2010. *Neuro Oncol* 15:ii1–ii56.
- Patterson KI, Brummer T, O'Brien PM, Daly RJ. 2009. Dual-specificity phosphatases: Critical regulators with diverse cellular targets. *Biochem J* 418:475–489.
- Roch F, Jimenez G, Casanova J. 2002. EGFR signalling inhibits Capicua-dependent repression during specification of Drosophila wing veins. *Development* 129:993–1002.
- Sahm F, Koelsche C, Meyer J, Pusch S, Lindenberg K, Mueller W, Herold-Mende C, von Deimling A, Hartmann C. 2012. CIC and FUBP1 mutations in oligodendrogliomas, oligoastrocytomas and astrocytomas. *Acta Neuropathol* 123:853–860.
- Sharrocks AD. 2001. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol* 2:827–837.
- Specht K, Sung YS, Zhang L, Richter GH, Fletcher CD, Antonescu CR. 2014. Distinct transcriptional signature and immunoprofile of CIC-DUX4 fusion-positive round cell tumors compared to EWSR1-rearranged Ewing sarcomas: Further evidence toward distinct pathologic entities. *Genes Chromosomes Cancer* 53:622–633.
- Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM. 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310:644–648.
- Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, Cao X, Wei JT, Rubin MA, Shah RB, Chinnaiyan AM. 2006. TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res* 66: 3396–3400.
- Tseng AS, Tapon N, Kanda H, Cigizoglu S, Edelmann L, Pellock B, White K, Hariharan IK. 2007. Capicua regulates cell proliferation downstream of the receptor tyrosine kinase/ras signaling pathway. *Curr Biol* 17:728–733.
- Wang K, Li M, Hakonarson H. 2010. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38:e164.
- Weiss WA, Burns MJ, Hackett C, Aldape K, Hill JR, Kuriyama H, Kuriyama N, Milshteyn N, Roberts T, Wendland MF, DePinho

- R, Israel MA. 2003. Genetic determinants of malignancy in a mouse model for oligodendroglioma. *Cancer Res* 63:1589–1595.
- Wiestler B, Capper D, Sill M, Jones DT, Hovestadt V, Sturm D, Koelsche C, Berton A, Schweizer L, Korshunov A, Weiss EK, Schliesser MG, Radbruch A, Herold-Mende C, Roth P, Unterberg A, Hartmann C, Pietsch T, Reifenberger G, Lichter P, Radlwimmer B, Platten M, Pfister SM, von Deimling A, Weller M, Wick W. 2014. Integrated DNA methylation and copy-number profiling identify three clinically and biologically relevant groups of anaplastic glioma. *Acta Neuropathol* 128:561–571.
- Yip S, Butterfield YS, Morozova O, Chittaranjan S, Blough MD, An J, Birol I, Chesnelong C, Chiu R, Chuah E, Corbett R, Docking R, Firme M, Hirst M, Jackman S, Karsan A, Li H, Louis DN, Maslova A, Moore R, Moradian A, Mungall KL, Perizzolo M, Qian J, Roldan G, Smith EE, Tamura-Wells J, Thiessen N, Varhol R, Weiss S, Wu W, Young S, Zhao Y, Mungall AJ, Jones SJ, Morin GB, Chan JA, Cairncross JG, Marra MA. 2012. Concurrent CIC mutations, IDH mutations, and 1p/19q loss distinguish oligodendrogliomas from other cancers. *J Pathol* 226:7–16.