Proteomic profiling of leukemic cells in Philadelphia chromosome positive leukemia

By

Mythreyi Narasimhan

(LIFE09201304016)

Tata Memorial Centre, Mumbai

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfillment of requirements

for the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

December 2020

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 **Ms. Mythreyi Narasimhan** entitled "**Proteomic profiling of leukemic cells in Philadelphia chromosome positive leukemia**" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

31/12/2020

Chairperson – Dr. S.V.Chiplunkar — Date: 31/12/2020

Guide/Convener – Dr. Rukimini Govekar — Date:

31/12/2020

External Examiner – Dr. Vikram Mathews — Date: 1/01/2021

Member – Dr. Hari Misra — Date:

Member -Dr. P.G.Subramanian — Date: 4/1/2021

4/1/2021

Invitee – Dr. Geetanjali Sachdeva — 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: 31/12/2020

Place: Navi Mumbai

Dr. Rukmini Govekar
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.

Mythreyi Narasimhan

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.

Mythreyi Narasimhan

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out under my supervision. 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.

Dr. Rukmini Govekar

(Guide)

List of Publications arising from the thesis

Journal

1. **'Clinical biomarker discovery by SWATH-MS based label-free quantitative proteomics: impact of criteria for identification of differentiators and data normalization method'**. Mythreyi Narasimhan, Sadhana Kannan, Aakash Chawade, Atanu Bhattacharjee, Rukmini Govekar, Journal of Translational Medicine 2019, 17(1):184.
2. **'Inhibition of p38-MAPK activity to alleviate imatinib resistance in chronic myeloid leukemia'.** Mythreyi Narasimhan, Vaishnavi Khamkar, Sarika Tilwani, Sorab N. Dalal, Dhanlaxmi Shetty, Subramanian P.G., Sanjay Gupta, Rukmini Govekar (Manuscript submitted).

Conferences

1. 14th Indo-Australian Biotechnology conference, ACTREC, India, 22nd – 23rd Oct 2018.
   Nature of Participation**: <u>Poster presentation</u> (Title - Identification of potential therapeutic target for Imatinib- resistant chronic myeloid leukemia using proteomic approach).**

2. Gordon Research Conference on Drug resistance, Rhode Island, U.S.A., 22nd – 27th July 2018.
   Nature of Participation: **<u>Poster presentation</u> (Title - Mass spectrometric identification of a kinase and its regulator as potential therapeutic targets for tyrosine kinase inhibitor-resistant chronic myeloid leukemia).**

3. 7th Annual meeting of Proteomic Society of India (PSI), VIT University, Vellore, India, 3rd – 6th Dec.2015.
   Nature of Participation: **<u>Poster presentation</u> (Title - Preliminary studies to identify potential therapeutic targets in the blast crisis stage of chronic myeloid leukemia using proteomics approach)**.

Other publications:

1. **'Enrichment-free detection of hydrophobic proteins from complex biological samples by optimization of gradient composition and duration in tandem LC-MS'.** Mythreyi Narasimhan, Vaishnavi Khamkar, Rukmini Govekar (Manuscript submitted).

Mythreyi Narasimhan

This thesis
is
dedicated
to
my beloved parents

ACKNOWLEDGEMENTS

I take this opportunity to convey my heartiest regards to those who have contributed, supported and guided me throughout the entire duration of my PhD research.

First and foremost I would like to convey my sincere and heart felt gratitude to my mentor Dr.Rukmini Govekar for relentlessly guiding and supporting me through every small aspect of my Ph.D. She is a great scientist and an amazing person to work with. Her personality of being very kind and optimistic always kept me motivated and made me enjoy working in lab. During this tenure I learnt a lot in both scientific and non-scientific aspects. I consider myself lucky for getting an opportunity to work in a lab which made me enjoy what I do. She would always be my role model and I thank her for making my Ph.D. journey so memorable and enjoyable.

I would like to thank Dr. Sudeep Gupta (Director), Dr. Shubhada Chiplunkar (Ex-Director) and Dr. Prasanna Venkatraman (Dy. Director) for providing me an opportunity to work in an institute with good infrastructure. I am extremely thankful to my Doctoral committee members : Dr. Shubhada Chiplunkar (Chairperson), Dr. Hari Misra (B.A.R.C), Dr. P.G. Subramanian and Dr. Geetanjali Sachdeva (NIRRH) for their scientific inputs and critical assessment which helped in shaping up the project. My special thanks to Dr. P.G. Subramanian and Dr. Sweta Rajpal from Hematopathology department, ACTREC, for generously helping me with kinase domain mutation studies. I would like to thank Dr. Dhanalakshmi Shetty, Cytogenetics lab, ACTREC for FISH analysis; Dr. Sorab Dalal and Ms. Sarika Tilwani for helping with 14-3-3 epsilon knockout studies; Dr. Sanjay Gupta and Mr. Ramchandra Amnekar for providing reagents for p38-MAPK studies. I also extend my thank Dr. Ajit Datar, Mr. Shailendra Rane and Mr. Ashutosh Shelkar from Shimadzu analytical India Pvt. Ltd. for assessment of intracellular imatinib concentration. I also thank Mr. Shashi Dolas and Ms. Savita Prabhu Chavan from mass spectrometry facility, ACTREC, for their immense support for carrying out all LC-MS experiments; Ms. Shamal Vetale, Mr. Ravindra Joshi and Mr. Prashant Masane for their help in flow cytometry based assays.

I take this opportunity to thank all RGL family members. I am grateful to Mrs. Poonam Kawle for her support, guidance and immense care. The excellent training provided by her during my days of traineeship made me so systematic in work. I all thank all my present and past trainees who

have contributed towards the project. A special thanks to project JRF Ms. Vaishnavi Khamkar, who has been an integral part of this project. I thank my labmates Rahul, Pravin, Pranali, Vaishnavi, Shraddha, Neha and Manish for making the lab environment very lively, cheerful, friendly and stress-free. The list is incomplete without a special acknowledgement for my foody gang members Aneri, Vaishnavi, Pranali, Shraddha and my longtime friend Mukund for all our routine outings and fun times, which were the best stress buster. Thanks for being there for me guys!!

I am expressing my heartiest thanks and warm regards to my fabulous batch mates JRF-2013 – Shalini, Asmita, Ram, Raghav, Ajay, Raghupathi, Anuja, Vasudha, Mahalakshmi, Burhan, Pravin, Rajashree, Rahul and Dhanashree . Our mandatory birthday celebrations and get togethers have been so much fun.

I am grateful and indebted to my parents for allowing me to pursue my passion in research and being a constant source of support and encouragement throughout this journey. I also thank my husband Nitish Kunte for always being there providing scientific inputs as well as cheering me up whenever I am stressed and my in-laws for being very considerate and supportive during the crucial period of my Ph.D.

Finally, I would like to thank all those with whom I have interacted in ACTREC and I am thankful for all the help provided during my tenure.

Mythreyi Narasimhan

CONTENTS

SYNOPSIS

Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

1. **Name of the Student:** Mythreyi Narasimhan
2. **Name of the Constituent Institution:** Tata Memorial Centre, ACTREC
3. **Enrolment No.:** LIFE09201304016
4. **Title of the Thesis:** Proteomic profiling of leukemic cells in Philadelphia chromosome positive leukemia.
5. **Board of Studies:** Life science

SYNOPSIS

INTRODUCTION:

Chronic myeloid leukemia (CML) is a myeloproliferative disorder of hematopoietic stem cell origin [1]. The transforming principle of CML is the reciprocal translocation between chromosome 9 and chromosome 22 [t9;22 (q34;q11)] which results in the formation of a shortened chromosome 22 called Philadelphia chromosome. This encodes a fusion protein Bcr-Abl, a constitutively active tyrosine kinase, which activates signalling pathways that lead to increased proliferation, inhibition of apoptosis and altered cellular adhesion to bone marrow or stroma [2, 3]. CML progresses through 3 phases – an initial chronic phase (CP), followed by a transient accelerated phase (AP) which inevitably leads to the terminal blast crisis (BC) [4]. Targeted therapy using tyrosine kinase inhibitor (TKI) Imatinib (IM), a drug that specifically inhibits tyrosine kinase activity of Bcr-Abl, is the most successful therapeutic strategy for treating CML. It not only brings about complete haematological response in 96% patients in chronic phase but also an unprecedented complete cytogenetic remission

(CCyR) in 86% and 10 year major molecular remission in about 93% of newly diagnosed patients in chronic phase, thereby making CML an epitome of targeted therapy [5, 6]. However, about 15% patients in CP [6] do not respond to imatinib due to Bcr-Abl gene amplification, acquisition of mutations in the BCR-ABL kinase domain or altered drug transport. Dose escalation of IM and introduction of second and third generation TKIs could overcome resistance due to Bcr-Abl gene amplification and altered drug transport to a great extent [7, 8]. However, in patients with resistance due to kinase domain mutations next-gen TKIs could induce or restore CCyR in only 40%–50% of patients [9-11]. The remaining non-responders in CP eventually progress to BC where about 80% patients are resistant to imatinib with median survival of 11 months [6, 12]. Drug resistance in BC, in addition to the above mechanisms, also occurs due to leukemic cells progressing to a Bcr-Abl independent phenotype, as a result of accumulation of additional molecular alterations during disease progression [13]. Contrary to the driver role of *BCR-ABL* in CP, it has been suggested that in BC previously Bcr-Abl dependent pathways such as proliferation, apoptosis, DNA damage surveillance and repair become Bcr-Abl independent. As a result, this Bcr-Abl independent pathway mediates resistance to TKIs by providing survival advantage to the cells despite inhibition of Bcr-Abl tyrosine kinase activity, due to which next generation TKIs though have increased cytogenetic response, could provide no survival advantage over imatinib to resistant CML-BC patients [12]. Thus, with emergence of more and more kinase domain mutations, limits to target kinase domain without compromising on specificity and existence of Bcr-Abl independent resistance in blast crisis, identifying and targeting molecules other than Bcr-Abl becomes crucial, for treating advanced phase CML. In cases where resistance is mediated by kinase domain mutation, Bcr-Abl still remains the key driving factor. Hence targeting Bcr-Abl downstream signalling components would help in inhibiting Bcr-Abl mediated oncogenesis. In cases where signalling pathways other than those mediated by Bcr-Abl drive the oncogenic signalling, identifying and targeting proteins of this alternate pathway would help in overcoming resistance.

AIM:

To understand the mechanism/s of imatinib-resistance in CML-blast crisis and identify potential therapeutic targets using proteomic approach.

OBJECTIVES:

- Comparative proteomic profiling to identify differentiators in (1) IM-sensitive cells untreated or treated with IM to delineate the components of Bcr-Abl pathway (2) IM- sensitive and resistant cells treated with IM to identify resistance-associated proteomic alterations.
- Deciphering interactions between differentiators in silico to detect the hub molecules therein, validate their functional association and check for their role in development of resistance.

RESULTS & DISCUSSION

CML-BC cell lines K562, KCL22 and KU812 representing different lineages erythroblast, myeloblast and basophilic blast respectively were used in this study.

I. ESTABLISHMENT OF APPROPRIATE BIOLOGICAL SYSTEM AND CONDITIONS FOR INVESTIGATION

1. Basic characterization:

Morphology and doubling time of all 3 cell lines were assessed and was found to be consistent with that of reported literature. The duration at which cells remain in the mid log phase with good viability, was chosen as the optimal harvest time, which was identified as 48hrs for K562 and 72 hrs for KCL22 and KU812 cell lines. The IC_{50} of imatinib (IM) for all 3 cell lines was determined by MTT assay and was found to be 0.5μM for K562 and KU812 and 0.3 μM for KCL22.

2. Development of IM-resistant cells:

The wild type CML-BC cell lines were termed IM sensitive & designated as K562/S, KCL22/S & KU812/S. IM resistant cells were developed from their sensitive counterpart by gradual dose escalation of IM up to 1 μM and designated K562/R, KCL22/R and KU812/R. Assessment of resistance by MTT assay revealed that at IC_{50} concentration of the sensitive cells, their corresponding resistant cells had more than 80% viability, thereby confirming resistance. The IC_{50} of K562/R, KCL22/R and KU812/R cells was found to be 5 μM, 2 μM and 4 μM respectively, which is ~10-fold higher than their sensitive counterpart. Once developed, the resistant cells were always maintained in the presence of IM and this concentration was decided based on their ability to remain stable without loss of viability for a longer duration. K562/R was found stably resistant to 0.75 μM IM while KCL22/R and KU812/R cells to 1 μM IM and hence were maintained under this constant IM pressure.

3. Optimization of IM-treatment condition for comparative evaluation of S and R cells:

An optimal IM treatment condition is that IM concentration and treatment duration in which Bcr-Abl activity is maximally inhibited without compromising on cell viability. Since the study involves comparison between drug treated 'S' and 'R' cell, it necessitates treating 'S' cells with the same IM concentration used to maintain the corresponding 'R' cells. So, K562/S should be treated with 0.75 μM IM, while KCL22/S and KU812/S with 1 μM IM. But from IM IC50 data for these cell lines it is evident that treating with this IM concentration for 48hrs would result in >50% cell death. Hence, a treatment duration less than 48hrs, where cell viability is maintained as well as Bcr-Abl activity is inhibited, had to be identified. Viability of K562/S cells treated with 0.75 μM IM, KCL22/S and KU812/S cells treated with 1 μM IM for 2, 6, 12, 18 and 24 hrs., were assessed by trypan blue dye exclusion method and Annexin/FITC – PI staining by flow cytometry. Viability was found to be unaffected up to 24 hr. treatment in all the 3 cell lines, indicating that any treatment duration up to 24 hrs with maximal Bcr-Abl activity inhibition could be considered optimal. Bcr-Abl tyrosine kinase activity was assessed based on its ability to phosphorylate the downstream substrate STAT5, by western blotting. Western blotting revealed that in all 3 cell lines Bcr-Abl activity was maximally

inhibited with 12 hr. IM treatment. Thus, the optimal IM treatment condition for comparative evaluation is 0.75 μM for K562/SR and 1μM for KCL22/SR and KU812/SR with 12 hr. treatment.

II. ASSESSMENT OF THE STATUS OF KNOWN RESISTANT MECHANISMS

4.a. Bcr-Abl kinase domain mutation:

The sensitive and resistant counterpart of all the 3 cell lines were assessed for the presence of kinase domain mutation using sequencing-based approach. G250E mutation was observed in the P-loop of kinase domain in KCL22/R cells, which is known to induce sub-optimal response to Imatinib. All other cells were negative for kinase domain mutation.

4.b. Bcr-Abl gene amplification and overexpression:

Interface and metaphase FISH analysis revealed no difference in Bcr-Abl gene status between IM-sensitive and resistant cells, indicating absence of gene amplification as a mechanism of resistance in these cells. Bcr-Abl protein levels assessed by western blotting revealed a significant increase in only in K562/R.

4.c. Status of drug transporters:

Status of Bcr-Abl importer OCT-1 and export protein P-glycoprotein were checked by western blotting. No difference in the level of transporters was observed in KCL22/R cells while a significant increase in P-glycoprotein was observed in both K562/R and KU812/R cells indicating the possibility of reduced intracellular IM in these cells. This was further confirmed by detecting the intracellular IM levels by LC-MS analysis. Table 1 summarizes the outcome of the assessment of known resistant mechanisms, which clearly indicate the existence of at least one mechanism of resistance in all resistant cell lines. Hence it is anticipated that Bcr-Abl activity in ‘R’ cells would not be inhibited by imatinib, as effectively as in ‘S’ cells. To confirm this, the status of Bcr-Abl activity was assessed in S, S+IM and R cells.

Table 1: Summary of known resistant mechanisms

Cell line	Bcr-Abl amplification	Bcr-Abl overexpression	Kinase domain mutation	Altered drug transport
K562/R	**X**	✓	**X**	✓
KCL22/R	**X**	**X**	✓	**X**
KU812/R	**X**	**X**	**X**	✓

5. Status of Bcr-Abl activity:

Western blotting of pSTAT5/STAT5 revealed that in K562/R and KU812/R cells, upon IM treatment, Bcr-Abl activity was inhibited to the same extent as in sensitive cells, while a partial inhibition was observed in KCL22/R cells. This partial inhibition could be attributed to the presence of kinase domain mutation, indicating that the resistance in KCL22/R is probably mediated by the residual Bcr-Abl activity. However, inhibition of Bcr-Abl activity despite reduced intracellular imatinib and increased Bcr-Abl protein level in K562/R and KU812/R cells indicates that resistance is these cells is not mediated by the tyrosine kinase activity of Bcr-Abl.

IM resistance could be induced despite Bcr-Abl inhibition either due to activation of novel signalling pathways that provide survival advantage to the cells or due to molecular alterations that restore Bcr-Abl downstream signalling by activating its components in a Bcr-Abl independent manner. Exploring such alterations would necessitate global profiling to enable identification of previously uncharted molecular alterations. Hence proteomic analysis of K562 S, S+IM and R cells were carried out to study the underlying changes. The observed alterations were further validated in KU812 cells.

III. PROTEOMIC ANALYSIS TO IDENTIFY MEDIATORS OF RESISTANCE AND THEIR FUNCTION VALIDATION

6. Proteomic analysis of K562 cell line:

The following comparison groups were used for proteomic analysis.

K562/S vs S+IM: To identify Bcr-Abl downstream signalling pathway components.

K562/S+IM vs R: To identify components that confer resistance.

Quantitative proteomic analysis, using labelled (iTRAQ) and label-free (SWATH) approach, had been carried out for both the comparison groups. Since both the approaches are complementary, the differentiators (p-value < 0.05) identified from both the methods were pooled and subjected to further analysis. Differentiators were categorized into different functional groups and the signalling molecules identified were further sub-categorized into those that are -

I. Unique to S vs S+IM comparison group: This include proteins that are involved in Bcr-Abl downstream signalling.

II. Common to both comparison groups: This include Bcr-Abl pathway components which are altered in resistance.

III. Unique to S+IM vs R comparison group: This include proteins altered in resistance which are not part of Bcr-Abl pathway.

This categorization revealed that some signalling molecules belonging to Bcr-Abl pathway were also altered in resistant cells. Further, to identify the Bcr-Abl interactors and their functional association with the other differentiators, a string analysis was carried out for both the comparison groups by introducing Abl as one of the inputs [as Bcr-Abl is not an annotated SWISS-PROT entry] and Bcr-Abl pathway was compiled *in silico*. The 2 key observations from this analysis are (a) in both the comparison groups, the functional association of Abl with multiple proteins converged onto 14-3-3 family proteins, indicating their potential as hub molecule. (b) p38 MAPK, a stress induced Ser/Thr kinase reported to be involved in resistance of multiple cancers, was not only found to be associated with Abl but also altered in resistance. Hence these two proteins were further studied to assess their association with IM-resistance.

7. Validation of 14-3-3 family proteins:

Western blotting analysis of 14-3-3 ε revealed a significant downregulation in R cells compared to S cells. To check if this downregulation of 14-3-3 ε is associated with IM-resistance, 14-3-3 ε was knocked out in K562/S cells using cripr-cas9 system and the effect of knock out on IC_{50} of IM was assessed by MTT assay. MTT assay revealed a significant increase in IC_{50} confirming that down regulation of 14-3-3 ε, a key component of Bcr-Abl pathway, could mediate development of resistant phenotype. This observation of downregulation of 14-3-3 ε was consistent in KU812/R, the basophilic blast cells in which like in K562/R, resistance is not mediated by Bcr-Abl tyrosine kinase activity.

8. Validation of p38 MAPK:

Western blotting did not reveal a significant change in the levels of p38-MAPK in K562/R cells, but its phosphorylation was found to be significantly increased in K562/R cells. So, to understand if increased phosphorylation of p38 lead to resistance, its activity in K562/R cells was inhibited using the inhibitor SB203580 and its effect on resistance to IM was assessed. Since the inhibitor binds to ATP binding domain of p38 and inhibits its downstream signalling, the efficiency of p38 inhibition was assessed based on the phosphorylation status of its downstream substrate MSK1. 1hr treatment with 10μM inhibitor significantly downregulated phosphorylation of MSK1, confirming inhibition of p38 activity. MTT assay of K562/R cells with and without p38 inhibitor revealed about 6-fold reduction in IC_{50} of IM, upon p38 activity inhibition. This clearly indicates the role of p38-MAPK in mediating IM-resistance and was further substantiated by the increase in p38-MAPK phosphorylation observed in KU812/R cells. Sensitization of K562/R cells to IM by inhibition of p38 activity thus indicates its potential as a therapeutic target to overcome IM- resistance

OUTCOME OF THIS STUDY

- This study for the first time identified inhibition of p38 activity, a key component of Bcr-Abl pathway, as a potential therapeutic strategy to overcome resistance in CML. The activity of the protein is elevated in Bcr-Abl pathway and is further upregulated in IM-resistant cells. Identification of a Bcr-Abl pathway component which can be inhibited would assist in controlling Bcr-Abl pathway in CML-CP with Bcr-Abl containing mutant kinase domain which cannot be inhibited by TKIs. Moreover, being overexpressed in BC, its inhibition could lead to imatinib sensitivity.
- Further, this study identified the association between downregulation of 14-3-3 ε and imatinib resistance in CML-BC. Identifying its upstream modulators would thereby help in better understanding the role of 14-3-3 ε in resistance.

REFERENCES

1. Pasternak G, Hochhaus A, Schultheis B, Hehlmann R: **Chronic myelogenous leukemia: molecular and cellular aspects**. *Journal of cancer research and clinical oncology* 1998, **124**(12):643-660.
2. Kurzrock R, Gutterman JU, Talpaz M: **The molecular genetics of Philadelphia chromosome-positive leukemias**. *The New England journal of medicine* 1988, **319**(15):990-998.
3. Pane F, Intrieri M, Quintarelli C, Izzo B, Muccioli GC, Salvatore F: **BCR/ABL genes and leukemic phenotype: from molecular mechanisms to clinical correlations**. *Oncogene* 2002, **21**(56):8652-8667.
4. Kurzrock R, Kantarjian HM, Druker BJ, Talpaz M: **Philadelphia chromosome-positive leukemias: from basic mechanisms to molecular therapeutics**. *Annals of internal medicine* 2003, **138**(10):819-830.

5. Jabbour E, Kantarjian H: **Chronic myeloid leukemia: 2018 update on diagnosis, therapy and monitoring**. 2018, **93**(3):442-459.
6. An X, Tiwari AK, Sun Y, Ding P-R, Ashby Jr CR, Chen Z-S: **BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: a review**. *Leukemia research* 2010, **34**(10):1255-1268.
7. Jabbour E, Kantarjian HM, Jones D, Shan J, O'Brien S, Reddy N, Wierda WG, Faderl S, Garcia-Manero G, Verstovsek S *et al*: **Imatinib mesylate dose escalation is associated with durable responses in patients with chronic myeloid leukemia after cytogenetic failure on standard-dose imatinib therapy**. *Blood* 2009, **113**(10):2154-2160.
8. Eechoute K, Sparreboom A, Burger H, Franke RM, Schiavon G, Verweij J, Loos WJ, Wiemer EAC, Mathijssen RHJ: **Drug Transporters and Imatinib Treatment: Implications for Clinical Practice**. *Clinical Cancer Research* 2011, **17**(3):406-415.
9. Rosti G, Castagnetti F, Gugliotta G, Baccarani M: **Tyrosine kinase inhibitors in chronic myeloid leukaemia: which, when, for whom?** *Nature reviews Clinical oncology* 2017, **14**(3):141-154.
10. Zaharieva M, Amudov G, Konstantinov S, Guenova M: **Modern Therapy of Chronic Myeloid Leukemia**. *Edited by Margarita Guenova and Gueorgui Balatzenko* 2013:227.
11. Yang K, Fu LW: **Mechanisms of resistance to BCR-ABL TKIs and the therapeutic strategies: A review**. *Crit Rev Oncol Hematol* 2015, **93**(3):277-292.
12. Hehlmann R: **How I treat CML blast crisis**. *Blood* 2012, **120**(4):737-747.
13. Soverini S, Mancini M, Bavaro L, Cavo M, Martinelli G: **Chronic myeloid leukemia: the paradigm of targeting oncogenic tyrosine kinase signaling and counteracting resistance for successful cancer therapy**. *Mol Cancer* 2018, **17**(1):49.

PUBLICATIONS IN REFEREED JOURNAL:

a. Published – Mythreyi Narasimhan, Sadhana Kannan, Aakash Chawade, Atanu Bhattacharjee, Rukmini Govekar "Clinical biomarker discovery by SWATH-MS based label-free quantitative proteomics: impact of criteria for identification of differentiators and data normalization method". Journal of Translational Medicine 2019, 17(1):184.

b. Accepted: NA

c. Communicated: NA

Other Publications:

a. Book/Book Chapter - NA

b. Conference/Symposium

- 14th Indo-Australian Biotechnology conference, ACTREC, India, 22nd – 23rd Oct 2018.

  Nature of Participation**: Poster presentation (Title - Identification of potential therapeutic target for Imatinib- resistant chronic myeloid leukemia using proteomic approach)**

- Gordon Research Conference on Drug resistance, Rhode Island, U.S.A., 22nd – 27th July 2018.

  Nature of Participation: **Poster presentation (Title - Mass spectrometric identification of a kinase and its regulator as potential therapeutic targets for tyrosine kinase inhibitor- resistant chronic myeloid leukemia)**

- 7th Annual meeting of Proteomic Society of India (PSI), VIT University, Vellore, India, 3rd – 6th Dec.2015.

  Nature of Participation: **Poster presentation** **(Title - Preliminary studies to identify potential therapeutic targets in the blast crisis stage of chronic myeloid leukemia using proteomics approach)**

N. Mythreyi

Signature of Student:

Date: 27/11/19

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. S.V. Chiplunkar	Chairman	S.V. Chiplunkar	27.11.19
2.	Dr. Rukmini Govekar	Guide & Convener		27.11.19
3.	Dr. Hari Misra	Member		18/12/19
4.	Dr. P.G.Subramanian	Member		17/12/19
5	Dr. Geetanjali Sachdeva	Invitee		23/12/19

Forwarded Through:

S.V. Chiplunkar 7/1/20

Prof. S.V. Chiplunkar,
Chairperson, Academic &
Training Programme, ACTREC

(Prof. Dr. S. V. Chiplunkar)
Chairperson, Academic &
Training Programme , ACTREC

11/1/2020

Prof. S. D. Banavali,
Dean (Academics)
T.M.C.

PROF. S. D. BANAVALI, MD
DEAN (ACADEMICS)
TATA MEMORIAL CENTRE
MUMBAI – 400 012.

Version approved during the meeting of Standing Committee of Deans held during 29-30 Nov 2013

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Chronic myeloid leukemia (CML) is a malignancy of hemopoietic cells marked by myeloproliferation. An oncogenic fusion protein is causally associated with CML. Targeting the oncoprotein is highly successful in controlling the disease in its initial stage. Thus, CML epitomises successful targeted therapy. However, the success of targeted therapy diminishes with disease progression and 80% patients in the late stage do not respond to the targeted therapy. These non-responders have a survival of 6-12 months and it is necessary to explore alternatives for treatment of these patients which is the focus of this study. Biology of CML, targeted therapy -successes and failures, literature reports on current approaches to identify therapeutic targets and the lacunae therein are introduced.

1.1. BIOLOGY OF CML

1.1.1. CLINICAL COURSE

CML is characterized by neoplastic transformation and clonal expansion of hematopoietic stem cell which differentiates into common myeloid progenitors (CMP) and common lymphoid progenitors (CLP) that give rise respectively to granulocytes, erythrocytes, monocytes, megakaryocytes and to lymphocytes [1, 2]. With progression of the disease there is alteration in the cytological profile of peripheral blood which make the progressive phases of the disease discernible.

CML progresses from an initial indolent chronic phase (CP) to an intermediate accelerated phase (AP) and terminal blast crisis (BC). In some cases, AP is indistinguishable from BC. Under untreated condition or in case of resistance to current therapies the progression from CP to BC is inevitable, which takes about 3-4 years and upon reaching BC, the median survival of patients is only about 6-12 months [3]. The key characteristics of different phases of CML are as follows.

Chronic phase:

Patients in this phase are mostly asymptomatic and diagnosed incidentally during routine blood tests or physical examination [2, 4]. CP is defined by the presence of less than 10% myeloid blasts in bone marrow and circulation [5]. The cells retain their ability to differentiate normally and an increase in myeloid progenitors and cells of granulocytic lineage is observed, with notable neutrophilic leukocytosis [3, 6]. In CP very less B-cells and rarely any T-cells originating from neoplastic clone are detected [2]. At this stage progenitor cells still depend on growth factors for survival and proliferation [7]. Spleenomegaly is observed in most cases [8].

Accelerated phase:

In this phase the number of immature cells increase with about 10-19% blasts in circulation or bone marrow [5]. It is also characterized by basophilia with >20% basophils in peripheral blood or thrombocytopenia not induced by treatment [9]. At this stage cells lose their terminal differentiation capability and become less dependent on growth factors for survival. This phase lasts for about 4-6 months [10]. Some patients progress from CP to BC with unidentifiable AP [9].

Blast crisis:

This stage is characterized by complete differentiation arrest [10] and the percentage of blasts in bone marrow and circulation increases to >20% [5]. Despite the myeloid preponderance observed in CP, in BC immature cells of either myeloid or lymphoid lineage proliferate rapidly resulting in myeloid blast crisis in 65% cases, lymphoid blast crisis in 30% cases and blasts of megakaryocytic, erythroid or mixed lineage in about 5% cases [6, 10].

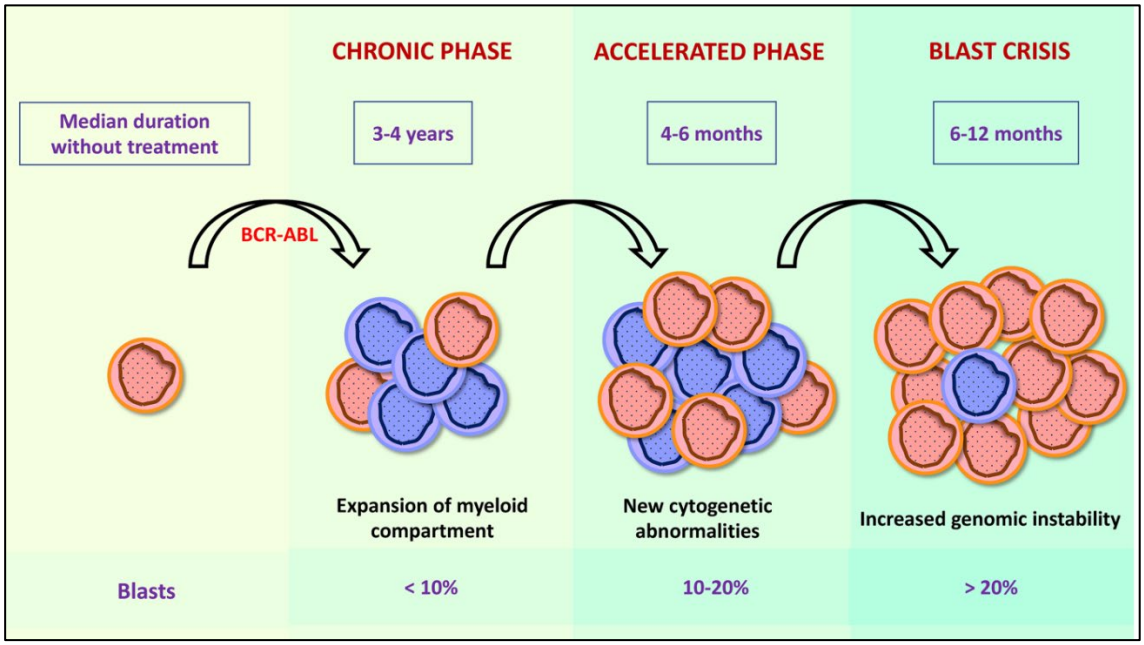


Fig. 1.1: *Clinical and cytological features of initial chronic phase, intermediate accelerated phase and terminal blast crisis under untreated condition.*

1.1.2. MOLECULAR PATHOBIOLOGY

The reciprocal translocation event [t(9;22) (q34;q11)] that occurs between the proto oncogene Abelson (ABL) in chromosome 9 and break-point cluster gene (BCR) in chromosome 22 is the hallmark of CML which results in shortened chromosome 22 called Philadelphia chromosome (Ph) (Fig.1.2) [6]. The association between Ph-chromosome and CML was established by Nowell and Hungerford in 1960 [8]. The translocated region on Ph chromosome encodes an oncogenic fusion protein Bcr-Abl, which is a constitutively active tyrosine kinase that interacts with proteins involved in various signalling pathways leading to altered cellular adhesion, increased proliferation and inhibition of apoptosis, thereby transforming the Ph+ hemopoietic stem cell [11]. Animal studies have shown that tyrosine kinase activity Bcr-Abl fusion protein is sufficient to induce chronic phase of CML in mice [2].

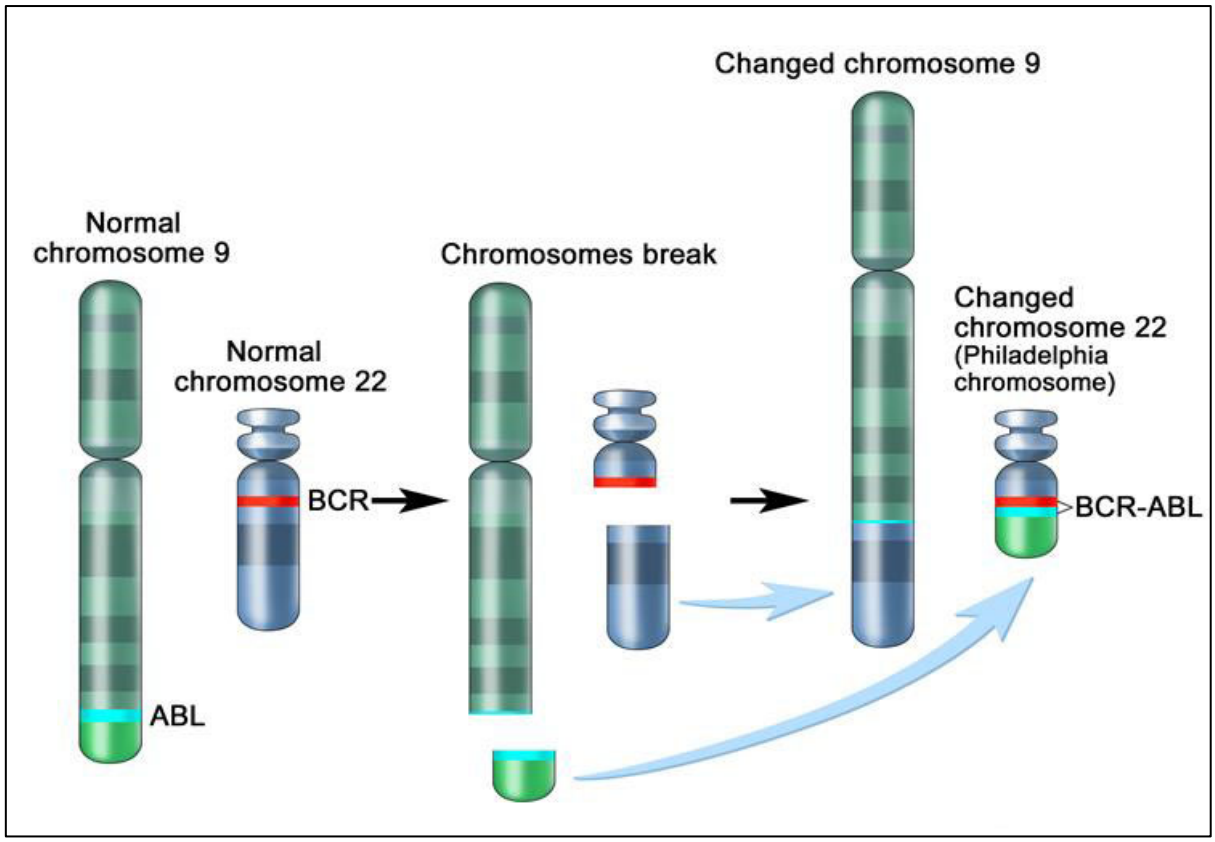


Fig. 1.2: *Chromosomal translocation between ABL in chromosome 9 and BCR in chromosome 22 leads to formation of Philadelphia chromosome which encodes BCR-ABL fusion protein (Adapted from [12]).*

About 90% CML patients harbour Ph-chromosome which is generally detected by cytogenetic analysis like fluorescent in-situ hybridization (FISH) [6]. About 5% patients harbour either simple variant translocation which occurs between chromosome 22 and a chromosome other than chromosome 9 or complex variant translocation which involve one or more chromosomes in addition to chromosome 9 and 22. A small percentage (2-5%) of patients exhibit hematological and morphological features of CML, with no Ph-chromosome detectable by cytogenetic analysis. In 1/3rd of these patients Bcr-Abl gene product could be detected using more sensitive molecular methods like RT-PCR [1, 13]. About 70% patients harbour transcriptionally active *ABL-BCR* gene whose function is not known [1, 14].

Though Bcr-Abl is the key alteration vital for CML initiation, disease progression to AP and BC is believed to be associated with accumulation of other chromosomal abnormalities [15]. This includes trisomy 8, trisomy 19, isochromosome 17q, additional loss of genetic material from 22q or loss of Y chromosome. Furthermore about 50% of CML-BC patients harbour multiple Ph chromosome [1, 13, 16].

1.1.3. *BCR, ABL* AND *BCR-ABL*: STRUCTURE AND FUNCTIONAL DOMAINS

***ABL*:**

ABL proto-oncogene is located on the long arm of chromosome 9 (band q34) [6], which spans 230Kb and contains 11 exons with 5'end oriented towards centromere. *c-Abl* has two first exons 1a and 1b separated by long (>200Kb) intron. Alternative splicing of these 2 exons result in 2 splice variants, a 6-Kb transcript with exon 1a and a 7-Kb transcript with exon 1b. The variants encode 2 isoforms of c-abl protein, both 145KDa, with major difference being in the presence of 19 amino acid long myristate group at the N-terminus of c-abl encoded by 1b splice variant but not by 1a splice variant [3, 16, 17].

c-abl is an ubiquitously expressed non-receptor tyrosine kinase which shuttles between cytoplasm and nucleus [3]. The functional domains of c-abl protein are depicted in Fig. 1.3. It has an N-terminal cap which is myristoylated in isoform 1b, followed by Src homology (SH) domains 1, 2 and 3. SH1 is the catalytic domain with tyrosine kinase activity while SH2 and SH3 domains interact with proteins containing phosphotyrosine residues and proline rich sequence respectively [15, 18-20]. Towards C-terminal end there are three nuclear localization signals (NLS) and one nuclear export signal that essential for shuttling of c-abl between nucleus and cytosol, DNA binding domain and actin binding domain [8, 20]. c-abl is associated with diverse functions including cellular response to

genotoxic stress, actin polymerization, DNA damage and repair, regulation of cell cycle, transmission of cellular information through integrin signalling and apoptosis[18].

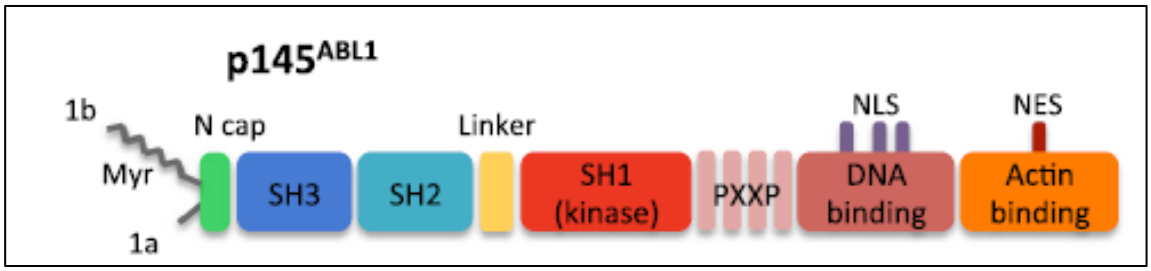


Fig. 1.3: Domain organization of ABL *- containing an N-terminal cap, tandem SH3, SH2 and SH1 (tyrosine-kinase) domains, proline-rich (PXXP) region, three nuclear localization signals (NLSs), one nuclear exporting signal (NES), a DNA-binding domain and an actin-binding domain. Adapted from [15].*

***BCR*:**

BCR gene on chromosome 22 (band q11) is 130Kb long with 23 exons. It transcribes into 4.5 Kb and 7 Kb splice variants, which encode 160 KDa and 130 KDa Bcr protein [16, 21]. Just like c-abl, Bcr protein is also ubiquitous with nuclear and cytosolic localization [3]. Fig.1.4 depicts the functional domains of Bcr protein. The first exon encodes three domains - a coiled coil domain spanning the first 61 amino acid residues which promotes oligomerization, a Ser/Thr kinase domain and SH2 binding domain. Exons 3-10 encode dbl-like and pleckstrin homology (PH) domain with GEF (Guanine nucleotide exchange factor) activity for Rho-GTPases and a calcium dependent lipid binding domain (CalB) [15]. The C-terminus contain Rac GAP (GTPase activating protein) homology domain. On the basis of functional domains, Bcr protein is speculated to have a role in signal transduction though its exact function is still unclear [7, 15, 18, 19, 21]

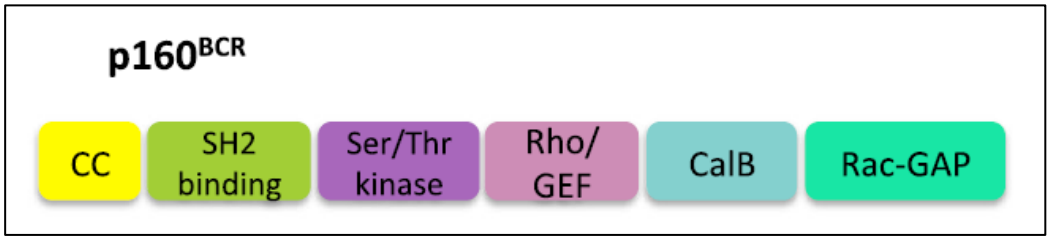


Fig. 1.4: Domain organization of BCR *- containing a coiled coil domain, SH2 binding domain, Ser/Thr kinase domain, Rho-GEF homology domain, calcium binding domain and Rac-GAP homology domain. Adapted from [15].*

BCR-ABL:

Philadelphia chromosome harbours the chimeric *BCR-ABL* gene, wherein 3'-region of *ABL* from chromosome 9 is juxtaposed to the 5'-region of *BCR* gene in chromosome 22 in a head-to-tail manner [3]. The breakpoint for translocation in *ABL* occurs within a single region as against *BCR* where there are three possible breakpoint regions. Depending on where the breakpoint occurs within *BCR* gene, three different Bcr-Abl fusion products are formed [17] (Fig. 1.5). In *ABL* gene, breakpoint occurs over the large $\geq$ 200Kb region either between the alternate exons 1a and 1b, downstream of exon 1a or upstream of exon 1b (Fig1.5a). Irrespective of where the breakpoint occurs, mRNA transcript of *BCR-ABL* always contain exons 2-11 (a2-a11) of *ABL*. Exon 1 of *ABL* despite being part of *BCR-ABL* gene on chromosome 22, is never included in mRNA transcript due to splicing [6, 18]. In 95% CML patients and in 1/3rd of Ph+ acute lymphoblastic leukemia (ALL) patients, breakpoint in *BCR* gene occurs within the major breakpoint cluster (M-Bcr) region which spans exons e12-16 (b1-b5) (Fig.1.5a). The break occurs within the intronic region immediate downstream of exon b2 (e13) or b3 (e14). As processing of *BCR-ABL* mRNA results in joining of *BCR* exons to exon 2 of *ABL*, 8.5Kb long b2a2 (e13a2) or b3a2 (e14a2) fusion transcripts are formed (Fig.1.5b). Both these encode the same 210KDa Bcr-Abl protein (p210$^{Bcr\text{-}Abl}$) [18, 21]. p210 Bcr-Abl contains a coiled-coil domain, SH2-binding domain and Ser/Thr kinase domain of Bcr and all domains of c-Abl except the N-

terminal cap region. Nuclear localization signal though retained, is inactivated in Bcr-Abl, resulting in its exclusive cytosolic localization [22].

The other two Bcr-Abl fusion products are p190 and p230 Bcr-Abl. If the breakpoint occurs at minor breakpoint cluster region (m-Bcr) i.e. within the 54.4Kb long first intronic region between alternative second exons e2' and e2, it results in the formation of a smaller *BCR-ABL* gene with 7Kb e1a2 fusion mRNA transcript. This encodes a 190KDa Bcr-Abl protein ($p190^{Bcr\text{-}Abl}$) which is associated with most cases of Ph+ pediatric ALL and 50% of Ph+ adult ALL. Breakpoint within the micro breakpoint cluster region (µ-Bcr) region which is downstream of exon 19 results in the largest *BCR-ABL* fusion gene with e19a2 mRNA transcript which encodes 230KDa Bcr-Abl protein. The extra *BCR* sequence in $p230^{Bcr\text{-}Abl}$ encodes additional 180 amino acids, compared to $p210^{Bcr\text{-}Abl}$ and is involved in development of chronic neutrophilic leukemia (CNL) in a subset of patients [16-19, 23].

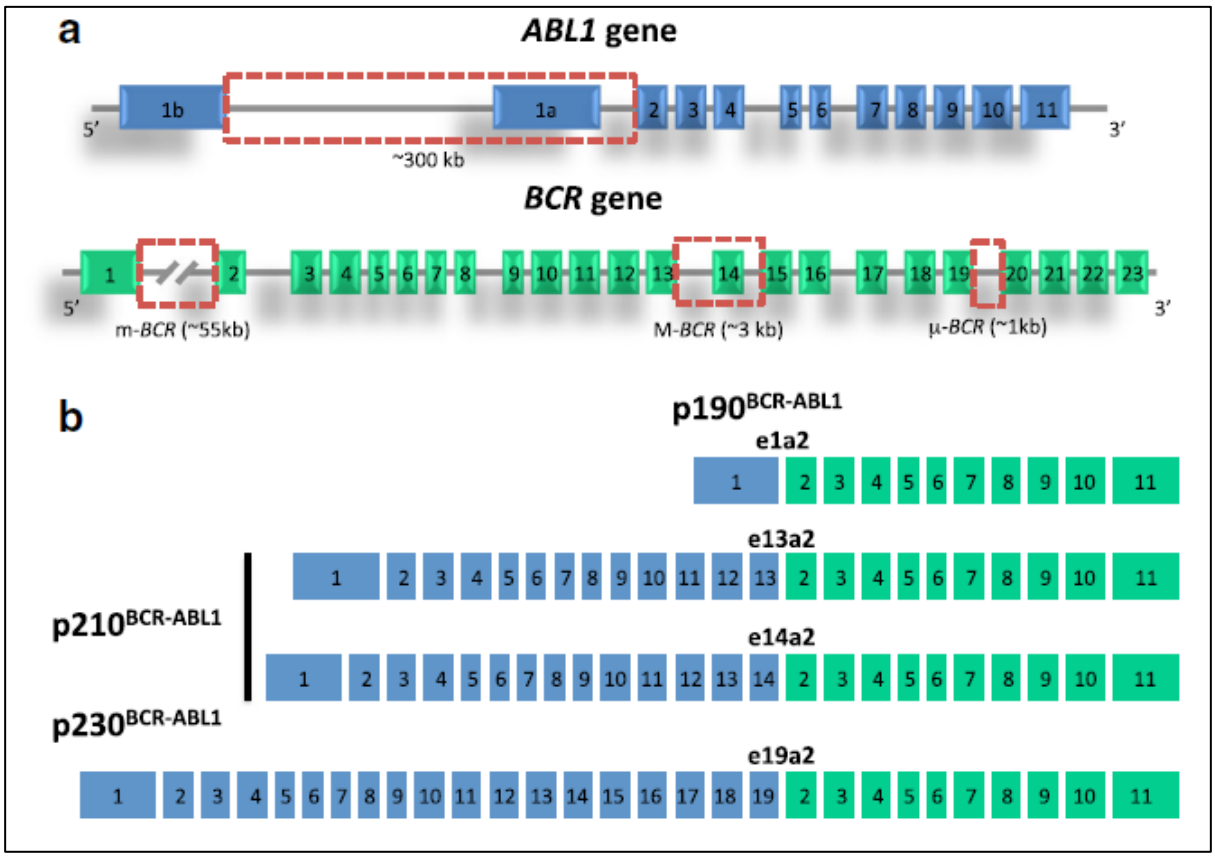


Fig. 1.5: *a. Breakpoints in ABL and BCR gene, b. BCR-ABL fusion transcripts. Adapted from [15].*

1.1.4. BCR-ABL MEDIATED TRANSFORMATION IN CML

Bcr-Abl mediates transformation through 3 important mechanisms - (a) constitutive activation of tyrosine kinase domain of Abl, (b) activating signalling pathways that induce increased proliferation and decreased apoptosis (c) altering cellular adhesion to bone marrow stroma and ECM [15].

1.1.4.1. Bcr-Abl tyrosine kinase activity

The coiled-coil domain of Bcr and loss of auto-inhibition of tyrosine kinase activity, contributes to the constitutive activation of Bcr-Abl. [19].

Coiled coil domain - This domain promotes homodimerization or tetramerization of Bcr-Abl protein, inducing mutual phosphorylation of tyrosine residues on kinase activation loop (Tyr-412) of each Bcr-Abl by transphosphorylation, thereby constitutively activating the tyrosine kinase [14, 24].

Loss of autoinhibition – Autoinhibition in c-abl is mediated through intramolecular interactions between N-terminal myristoyl cap, SH2 and SH3 domains. SH3 domain interacts with the linker region between SH2 and kinase domain, which contains proline-rich residue (PXXP). As a result, the linker region is sandwiched between SH3 domain and N-terminal lobe of kinase domain. This conformational change inhibits binding of substrate to the kinase domain. Binding of Abl inhibitory proteins (Abi 1 and 2) to SH3 domain further stimulate its kinase inhibitory function. In Bcr-Abl, this inhibitory function of SH3 domain is abolished as a result of fusion of Bcr sequences upstream. Further, constitutive activation of Bcr-Abl perturbs the conformation of the linker region preventing the sandwich formation.

SH2 domain mediates c-abl autoinhibition by interacting with the C-terminal lobe of kinase domain via hydrogen bonds. Binding of N-terminal myristoyl group of Abl 1b, to the hydrophobic pocket in the C-terminal lobe of kinase domain induces a 90^{o} bending of the C-terminal α-I helix of kinase domain. This creates a docking site for SH2 domain, thereby maintaining Abl in an inhibited state.

This assembly does not form in Bcr-Abl as it lacks N-terminal myristoyl group. Autophosphorylation of tyrosine residues in kinase domain also disrupts the intramolecular interaction between SH3 and SH2 domains [15, 17, 18].

1.1.4.2. Bcr-Abl downstream signalling

Pro-proliferation and anti-apoptotic signalling

Activation of oncogenic signalling pathways by Bcr-Abl is mediated not just by phosphorylating Tyr-residues of downstream substrates but also through autophosphorylation of Tyr-residues within Bcr-region of Bcr-Abl [21]. Autophosphorylation on Y328 and Y360 in Ser/Thr kinase domain of Bcr inhibits its Ser/Thr kinase activity [25]. Hence Bcr-Abl has only tyrosine kinase activity. The crucial autophosphorylation event occurs on Y177 within Ser/Thr kinase domain of Bcr. This creates binding site for the SH2 domain containing adapter protein Grb2, which activates diverse signalling pathways (Fig.1.6) [17, 19]. The importance of phosphorylation of Y177 in Bcr-Abl mediated oncogenesis was further emphasized by mutational studies where mutation of Y177 lead to reduced transformation potential due to loss of Grb2 binding site [20].

One of the major pathways activated by Bcr-Abl through Grb2 is the Ras signalling pathway. Ras activates downstream kinases Raf, MEK, ERK and in-turn stabilizes c-myc, which is upregulated in CML [20]. Ras signalling is also activated by two other adapter proteins CrkL and Shc, which bind to proline rich sequence and SH2 domain of Abl respectively. Studies with dominant negative forms of Ras, Grb2 or Raf have shown that, inability to activate this pathway inhibits Bcr-Abl induced transformation, indicating that Ras signalling is the key mitogenic pathway required for Bcr-Abl mediated oncogenesis. Ras signalling also plays a role in mediating growth factor independent survival, through its anti-apoptotic effect [22].

Another key pathway activated by Bcr-Abl via Grb2 is PI3K signalling pathway. Grb2 mediates constitutive activation of PI3K/Akt signalling by activating the scaffold protein Grb2 associated

binding protein 2 (GAB2) [20]. PI3K signalling is also activated by Shc, CrkL and Cbl proteins. This pathway activates Ser/Thr kinases Akt and p70S6K. Akt exhibits anti-apoptotic effect by phosphorylating pro-apoptotic protein 'Bad' and preventing it from binding to anti-apoptotic protein Bcl-xL, while p70S6K promote cell proliferation [18, 22].

Under physiological conditions, activation of Abl by growth factors like GM-CSF or IL-3 results in activation of JAK-STAT pathway. In CML on the contrary, STAT5 being one of the first downstream effectors of Bcr-Abl [17] is constitutively phosphorylated and STAT1 to a lesser extent, thereby mimicking the growth factor stimuli, resulting in direct activation of JAK-STAT signalling pathway. Bcr-Abl also stimulates JAK2 protein. It was found that JAK-STAT signalling is also essential for Bcr-Abl mediated transformation, as use of mutant inactive JAK2 or STAT5 diminished the oncogenic potential of Bcr-Abl. The role of activated STAT5 in Ph+ cells primarily seems anti-apoptotic, due to its involvement in transcriptional activation of anti-apoptotic protein Bcl-xL [18, 20]. Thus activation of Ras, PI3K and JAK-STAT signalling by Bcr-Abl could be the key for growth factor independence in CML cells [18].

Cell adhesion signalling

The key mediator of adhesion of hematopoietic progenitor cells to bone marrow stroma and extracellular matrix (ECM) is β1-integrin. The affinity of β1-integrin towards extracellular ligands for adhesion is regulated by inside-out signalling, in which binding of focal adhesion proteins to cytoplasmic domain of β1-integrin brings about conformational change in its extracellular domain, thereby increasing ligand binding affinity [27]. This inside-out signalling also exhibits a negative feedback regulation on cell proliferation. For instance, it leads to upregulation of $p27^{Kip-1}$ which inhibits cdk2, thereby preventing the cells from entering cell cycle. In CML, Bcr-Abl interacts with proteins involved in β1 integrin signalling like FAK, paxillin, vinculin and talin and prevent their interaction with cytoplasmic domain of β1 integrin, thereby inhibiting inside-out signalling and in turn efficient adhesion. Studies have also shown that, Ph^+ cells express an adhesion inhibitory variant

of β1 integrin with unaffected outside-in signalling but impaired inside-out signalling and proliferation inhibitory signals. CML is thus characterized by pre-mature egression of hematopoietic progenitor cells which leads to increase in number of blast cells in circulation [18, 20].

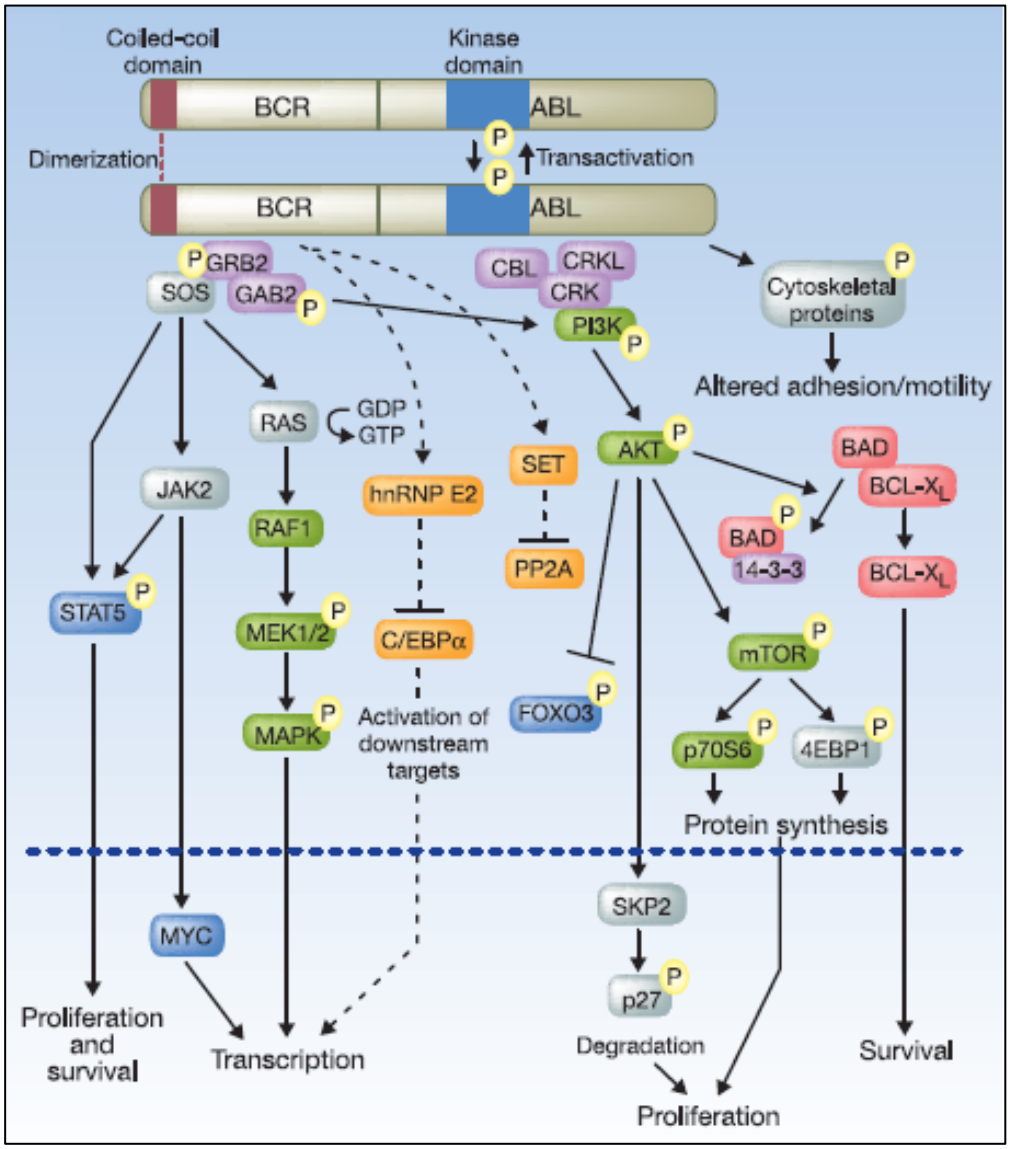


Fig. 1.6. Bcr-Abl downstream signalling. *Autophosphorylation of tyrosine residues within Bcr-region along with phosphorylation of downstream substrates by Bcr-Abl leads to constitutive activation of Ras, PI3K and JAK-STAT signalling pathways thereby mediating transformation. Adapted from [26].*

1.2. CML THERAPY: TARGETING BCR-ABL

In the past decade, CML therapeutics has undergone a significant advancement. The success of therapies targeting Bcr-Abl has established CML as an epitome of targeted therapy. The success of treatment response is determined by the ability and duration to attain the following response criteria:

- **Complete hematological remission (CHR)** is defined by normalization of blood counts and lack of CML-related symptoms.
- **Cytogenetic remission (CyR)** is graded based on the percentage of Ph^+ metaphases determined by FISH. The remission is termed minor if the Ph^+ metaphases are > 35%, major (MCyR) if Ph^+ metaphases are less than 35% and complete (CCyR) if no Ph^+ metaphases could be detected.
- **Major molecular remission (MMR)** is characterized by the presence of < 0.1% of detectable *Bcr-Abl* mRNA transcript [28, 29].

CML therapies could barely reach haematological remission before 2001. Till mid 19th century CML was treated using Fowler's solution which is 1% arsenic trioxide. With discovery of X-Rays, radiation therapy was used to alleviate the symptoms. With the advent of chemotherapy, these treatment modalities were replaced by busulfan and hydroxyurea, which for many decades remained the key therapeutic agents for CML [30]. These cytotoxic drugs inhibited the growth of both normal and leukemic cells but failed to eradicate Ph^+ clones. As a result, they could induce hematological remission in patients but were unsuccessful in altering the natural course of the disease as patients eventually progressed to BC, although the rate of progression was marginally reduced. Both these drugs had associated toxicities. While busulfan was known to cause severe myelosuppression in ~10% patients, hydroxyurea was comparatively well tolerated with fewer side effects [11, 16, 29, 31]. Interferon-α (IFN-α) treatment introduced in 1980s was found to be more effective than the earlier chemotherapeutic drugs for treating the chronic phase of CML. It could bring about hematological remission in majority of CP patients and for the first time complete cytogenetic remission in about

10-30% patients in CP, thereby acutely reducing the disease progression rate and improving survival [11, 14, 30]. IFN-α was found to act by restoring the normal adhesion function of β1 integrin [32]. The major drawback with IFN-α treatment is the adverse toxicity leading to intolerance [11]. Bcr-Abl tyrosine kinase activity, which drives the transformation, was not inhibited by any of these drugs. Allogenic stem cell transplant is the only curative treatment available for CML but is highly restricted due to difficulty in finding compatible donors, age and associated mortality and morbidity and high treatment cost [30, 33].

Post-2001 began the era of targeted therapy. Bcr-Abl tyrosine kinase activity inhibition has been the most successful therapeutic strategy in CML. Tyrosine kinase inhibitor (TKI) therapy using Imatinib mesylate (IM), an ATP-mimetic belonging to the class of 2-phenyl aminopyrimidine which acts by inhibiting the Bcr-Abl tyrosine kinase activity, was considered the 'gold standard' therapy [11, 34].

1.2.1. IM- MODE OF ACTION

The kinase domain of Abl consist of –

- A phosphate binding loop (**P-loop**), which is rich in glycine and critical for co-ordination of ATP -Mg^{2+} complex, located in the N-terminal lobe.
- A peptide **substrate binding site** at the C-terminal lobe.
- An **ATP-binding site** in the cleft between both the lobes.
- The most flexible activation loop (**A-loop**), which is modulates the switch between active and inactive conformation of the kinase.

The A-loop consists of a conserved 'DFG motif' (Asp 381 – Phe 382 – Gly 383) at its N-terminus and a central Tyr-393 (Y-393) residue. A single phosphorylation at Y-393 leads to an electrostatic interaction with the neighbouring arginine residue and stabilizes the extended – open form of A-loop, which provides access to the peptide substrate for binding, thereby resulting in assumption of active

conformation of Abl. In this active state, the DFG motif is oriented towards the ATP binding site ('DFG in' conformation), where it co-ordinates with Mg^{2+} ion, whose catalytic activity is essential for phosphorylation. Further, C-terminal portion of the loop forms a platform to facilitate substrate binding. When Y-393 is not phosphorylated, the A-loop folds in towards the ATP binding site and occludes substrate binding, which results in an inactive conformation. Also, aspartate residue of DFG motif rotates away from the active site and assumes a 'DFG out' conformation, which prevents Mg^{2+} ion from co-ordinating with ATP phosphate at the active site [15, 35-37].

IM is a competitive inhibitor of ATP and inhibits Bcr-Abl kinase activity by binding to its ATP-binding site. IM binds to Bcr-Abl in its inactive (DFG out) conformation. Though Bcr-Abl is constitutively active, after each substrate phosphorylation step the conformation changes to inactive form due to dephosphorylation of Tyr 393 in A-loop and transient protonation of Asp 363 in the catalytic site. IM binds to inactive conformation of Bcr-Abl during this transient period and prevents the switch to active conformation [36, 38]. Though the active conformation is conserved among various kinases, their inactive conformation differs significantly. Even closely related kinases like Abl and Src family kinases have quite distinct inactive conformation. Hence IM is highly specific to Abl and inactive against most tyrosine kinases with an exception of PDGFR and c-kit and all other Ser/Thr kinases [34].

1.2.2. IM– RESPONSE RATE

IM could not only induce haematological response in 96% patients in chronic phase but also bring about a remarkable complete cytogenetic remission of 86% and 10 year major molecular remission in about 93% of newly diagnosed CP patients [13, 34, 39]. The associated side effects are quite mild compared to IFN-α [14].

However, IM fails to eradicate quiescent leukemic stem cells (LSC) which could eventually lead to relapse upon treatment discontinuation [40]. Further, IM is not very effective against patients in advanced stage of CML. A study with 484 CML-BC patients revealed that IM could induce hematological remission in only 50-70% patients and cytogenetic response in 12-17% patients, with 1-year survival rate of 22-36%. Despite tremendous success in treating CML-CP patients, the median survival of BC patients treated with IM was improved to be only 7-11 months as against 3-4 months in pre-IM era [41]. This lack of response is attributed to development of primary and secondary resistance to IM. Primary resistance is defined as the failure to achieve time dependent endpoints of complete haematological, cytogenetic and molecular remission upon initiation of TKI therapy, while secondary (acquired) resistance involves gradual loss of response during the course of treatment [28]. The rate of imatinib resistance increases with disease progression, with primary and secondary resistance being 9% and 17% respectively in CP, 24% and 51% in AP and 66% and 88% in BC [11].

1.2.3. MECHANISM OF RESISTANCE TO IM: BCR-ABL DEPENDENT MECHANISMS

Genetic alterations in *BCR-ABL,* leading to increased expression or kinase domain mutation, constitute Bcr-Abl dependent resistant mechanisms.

Increased *BCR-ABL* expression

BCR-ABL gene amplification or presence of double Ph^+ chromosome causes increase in *BCR-ABL* expression and protein level, due to which, the amount of IM administered is not sufficient to inhibit Bcr-Abl activity. As a result, the Bcr-Abl activity persists despite presence of TKI, leading to resistance. This resistant mechanism is less common in patients and is less challenging than other resistant mechanisms [11, 28].

Kinase domain mutations

Kinase domain mutations are one of the most common cause for IM resistance and accounts for 50% of TKI resistance observed in patients [42]. There are about 90-point mutations identified in Bcr-Abl

kinase domain [11] and not all of those mutations lead to resistance. Only those mutations which either affect the amino acids whose interaction with IM is vital for its binding or affect the ability of Bcr-Abl to attain inactive conformation, thereby preventing IM binding, leads to resistance [30, 37]. The kinase domain mutation rate increases with disease progression and is more common in advanced phases (AP and BC). One of the common mutations found in about 2-20% of CML resistant cases, with very poor prognosis, occurs at the 'gatekeeper residue' T315 wherein threonine is replaced with isoleucine (T315I) [36]. Threonine 315 is called a gatekeeper residue due to its strategic positioning, which controls the accessibility to ATP binding pocket. Its '-OH' group forms a hydrogen bond with IM and its side chain assists binding of IM to the hydrophobic region adjacent to ATP-binding site. The gatekeeper residue is thus indispensable for IM binding but not for ATP binding [15]. Presence of isoleucine in place of threonine increases the bulkiness and thus poses steric hindrance to binding of IM [28, 30]. Mutations in ATP binding P-loop (Q252R/H, Y253F/H, E255K/V, M244V and G250E) accounts for 36-48% of all mutations and is also associated with poor prognosis. P-loop mutations destabilize the loop conformation and arrangement of residues essential to attain the inactive conformation, thus preventing conformational change required for IM binding [11, 28, 37]. Mutations are also found in A-loop (H396R) which stabilizes the kinase in its active conformation and in C-lobe (M351T, E355G, F359 V/C/I and F486S) which disturbs the relative orientation of N and C-lobe for IM binding [37].

1.2.4. MECHANISM OF RESISTANCE TO IM: BCR-ABL INDEPENDENT MECHANISMS

About 50% IM- resistant patients do not harbour any kinase domain mutation or express increased amount of Bcr-Abl [43]. The cause for resistance in such patients is attributed to Bcr-Abl independent mechanisms, which consist of altered drug transport and activation of alternate signalling pathway.

Altered drug transport

The ability of IM to inhibit Bcr-Abl tyrosine kinase activity is highly dependent on its intracellular concentration. Reduction of intracellular IM concentration due to increased drug efflux or decreased drug influx is found to be associated with IM resistance. ABCB1, also called P-glycoprotein or MDR1, is an ATP-dependent multidrug efflux protein belonging to the family of ABC transporters, mediate efflux of IM. Overexpression of ABCB1 is observed in CML-BC patients. Similarly reduction in levels of IM influx protein OCT1 (organic cation transporter 1) in CML patients has been correlated with poor IM response [11, 33], while high level of hOCT1 in patients is predictive of improved MMR rate and overall survival [28].

Activation of alternate signalling pathways

In about 40% patients, resistance is observed despite inhibition of Bcr-Abl tyrosine kinase activity by Imatinib [28]. As CML progresses to blast crisis, its dependence on Bcr-Abl for oncogenic signalling reduces. This could be due to genetic alterations which lead to either activation of oncogenic signalling pathways other than BCR/ABL [15] or those which can activate downstream components of BCR/ABL pathway despite inhibition of its activity. It is also proposed that the residual, persistent LSCs mediate activation of survival signals in the microenvironment, stimulating Bcr-Abl independent survival [40]. There have been quite a few studies on identifying the alternate signalling pathways that confer resistance. Donato *et al.*, have demonstrated that in an IM-resistant cell line, whose Bcr-Abl activity is inhibited, resistance is conferred by overexpression of Src family kinase Lyn through activation of Bcr-Abl downstream components in a Bcr-Abl independent manner. Overexpression of Lyn kinase was also observed in resistant patients [44, 45]. Bcr-Abl independent activation of STAT3 observed in primary cells from resistant CML patients, was to mediate survival of TKI resistant cell lines. Bcr-Abl independent activation of PI3K/Akt/mTOR and upregulation transcription factor FOXO1 were observed in some resistant patients without any kinase domain mutations [28]. Green *et al.*, identified that activation of RAF-MEK-ERK pathway by PKCη leads to development of Bcr-Abl independent resistance in a CML-BC cell line [43].

1.2.5. ATTEMPTS TO OVERCOME IM RESISTANCE

IM dose escalation:

Resistance due to increased Bcr-Abl levels could be overcome by increasing the dosage of IM. This is beneficial to a fraction of patients who had initially achieved cytogenetic response with standard IM dose but eventually resulted in cytogenetic failure, due to Bcr-Abl gene amplification with no kinase domain mutation. High dose of IM was found to be well tolerated by patients [46]. Some chronic phase patients with resistance due to altered drug transport are also found to be benefited by IM dose escalation [47].

Second and third generation TKIs:

To deal with resistance posed by kinase domain mutations, 2nd and 3rd generation TKIs were developed which are structural variants of IM that differ in their amino acid requirements within kinase domain for binding. Second generation TKIs include dasatinib, nilotinib and bosutinib. Unlike IM, dasatinib and bosutinib could bind to Bcr-Abl kinase domain both in its active as well as inactive conformation. As a result, they are dual specificity inhibitors, which inhibit Src family kinases along with Abl. Nilotinib on the other hand is a structural analogue of IM and a more specific inhibitor of Abl, which binds and stabilizes its inactive conformation [48]. Dasatinib is about 325-fold more potent towards wild-type Bcr-Abl than IM while Nilotinib has 50-fold higher potency than IM in-vitro. Dasatinib and nilotinib are used as second-line therapy for patients who do not respond to IM. Many countries use dasatinib and nilotinib for first-line treatment just like IM. Bosutinib which is approved only as a third-line therapy, is used to treat patients who fail to respond to any of the other three TKIs [30, 48]. Dasatinib, nilotinib and bosutinib are effective against many kinase domain mutations with the exception of T315I mutation, as all three 2nd generation TKIS retain their dependency on hydrogen bonding with Thr315 for binding [49, 50].

Third generation TKI includes ponatinib, which like IM binds to inactive conformation and is 520 times more potent than IM in-vitro [51]. It is the only available TKI that is effective against T315I as

it doesn't require hydrogen bond with Thr315 for binding. It has a long flexible ethynyl tricarbon linker, which despite the presence of bulky sidechain of isoleucine, facilitate its accommodation in catalytic domain [30]. Ponatinib was found to be effective against all other kinase domain mutations as well and seem to fail only in advanced phase CML patients harbouring multiple kinase domain mutations[28]. The major drawback with ponatinib is its adverse side effects which include vascular occlusion, hepatoxicity, arterial and venous thrombosis and heart failure. As a result, its use involves a huge risk-benefit trade off and is highly restricted to patients harbouring T315I mutation, who fail to respond to any other available treatment options [15, 29].

Apart from being effective against kinase domain mutations, next generation TKIs could also overcome Bcr-Abl independent resistance to an extent. Both nilotinib and ponatinib doesn't depend on hOCT1 and P-glycoprotein for their cellular import and export respectively and hence can efficiently overcome resistance due to drug transporters. Dasatinib on the other hand, though is independent of hOCT1 for import, is a substrate for P-glycoprotein and is thus effective only against patients with low hOCT1 level [28, 36, 47, 52]. Dasatinib and bosutinib are effective against patients with resistance due to activation of alternate signalling pathways by Src family kinases, as they effectively inhibit Src family kinases along with Abl [30, 48].

1.2.6. CLINICAL SUCCESS IN OVERCOMING RESISTANCE

A considerable proportion of CML-CP patients, with cytogenetic failure to standard IM dosage, were found to respond effectively to increased IM dosage. CCyR was observed in 40% patients. The response was found durable with 88% patients achieving MCyR with sustained response for at least 2 years [46]. IM resistant patients who fail to respond to increased IM dose, due to Bcr-Abl kinase domain mutations or reduced intracellular IM concentration respond to second and third generation TKIs. IM-resistant CML-CP patients treated with dasatinib as second line treatment, resulted in CCyR in 44-53% patients and MMR in 29-43% patients, while those treated with nilotinib showed CCyR

in 31-45% patients and MMR in 28% patients. Bosutinib though has been approved only as a third line therapy after failure to all 3 TKIs, its use following IM failure has shown to induce CCyR in 41-48% patients and MMR in 64% patients. [53]. The phase 2 trial of ponatinib for patients in chronic phase, who failed to respond to 3 or more TKIs, reported MCyR in 51% patients resistant to dasatinib or nilotinib and in 70% patients with T315I mutation [54]. Use of next generation TKIs could thus significantly improve cytogenetic and molecular response rates in resistant CML-CP patients [29]. Like IM, dasatinib and nilotinib have also been approved as first line therapy for CML-CP patients. Both these TKIs help in achieving molecular response faster than IM but no significant difference was observed in 5 year overall survival rate [53].

Progression to advance stages of CML (AP and BC), leads to suboptimal response not just to IM but to other TKIs as well [55]. 5 studies on 484 BC patients treated with IM, showed the following response rates: CHR – 50-70%, MCR - 12-17% and a median survival of 6.5-10 months. The response rate from 3 studies, using dasatinib for treatment of 400 BC patients pre-treated with IM, was CHR in 33-61% patients, MCyR in 35-56% patients and a median survival rate of 8-11 months. The response rate of 169 IM-resistant BC patients treated with nilotinib, from 2 different studies, was found to be CHR in 60% patients, MCyR in 38-52% patients and a median survival of about 10 months [41]. Ponatinib was found to induce MHR in 55% patients and MCyR in 39% patients in accelerated phase while MHR in 31% and MCyR in 23% patients in blast crisis [54]. However, the response is transient in BC patients [56]. Further, PACE (Ponatinib Ph^+ ALL and CML evaluation) trial reported that only 27% patients resistant to multiple TKIs, without any kinase domain mutation, could achieve MMR with ponatinib [57] .

1.3. NEED FOR ALTERNATE THERAPEUTIC TARGETS

The steps taken to overcome resistance, like development of next generation TKIs, though have increased cytogenetic response, provide no survival advantage over IM to resistant CML-BC patients,

with their median survival rate restricted to a maximum of 11 months [41]. Currently there is also a rising concern about the emerging clinical reports of compound mutations (multiple kinase domain mutations) in patients who underwent sequential treatment with different TKIs. Even ponatinib, the only TKI that is effective against almost all known kinase domain mutations, is ineffective against T315I-inclusive compound mutations. PACE trial reported that the frequency of compound mutations is more in BC than in CP, thereby indicating increased risk of ponatinib resistance in advanced phase CML patients [56]. Further, significant proportion of BC patients, without any kinase domain mutation, are refractory to all five available TKIs. This is attributed to the additional complex cytogenetic changes in BC, which leads to activation of signalling pathways, unrestrained by TKIs [55]. Thus, despite all attempts, TKI failure is still a major challenge for blast crisis patients. This also is a matter of concern in Indian population with unique features of CML. While in western countries most patients are asymptomatic at presentation and diagnosed during routine blood test or physical examination [13] in India most patients when presented at clinic are symptomatic, with high risk disease and thus more at a risk of being resistant to TKIs. Further the incidence age in Indian CML patients (32-42 years) is about a decade earlier than in the West (55years) [58, 59]. TKI therapy is not curative but is a maintenance therapy. Thus, the younger population with disease will be on treatment for a longer duration and thus at higher risk of developing secondary resistance to TKIs.

CML accounts for 15% of all adult leukemias in the west but is the most common adult leukemia in India (30-60% of all adult leukemia) [13] and considering the population in India, number of patients with resistance to TKIs would be significantly high. It is therefore imperative to conduct a comprehensive molecular analysis of TKI-resistant CML-BC, so as to identify additional therapeutic targets. Hence, we hypothesize that in CML-BC, where signalling pathways other than those mediated by Bcr-Abl drive the oncogenic signalling, identifying and targeting proteins of this alternate pathway that provide survival advantage in a Bcr-Abl independent manner would be the ideal therapeutic strategy to overcome resistance. In CML-BC wherein molecular alterations activate

downstream components of Bcr-Abl, thereby keeping the pathway active even if Bcr-Abl activity is inhibited, identifying and targeting components from the BCR/ABL pathway would help in managing resistance. These strategies also circumvent the problem of developing additional TKIs as kinase domain inhibitors are majorly limited by their inherent lack of specificity. Delineation of Bcr-Abl pathway can also serve to find therapeutic targets for non-responders to TKI in CML-CP wherein *BCR-ABL* is still the driver mutation.

1.3.1. INVESTIGATORY DRUGS IN CLINICAL TRIAL

Based on our understanding of Bcr-Abl functional domains and signalling pathways activated, some drugs have thus entered clinical trial for patients unresponsive to TKIs.

- **Bcr-Abl allosteric inhibitor:**

Asciminib (ABL001) is a novel allosteric inhibitor of Bcr-Abl targeting the myristoyl pocket and is the most anticipated new therapy. It mimics the negative regulation of Abl by N-terminal myristoyl group, which is lost in Bcr-Abl, thereby restoring its auto-inhibition. As the efficacy of allosteric inhibitor is unaffected by the mutations in kinase domain including T315I, this drug has entered Phase I clinical trial for CML-CP and AP patients unresponsive or intolerant to 2 or more TKIs, wherein it is being investigated as a single agent as well as in combination of TKIs.

- **Omacetaxine:**

This is a protein synthesis inhibitor which blocks the translational elongation process by binding to ribosome aminoacyl-t-RNA acceptor site. Omacetaxine has been approved by FDA (only in the US) directly after phase II trial, with strict guidelines, selectively for CML-CP and AP patients resistant or intolerant to 2 or more TKIs [29, 60, 61].

- **Hsp-90 inhibitor:**

Hsp-90 interacts with Bcr-Abl and stabilizes it. Increased Hsp-90 level in CML was found to prevent degradation of Bcr-Abl. Hsp-90 inhibitors like paclitaxel, 17-AAG (Tanespimycin) and STA-9090 (Ganetespib) prevents binding of Hsp-90 to Bcr-Abl, thereby promoting degradation of Bcr-Abl via ubiquitin proteasomal pathway [49, 62]. 17-AAG has low efficacy as a monotherapy but increases the rate of apoptosis when used in combination with HDAC inhibitors or imatinib in-vitro and in-vivo. STA-9090 is more potent than 17-AAD in reducing CML cell proliferation of a single agent. Both these have thus entered clinical trial for treating patients with advance phase of disease or relapse [60].

- **HDAC inhibitors:**

Histone deacetylase (HDACs) are found to be overexpressed in several cancers [60]. HDAC inhibitors like valproic acid, pracinostat and vorinostat help in generating hyperacetylated histones which causes cell cycle arrest and induce apoptosis in tumor cells by modulating the expression of various cell cycle regulators including p21 and p27 [49]. Though there is no direct evidence of involvement of HDACs in Bcr-Abl, studies evaluating HDAC inhibitors in CML cell lines and mouse models provided encouraging results in overcoming TKI resistance, in combination with TKIs. Hence these have entered clinical trials as combination therapy with TKIs [60].

- **Inhibitors of MAPK signalling pathway:**

RAS-MAPK signalling is one of the key pathways activated downstream of Bcr-Abl, essential for transformation. Hence inhibitors of this pathway namely farnesyl transferase inhibitors (FT-Is), Raf-1 inhibitor and MEK inhibitor have gained attention. Farnesyl transferase activity is essential for membrane migration of proteins like RAS and their activation. FT-Is tipifarnib and lonafarnib thus act by inhibiting RAS signalling pathway and demonstrated promising anti-leukemic activity against CML in-vitro and in-vivo, which led to their phase I clinical trial. Both these drugs, as monotherapy were found less beneficial in clinical setting but proved useful in combination with IM [49, 60]. Raf-1 inhibitor sorafenib is an FDA approved drug for hepatocellular and renal carcinoma. Following the

in-vitro demonstration of its ability to induce apoptosis in IM-resistant cells, sorafenib has entered phase II clinical trial as a single agent for CML-CP patients resistant to IM [49]. Identification of Bcr-Abl independent activation of RAF/MEK/ERK pathway by Green et.al., [43] led to investigation of MEK inhibitor trametinib for Bcr-Abl independent resistance. Promising in-vivo results has led to its clinical trial for CML patients resistant to TKIs [49].

- **mTOR inhibitors:**

mTOR is a Ser/Thr kinase that mediates PI3K/Akt pathway, which is another key Bcr-Abl downstream signalling pathway. mTOR activation has also been demonstrated as the cause for Bcr-Abl independent resistance in-vitro [57]. Hence number of mTOR inhibitors are in clinical trial to overcome resistance in CML. Since till date there has been no data available for mTOR inhibitors on CML patients, their efficacy remains to be established. Rapamycin which inhibited cell growth in Ph^+ cell lines with or without T315I mutation and found to be effective against IM-resistant patients in a pilot study, has entered phase I/II clinical in combination with cytarabine or etoposide for AP and BC patients. The therapeutic efficacy of everolimus is being evaluated as a single agent against CML-BC patients or in combination with IM for CML-CP patients in two different phase I/II clinical trials. BEZ235, a dual PI3K-mTOR inhibitor and temsirolimus in combination with IM are in phase I clinical trials for AP and BP patients respectively [49] [60].

- **JAK2 inhibitor:**

JAK-STAT signalling is another important pathway constitutively activated by Bcr-Abl. JAK2 inhibitor was assessed for their ability to inhibit Bcr-Abl downstream signalling. Following the success of JAK 2 inhibitor Ruxolinitib, in combination with TKIs to restore TKI sensitivity in resistant cells in-vitro and its promising phase I clinical trial with nilotinib that resulted in molecular remission in patients, many such phase I/II clinical trials are now being carried out with patients in different phases of CML [60].

- **Aurora kinase inhibitors:**

Aurora kinase (AURK) is a Ser/Thr kinase. Though the correlation between Bcr-Abl and AURK in CML progression is quite unclear, its inhibitors are investigated in clinical trials. Tozasetib could potentially revert patients in advanced stage CML with T315I mutation to chronic phase. This led to its phase II clinical trial which brought about hematological and cytogenetic response in advanced phase CML patients. Danusertib is a dual AURK/Abl inhibitor effective against T315I mutation. In phase I study this drug could induce modest response in T315I mediated TKI-resistant AP and BC patients [28, 60].

Since approval of imatinib by FDA for treating CML in 2001, so far, no non-TKI based drug has been approved for treatment in resistant CML with the exception of omacetaxine which has been approved only in the US. Based on our limited understanding of Bcr-Abl dependent signalling pathways and very little knowledge on Bcr-Abl independent pathways that confer resistance, drugs targeting few signalling molecules have entered clinical trial. Recent genetic studies have shown that Bcr-Abl independent resistant mechanisms are highly heterogenous among patients [57], indicating that a significant fraction of TKI-resistant advanced phase patients would still be unresponsive to the above drugs due to existence of unidentified Bcr-Abl independent resistant mechanisms. Further, though these drugs provided promising results in in-vitro and in-vivo setting, their efficiency against TKI-resistant CML-BC patients is unclear and many drugs fail in phase II clinical trial due to previously unknown side effects and inefficient efficacy [63]. Hence, there is a dire need to keep the quest for identifying more therapeutic targets for TKI-resistant BC active, by gaining in-depth understanding of Bcr-Abl downstream and alternate signalling pathways.

1.4. PROTEOMIC APPROACH TO IDENTIFY ALTERNATE THERAPEUTIC TARGETS

The above-mentioned therapeutic targets were selected from hypothesis-driven investigation wherein prior knowledge of involvement of the protein in drug resistance existed. Limitation of the hypothesis-driven approach is the inability to discover novel molecular alterations. 'OMICS' analysis provides a systematic, global and unbiased approach to identify and understand molecular alterations leading to a diseased state. Since proteins are functional molecules, changes in their level and activity directly influence the normal functioning of cells. Hence, proteomics is the approach of choice to discover novel biomarkers or therapeutic targets [64].

1.4.1. EXPERIMENTAL STRATEGIES FOR QUANTITATIVE PROTEOMICS

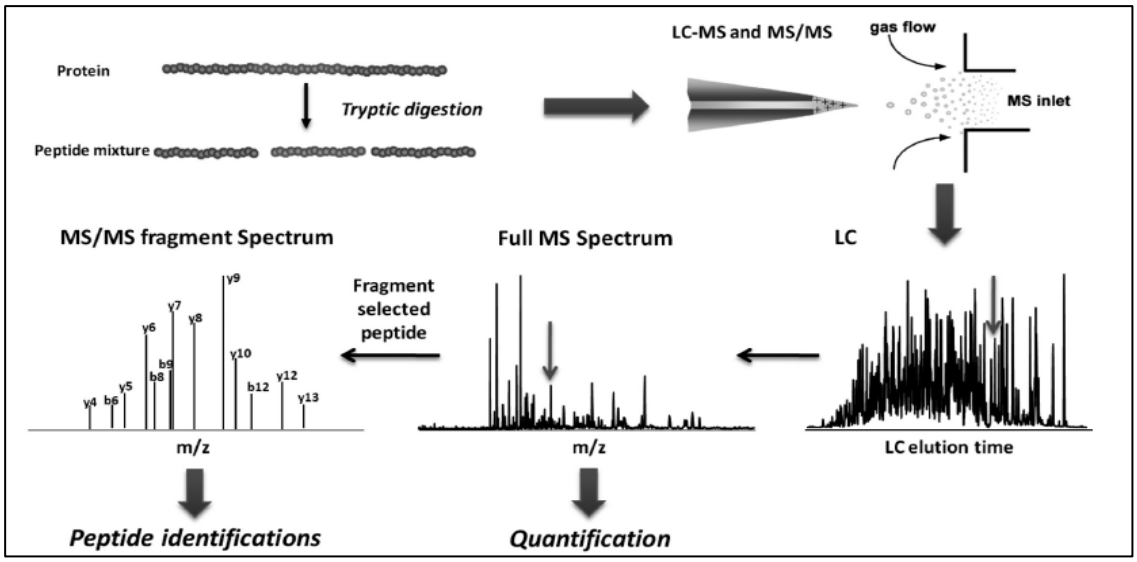


Fig.1.7. LC-MS based proteomic analysis workflow. *Proteins from complex biological mixture are digested into peptides using trypsin. The peptides are injected into LC-MS where they are separated in LC and ionised by electrospray ionisation before entering the mass spectrometer, wherein the MS and MS-MS spectra are generated which help in protein identification by searching against protein sequence database and quantification. Adapted from [65].*

LC-MS based proteomic workflow is shown in Fig. 1.7, wherein proteins are digested to peptides for identification and quantification [65]. LC-MS based quantitative proteomic profiling could be carried out using labelled approaches like stable isotope labelling with amino acids in culture (SILAC), isobaric tag for relative and absolute quantification (iTRAQ), tandem mass tag (TMT) and isotope coded affinity tag (ICAT) or label-free approaches like sequential window acquisition of all theoretical fragment-ion spectra (SWATH). Labelled approaches use data dependent acquisition strategy (DDA) which involves stochastic selection of precursors from MS1 survey scan for fragmentation[66] while label-free approach utilize data independent acquisition (DIA) wherein all precursors within a specified mass range are subjected to fragmentation [67].

1.4.1.1 Labelled quantification by iTRAQ

Quantitative proteomic profiling by iTRAQ involves labelling trypsin digested peptides from the comparison groups with distinct isobaric tags that bind to primary amines. 4-plex iTRAQ experiment facilitates comparison for four different samples while 8-plex facilitates comparing eight different samples. In a 4-plex, each tag contains a unique charged reporter group with mass ranging from 114 to 117 Da, a neutral balance group with mass ranging from 31 to 28 Da respectively so as to maintain a total mass of 145 Da thereby making the tags isobaric and an amine-specific peptide reactive group (Fig. 1.8a). The labelled peptides from the comparison groups were pooled and subjected to LC-MS/MS analysis, wherein in MS analysis each identical peptide from all the comparison groups appear as a single precursor, owing to the overall same m/z maintained in all tags by the balance group. However, fragmentation in MS/MS along with generating strong y- and b-ions from peptides that help in identification through sequence database searching, also breaks off the iTRAQ reporter ions from the balance group resulting in distinct ions with m/z 114, 115, 116 & 117. The relative intensities of these reporter ions are proportional to the relative abundances of each peptide in the corresponding comparison group (Fig. 1.8b). The reporter ion signals thus provide relative quantification of the peptides [68].

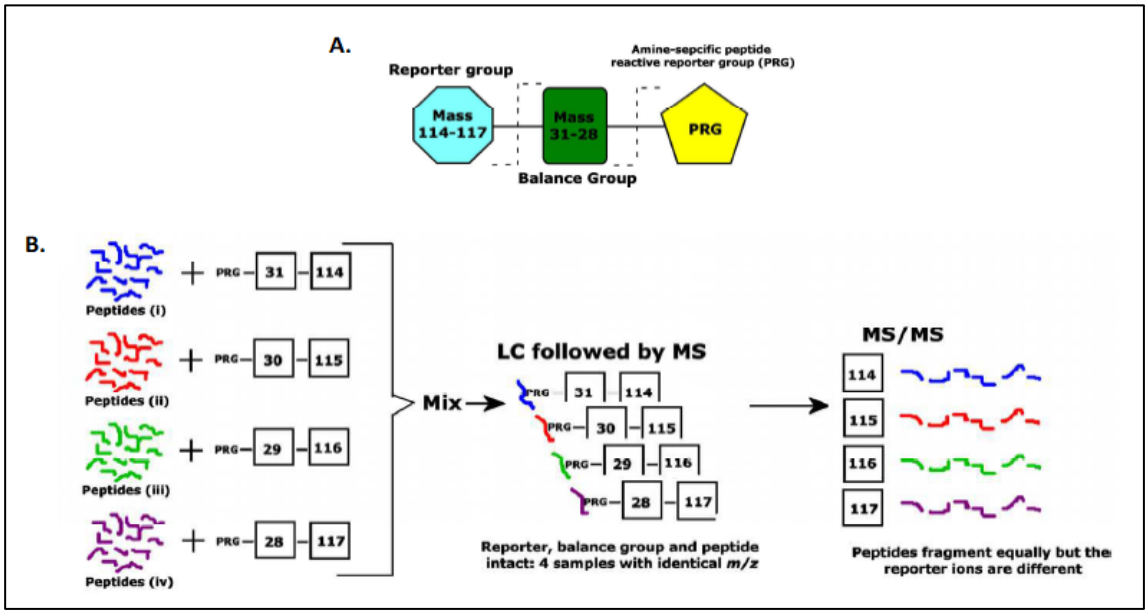


Fig.1.8. Quantitative proteomic profiling by iTRAQ*. A. iTRAQ tag containing a reporter group, balance group that help in maintaining the tag isobaric and peptide reactive group which help in binding to the peptides via amine group. B. Comparative groups labelled with different iTRAQ tags were pooled and subjected to LC-MS/MS analysis wherein during MS/MS fragmentation the reporter group separates from the balance group, whose intensity is proportional to the relative abundance of the peptide. Adapted from [68].*

1.4.1.2. Label-free quantification by SWATH

In SWATH-MS based label-free quantification, trypsin digested peptides of each sample from the comparison group is run individually in LC-MS. SWATH-MS provides in depth profiling through a cyclic acquisition throughout the LC retention time, with each cycle consisting of a MS1 survey scan which detects all peptide precursors eluting at a given time point, followed by generation of a series of MS2 fragment ion spectra, wherein the entire peptide mass range in the first quadrupole (Q1) (400-1200 m/z) is divided into multiple wider precursor isolation windows (mostly 25Da wide) and all precursor ions within each window are subjected to fragmentation sequentially in the collision cell (Q2), in a systematic and unbiased manner [67].

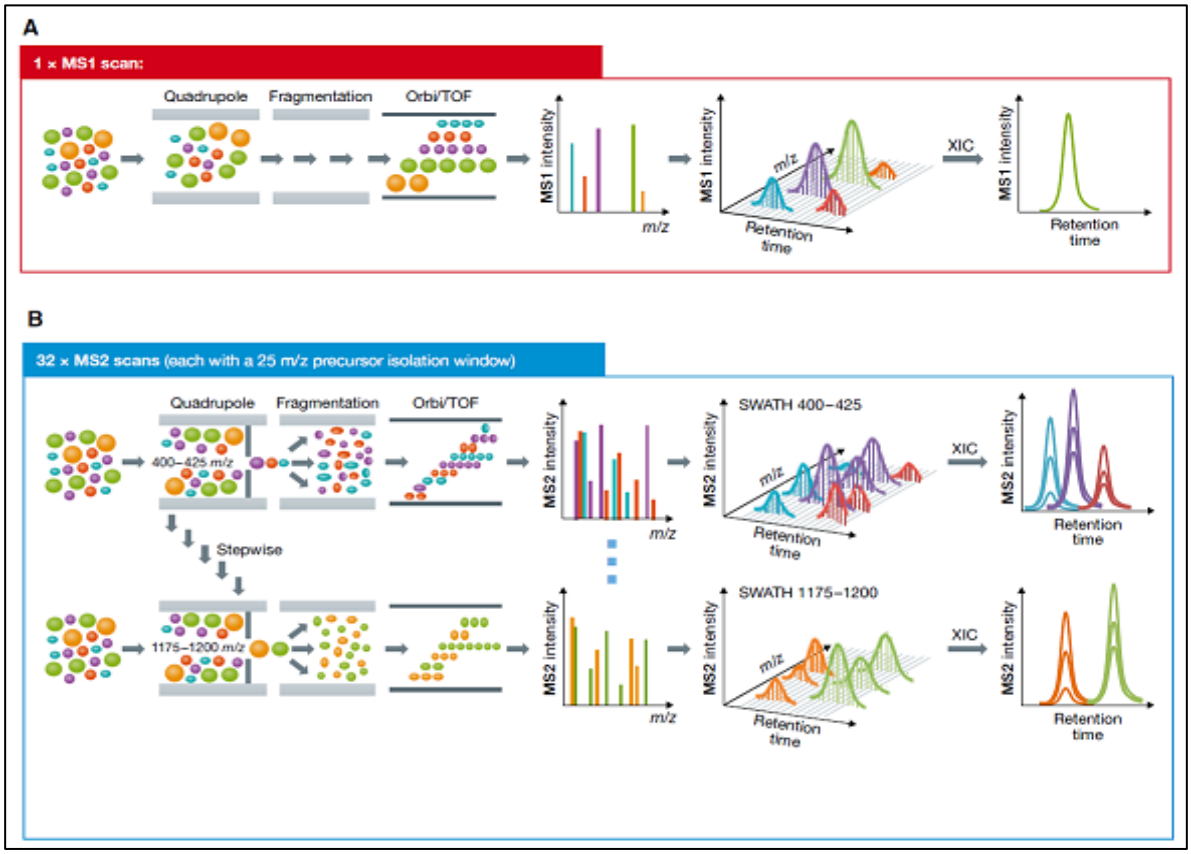


Fig. 1.9. Principle of SWATH-MS. *SWATH-MS involves (A) generation of one MS1 spectrum that scans all peptide precursors eluting at a given time point. (B) For generation of a comprehensive MS/MS spectra, the entire range of 400-1200 m/z was divided in 32 overlapping precursor isolation windows (25Da wide) and all ions within each 25Da window were fragmented through repeated cycling of consecutive precursor isolation window. Adapted from [67].*

This results in not only a comprehensive and continuous information on all detectable fragment and precursor ions but also a mixed and highly convoluted MS2 spectra in which there is a loss of link between the fragment ions and the corresponding precursors they originated from [66, 67]. Such spectra could not be analysed using standard database search tools used for spectra obtained through DDA. Hence SWATH-MS analysis requires generation of a spectral ion library beforehand using DDA approach. This peptide spectral library serves as a reference against which the MS2 fragment ion spectra generated from SWATH-MS with their corresponding retention time, could be matched to obtain peptide or protein identification and quantification results. Therefore, the depth of

identification and quantification by SWATH-MS is limited by the depth of spectral ion library generated using shotgun proteomics [69]. This approach is termed as targeted data extraction.

iTRAQ based comparative proteomic analysis involves less LC-MS run time with no inter-run variability among the samples, due to pooling of labelled samples before injection into LC and provide accurate and precise quantification. Label-free SWATH-MS, on the other hand, is a cost-effective approach which has no restriction on the number of samples that could be compared and with generation of deeper spectral ion library could provide near-complete proteome coverage [65] [70]. Thus, due to the advantages associated with both labelled and label-free quantitative proteomic approaches, they complement each other and thereby employing both these approaches would help in generating in-depth proteomic profiles. Clinical success of quantitative proteomic profiling using SWATH and iTRAQ is evident from the wide range of studies aimed at identifying disease biomarkers and therapeutic targets for multitude of diseases including cancer. Some of these include identification of salivary biomarkers in oral cancer patients [71], serum biomarkers associated with progression of metastatic prostate cancer [72], biomarkers for vernal keratoconjunctivitis (VKC) from tear fluid [73], calcium/calmodulin-dependent protein kinase kinase 2 (CAMKK2) as the potential therapeutic target for gastric adenocarcinoma [74] using iTRAQ based quantitative profiling and use of SWATH based quantitative profiling for detecting carbonic anhydrase 2 as a biomarker for nasopharyngeal carcinoma [75], serum biomarkers for growth hormone deficiency in children [76], CHTOP (chromatin target of protein arginine methyltransferase) as a therapeutic target for chemo resistant epithelial ovarian cancer [77] and novel molecular networks contributing to pathogenesis of ischemic retinopathies [78].

Thus, generating comparative proteomic profiles of CML-BC cell line with active and inhibited Bcr-Abl as well as sensitive and resistant to IM, using iTRAQ and SWATH-MS, would help in delineating Bcr-Abl downstream signalling and alternate signalling pathways that confer resistance respectively.

2. REVIEW OF LITERATURE

Identification of additional therapeutic targets for TKI-resistant CML-BC is clinically important. Proteomic profiling is the appropriate approach to identify relevant targets in an unbiased manner. However, this approach is minimally explored for identification of therapeutic targets for CML. Whether to gain deeper insights on Bcr-Abl downstream signalling or on molecular alterations that confer resistance to IM, so as to identify alternate therapeutic targets for IM-resistant CML, only nineteen proteomic studies have been reported. To better comprehend the outcome and strategies involved, the studies can be categorized on the basis of their objective, proteomic approach employed and biological material used.

2.1. STUDIES DIFFERED IN THEIR OBJECTIVE

Among the proteomic studies on CML, two studies [79, 80] were aimed at finding predictive biomarkers for monitoring therapeutic response in chronic phase patients, wherein these pilot studies identified proteins like alpha-1-antitrypsin, CD5 molecule-like protein, transthyretin proteins [79], Myc and receptor tyrosine kinase TYRO3 [80] as putative biomarkers. Ten studies were found to be focused on delineating Bcr-Abl downstream components. Of these four were only concerned with identifying direct interactors of Bcr-Abl, wherein Brehme *et.al.,* reported Shc1, Sts-1, Crk-I, p85, c-Cbl, SHIP-2 and Grb2 as core interactors [81], while Preisinger *et.al.,* Patel *et.al.,* and Kuzelova *et.al.,* identified few cytoskeletal regulatory proteins [82, 83] and few cellular adhesion molecules [84] as interactors respectively. Studies by Pizzatti *et.al.,* Park *et.al.,* Griffiths *et.al.,* Xiong *et.al.,* and Arvaniti *et.al.,* though identified global proteomic changes brought about by Bcr-Abl [85-89], most of it were restricted to either identification of differentiators or their functional categorization without any further validation to confirm their involvement in Bcr-Abl pathway. The only study that has identified and suggested a targetable Bcr-Abl downstream component was by Balabanov *et.al.,* wherein eIF5A, the only eukaryotic protein activated by the post translational modification hypusination was found to be downregulated upon IM treatment, based on which inhibition of

hypusination through pharmacological agents along with inhibition of Bcr-Abl by IM, was proposed as the promising new approach to overcome Bcr-Abl mediated resistance. The study further demonstrated in vitro that synergistic use of hypusination inhibitor and IM could selectively induce apoptosis in cells harbouring wild type as well as M351T mutant Bcr-Abl though was not effective against T315I mutation [90].

To understand the molecular alterations that confer IM-resistance, seven studies have been carried out. Studies by Park *et.al.*, Rosenhahn *et.al.*, Zhang *et.al.*, and Ferrari *et.al.*, reported list of proteins differentiallly expressed in resistance but did not identify a key molecule or demostrate its involvement with resistance by inhibitng the same [87, 91-93] while Colavita *et.al.*, demonstrated that increase in GSH concentration modulates redox balance in resistant cells [94]. Hsp70 was found to be upregulated in IM-resistant cells by Pocaly *et.al.*, [95] and further demonstrated that inhibition of Hsp70 using siRNA reduced the viability of the resistant cells drastically, thereby implying the potential of inhibition of Hsp70 as a potential therapeutic strategy to overcoming IM-resistance [96]. Similarly, Toman *et. al*., identified upregulation of Na+/H+ exchange regulatory factor 1 (NHERF1 also known as SLC9A3R1) in IM-resistant cells. Functional analysis revealed that this upregulation leads to alteration of ion homeostasis with increase in pH and decrease in cytosolic calcium concentration compared to the IM-sensitive counterpart. Further calcium channel blockers and modulators of calcium homeostasis were found to selectively inhibit IM-resistant cells. Hence, this study proposed calcium homeostasis as the cause for resistance and hence the potential therapeutic target for IM-resistant CML [97].

Thus, among the studies aimed at identifying therapeutic targets, only three studies [90, 95, 97] have demonstrated the potential of either inhibiting a molecule (Hsp70) or biological process (hypusination of eIF5A or calcium homeostasis) to overcome TKI-resistant CML. This indicates the need for further exploration to identify and validate therapeutic targets.

2.2. STUDIES EMPLOYED DIFFERENT PROTEOMIC APPROACHES

The proteomic studies on CML have employed either limited or global approach. In the limited approach employed by Brehme *et.al.,* and Preisinger *et.al.,* which involved immunoprecipitation (IP) of whole cell lysates with Bcr-Abl antibody followed by MS analysis, resulted in identification of nine interactors [81] and eight interactors [82] respectively. The strategy of Patel *et.al.,* involved IP with Bcr-Abl antibody followed by detection using protein array containing antibodies corresponding to 224 proteins, from which 31 interactors were identified [83] while Quintas-Cardama *et.al.,* used reverse phase protein array which was probed with 112 antibodies and detected 20 proteins whose expression altered across different phases of CML [98]. These strategies are highly limiting as IP helps in identifying only the interactions of the target protein which can be captured in the experimental system while protein arrays are restricted to detection of changes in preselected panel of proteins, diminishing the possibility of detecting novel changes. This limitation has been overcome to a great extent in studies employing global proteomic approach.

Of the thirteen studies involving global proteomic analysis, ten studies [84, 85, 87, 90, 92-95, 97, 99] have employed two dimensional gel electrophoresis coupled with mass spectrometry approach (2D-MS) while only three studies [86, 88, 89] have employed LC-MS approach. Though 2D gel electrophoresis is a good technique to identify differentiators, the total number of proteins detected in all the above studies is restricted to a maximum of few hundreds. The issue of limited proteome coverage be attributed to the factors such as single pH range used for first dimensional separation with no further zoom in approach, size and percentage of gel in second dimensional separation both of which limits the number of proteins that could be resolved and picked up for further detection in MS. Even the few hundreds of identified proteins would constitute the most abundant proteins, leading to loss of information from less abundant yet biologically important proteins, thereby making it difficult to delineate the signalling pathways. Gel-free LC-MS approach on the other hand comparatively increases the proteome coverage facilitating identification of thousands of proteins

from whole cell lysate. LC-MS studies by Xiong *et.al.,* and Arvaniti *et.al.,* have identified a total of 1344 and 986 proteins respectively [88, 89].

Thus, use of LC-MS based approach would overcome the lacunae of limited proteome profile observed in the above studies and help in generating deeper proteomic prolife to delineate signalling pathways.

2.3. PROFILES FOR CML HAVE BEEN GENERATED BY ASSAYING DIVERSE BIOLOGICAL MATERIALS

Based on the type of biological material used for proteomic analysis, studies could be categorized into patient based and cell line-based studies. Four patients based CML studies have been carried out which differ in the comparison groups used. Pizzatti *et.al.,* compared bone marrow mononuclear cells from 13 healthy donors with 17 CML-CP patients to understand the protein changes specific to chronic phase [85] while Zhang et.al., compared the profiles of bone marrow mononuclear cells from 25 CML-CP and 20 CML-BC patients to identify differentially expressed proteins [92]. $CD34^{+}$ cells from CML-CP patients cultured in vitro with and without IM treatment were studied to understand the effect of IM treatment by Griffiths *et.al.*[86]. Quintas-Cardama *et.al.,* profiled leukemia enriched fractions from 25 CP, 5 AP and 10 BC patients using protein array to understand the biology of CML [98]. About thirteen studies were carried out using various CML-BC cell lines of which K562 has been most commonly used. Brehme *et.al.,* and Preisinger *et.al.,* used K562 cell line to identify Bcr-Abl interacting partners [81, 82]. Comparative profiling of K562 cells with and without IM treatment were used in five different studies [87-90, 99] to delineate Bcr-Abl mediated changes while two studies [87, 95] involved comparison of IM-sensitive and resistant K562 cells to understand the mechanism of resistance. Similarly comparative proteomic profiles were generated from IM-sensitive and resistant counterparts of KCL22 [94], LAMA84 [93] and CML-T1 cells [97] to delineate Bcr-Abl independent IM-resistant mechanisms. BV173, KCL22 and KYO-1 cell lines were used by Patel

et.al., to identify Bcr-Abl interactors [83] while untreated and IM treated JURL-MK1 cell line was used to study the changes associated with cell adhesion and actin polymerization [84].

About 75% proteomic studies on CML have employed a cell-based approach, probably due to the difficulties associated with using patient samples. With respect to patient-based studies, the major challenge has been obtaining homogenous cellular population and sufficient material for proteomic analysis. While homogenous cell population would help in providing a better picture of the proteomic changes, most CML patient based proteomic studies have been carried out using mononuclear cells containing heterogenous population of cells, as enrichment to obtain homogenous population would result in insufficient material for proteomic analysis. On the contrary, cell lines are not only less heterogenous but are also available in ample amounts for generation of in-depth proteomic profile. Further, the key molecules identified from cell line-based study could easily be validated in scarcely available patient samples. This thus implies that cell lines are the ideal biological material for exploratory proteomic studies.

3. AIM AND OBJECTIVES

AIM OF THE STUDY

To understand the mechanism/s of imatinib-resistance in CML-blast crisis and identify potential therapeutic targets using proteomic approach.

OBJECTIVES

I. Generation of comparative proteomic profiles of (1) IM-sensitive cells untreated or treated with IM to delineate the components of BCR/ABL pathway (2) IM- sensitive and resistant cells treated with IM to identify resistance-associated proteomic alterations.

II. Deciphering interactions between differentiators *in silico* to detect the hub molecules therein, validate their functional association and check for their role in development of resistance.

4. MATERIALS AND METHODS

4.1. LIST OF REAGENTS AND ANTIBODIES

Table 4.1: List of routinely used chemicals and reagents

S.No.	Reagents	Catalogue	Company
1	Acetone	SC8F680247	Merck
2	Acetonitrile (ACN)	SF6SF56548	Merck
3	Acrylamide	A8887	Sigma
4	Ammonium formate	70221	Sigma
5	Ammonium persulfate (APS)	28875	SRL
6	Annexin FITC/PI Kit	556547	BD Pharmingen
7	Antibiotic-Antimycotic (Anti-Anti) 10X		GIBCO
8	Bisacrylamide	75821	USB
9	Bovine serum albumin (BSA)	A7906	Sigma
10	Bradford Reagent	PG-035	Genetix
11	Calcium chloride ($CaCl_2.H_2O$)	21097	Sigma
12	Disodium phosphate (Na_2HPO_4)	30158	Merck
13	Dithiothreitol (DTT)	D9779	Sigma
14	Fast green FCF	F7252	Fluka
15	Fetal bovine serum (FBS)	10270-106	GIBCO
16	Formic acid (FA)	94318	Fluka
17	Glycerol	77453	SRL
18	Hydrochloric acid (HCl)	29505	Fisher-Scientific
19	Imatinib (IM)	9084-S	Cell signaling
20	Iodoacetamide (IAA)	I1149	Sigma
21	iTRAQ kit	4352135	Sciex
22	Methanol	SD8F680292	Merck
23	MTT reagent	TC-191	HiMedia
24	Potassium chloride (KCl)	20198	SD fine chemicals
25	Potassium dihydrogen phosphate (KH_2PO_4))	20203	SD fine chemicals
26	Puromycin (culture grade)	P8833	Sigma
27	RPMI 1640	230400-021	GIBCO
28	Sodium acetate (CH_3COONa)	4303	Glaxo lab

29	Sodium bicarbonate ($NaHCO_3$)	20247	SD fine chemicals
30	Sodium chloride (NaCl)	15915	Fisher-Scientific
31	Sodium dodecyl sulfate (SDS)	L3771	Sigma
32	Triethylammonium bicarbonate buffer (TEAB)	T7408	Sigma
33	Tris	T1378	Sigma
34	Trypan blue	T8154	Sigma
35	Trypsin (Proteomic grade)	T6567	Sigma
36	Tween-20	P1379	Sigma
37	Urea	U5378	Sigma
38	Western blotting detection reagent (ECL Prime)	RPN2232	GE Healthcare
39	β-Mercaptoethanol	15433	Merck

Table 4.2. List of antibodies

S.No.	Name	Company	Catalogue no.
1	STAT5	Life Technologies	335900
2	Phospho STAT5	Abcam	ab32364
3	c-Abl	Abcam	ab85947
4	P-Glycoprotein	Abcam	ab170904
5	OCT-1	Abcam	ab181022
6	14-3-3 Epsilon (ε)	Santacruz	sc-23957
7	14-3-3 Gamma (γ)	Santacruz	sc-398423
8	p38 MAPK	Cell signalling technology	9212
9	Phospho p38 MAPK	Cell signalling technology	9211
10	Msk	Abcam	ab99412
11	p-Msk	Millipore	04-384

STRATEGY FOR IDENTIFICATION OF THERAPEUTIC TARGETS FROM BCR/ABL PATHWAY AND OTHERS

4.2. Systematic review and meta-analysis

4.3. Experimental identification of therapeutic targets using CML-BC cell lines

4.3.1. Routine maintenance of cell lines.

4.3.1.1. Cell lines.
4.3.1.2. Revival of cell lines.
4.3.1.3. Cell count of viability.
4.3.1.4. Sub-culturing.
4.3.1.5. Freezing down.

4.3.2. Determination of doubling time and optimal harvest time.

4.3.3. Development of IM resistant cells.

4.3.3.1. IC_{50} of IM by MTT assay

4.3.4. Optimization of IM treatment conditions for comparison of S and R cells

4.3.4.1. Based on cell viability.
4.3.4.2. Based on BCR/ABL activity.

4.3.5. Status of known mechanism of resistance

4.3.5.1. *BCR-ABL* gene amplification.
4.3.5.2. Bcr-Abl overexpression.
4.3.5.3. *BCR-ABL* kinase domain mutation.
4.3.5.4. Altered drug transport.

4.3.6. Proteomic analysis.

4.3.6.1. Protein estimation.
4.3.6.2. Sample preparation.
4.3.6.3. IDA.
4.3.6.4. SWATH.
4.3.6.5. iTRAQ.

4.3.7. Bioinformatic analysis (STRING).

4.3.8. Validation by western blotting.

4.3.9. Functional validation.

4.3.10. Statistical analysis.

4.2. SYSTEMATIC REVIEW AND META ANALYSIS: SEARCH FOR THERAPEUTIC TARGETS FROM REPORTED PROTEOMIC ANALYSIS OF CML CELL LINES

4.2.1. SEARCH STRATEGY AND CRITERIA FOR SELECTION OF STUDIES

A systematic review was carried out to identify CML blast crisis cell line studies employing proteomic analysis to understand either Bcr-Abl downstream signalling pathway or Bcr-Abl independent signalling pathways that confer resistance, using Pubmed's advanced search tool. The search strategy with keywords used is mentioned in Table 4.3. Articles published till May 2019 were considered. The identified studies were then subjected to a preliminary title-based screening to exclude unrelated and review articles. This was followed by abstract level screening by two researchers independently based on two inclusion criteria (1) use of CML-BC cell lines as biological material (2) generation of mass spectrometry based comparative proteomic profiles of either untreated and imatinib treated cells or imatinib sensitive and resistant cells. The abstracts which were approved by both researchers were taken for final level of selection involving full-text screening of articles. Articles that failed to fulfil the inclusion criteria upon full-text screening were excluded. The finalized articles were then categorized based on the comparison groups involved and a data extraction sheet containing a list of differentiators identified from each study within the comparison group was prepared.

Table 4.3: Search strategy for systematic review

Search	Query
#6	Search (#4 OR #5)
#5	Search (#1 AND #3)
#4	Search (#1 AND #2)
#3	Search Mass spectrometry
#2	Search (Proteome) OR Proteomic
#1	Search ((Chronic Myeloid leukemia [MeSH Terms]) OR BCR-ABL) OR BCR/ABL

4.2.2. META-ANALYSIS

The differentiators identified from studies within a comparison group were assessed for their overlap with the help of venn diagram plotted using the program InteractiVenn [100]. The differentiators were then pooled together and their functional association was assessed by STRING (v.11) analysis [101]. The analysis parameters used were as follows: Organism – Homo sapiens, active interaction sources – experiments and databases, minimum required interaction score – high confidence (0.7).

4.3. EXPERIMENTAL IDENTIFICATION OF THERAPEUTIC TARGETS USING CML CELL LINES

4.3.1 ROUTINE MAINTENANCE OF CELL LINES

4.3.1.1 Cell lines

CML-BC cell lines K562, KCL22 and KU812 (Table 4.4) were a generous gift from Dr. Tadashi Nagai, Jichi Medical University, Tochigi, Japan. All three suspension cell lines were maintained in complete RPMI (cRPMI) medium.

Table 4.4: Details of CML-BC cell lines

Cell line	Lineage
K562	Erythroleukemic blast
KCL22	Myeloblast
KU812	Basophilic blast

Reagents:

(a) RPMI 1640 medium:

Powdered media along with 2 gm $NaHCO_3$ was dissolved in 800ml autoclaved distilled water and the final volume was made up to 1000ml. Media was filter sterilized, aliquoted in sterile tubes and stored at 4ºC. The media was made complete before use by adding 10% FBS and 1% antibiotic.

(b) Phosphate-buffered saline (PBS):

137mM NaCl, 2.7mM KCl, 8mM Na_2HPO_4 and 1.5mM KH_2PO_4 were dissolved in 1000ml distilled water and sterilized by autoclaving.

(c) Freezing media:

90% FBS + 10% DMSO

(d) Imatinib:

10mM IM stock was prepared by solubilizing a vial of IM in 847.9 µl DMSO.

4.3.1.2. Revival of cell lines

Cells cryopreserved in liquid nitrogen were thawed at 37ºC, transferred to microcentrifuge tube and centrifuged at 1500rpm for 5 min. The supernatant was discarded and the pellet was washed with 1ml cRPMI to remove residual DMSO. The pellet was then reconstituted in 1ml cRPMI and cell number, viability was calculated. Depending on the cell count, cells were seeded in the appropriate tissue culture dishes as mentioned in Table 4.5 and incubated in humidified incubator with 5% CO_2 at 37ºC.

Table 4.5: No. of cells seeded in different tissue culture dishes

Tissue culture dish	No. of cells seeded	Media volume
60mm plate	0.3 - 0.5 x 10^6	3 ml
90mm plate	1 x 10^6	8 ml
T-25 flask	0.5 x 10^6	4 ml
T-75 flask	1 x 10^6	8 ml

4.3.1.3. Cell count and viability by trypan blue dye exclusion method

From 1ml cell suspension, 10µl cells were mixed with 10 µl trypan blue dye (1:1 v/v) and loaded on to haemocytometer for determining cell count and viability. In case of dense cell suspension, cells were diluted in PBS prior to mixing with trypan blue. Cells in WBC chambers on either side of haemocytometer were counted. Live cells remain unstained while dead cells take up trypan blue and appear blue in colour. The cell count and percent viability were calculated as follows.

Cells/ml = (No. of cells counted/ No. of chambers counted) X Dilution factor X10^4

% viability = [No. of live cells / Total no. of cells (Live + dead)] X 100

4.3.1.4. Sub-culturing of cell lines

Cells were harvested in the log phase at about 70% confluence. Media containing cells was collected in a 15ml tube. The tissue culture dish was washed with PBS to collect the remaining cells and pooled in the same 15ml tube which was centrifuged at 1500 rpm for 5 min. The supernatant was discarded, pellet was reconstituted in 1ml cRPMI, cells were counted using haemocytometer and reseeded.

4.3.1.5. Freezing down of cell lines

To freeze down, a pellet of minimum $2X10^6$ cells was resuspended in 1ml freezing medium and aspirated to get a single cell suspension. This was transferred to a freezing vial and slow freezing was carried out by placing the cells at 4ºC for 20 min, -20ºC for 2 hrs. and -80ºC overnight. Cells were then transferred to liquid nitrogen for long time storage.

4.3.2. DETERMINATION OF DOUBLING TIME AND OPTIMAL HARVEST TIME

Cells (~0.2×10^6) were seeded in 60mm plate, with a total of 15 plates. Every 24 hrs, 3 plates were harvested (technical replicates) for 5 consecutive days (day 0 – day 4), cell count and viability were recorded for cells from each plate independently as mentioned in section 4.3.1.3 A plot of growth curve and viability were plotted with incubation time (hrs) on X-axis and cell count or viability on Y-axis respectively. Optimal harvest time was identified based on these graphs. Further, doubling time was calculated using the formula

$$t_d = (t \times \ln 2) / \ln (C2/C1)$$

where t_d - doubling time

t – Incubation time in hrs.

C1 - Cell count at the beginning of incubation time (exponential phase)

C2 - Cell count at the end of incubation time

4.3.3 DEVELOPMENT OF IMATINIB (IM) RESISTANT CELLS

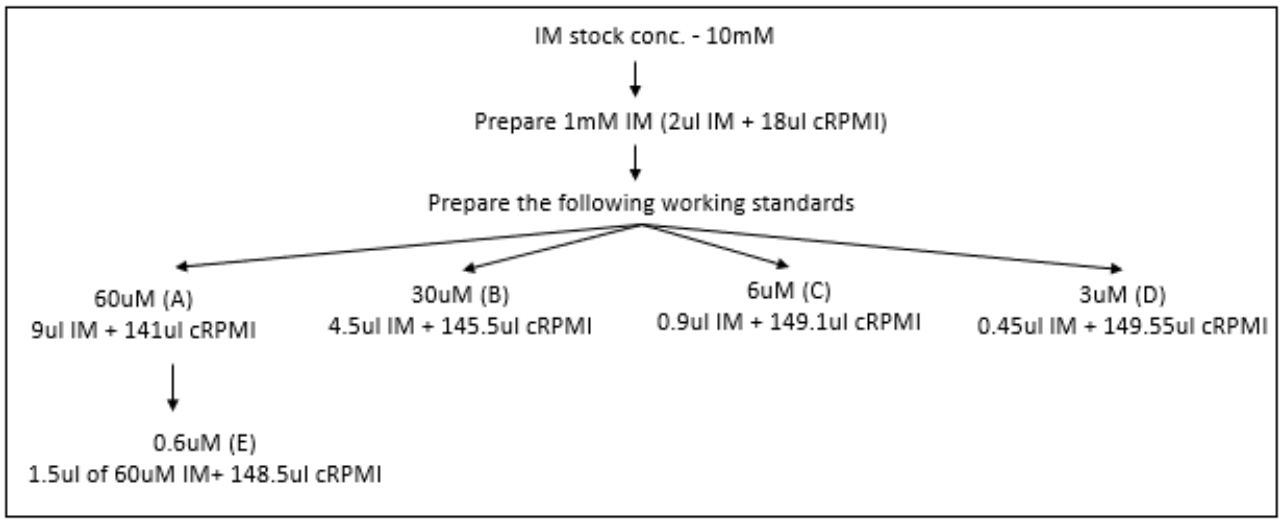
IM resistant K562, KCL22 and KU812 cells were developed from their corresponding drug sensitive cells by gradual IM dose escalation. Starting with 0.1µM, IM dosage was incremented by 0.1µM every 3 passages to a maximum of 1 µM. If viability dropped less than 50% during dose escalation, the cells were maintained in the absence of IM until viability reaches to about 80%. The entire process of developing resistant cells spanned a period of about 5 months. Once the final concentration was reached, resistance was confirmed by MTT assay and the cells were always maintained in that IM concentration where viability remains more than 80%. The resistant cells were labelled as K562/R, KCL22/R, KU812/R and their corresponding sensitive counterpart as K562/S, KCL22/S and KU812/S.

4.3.3.1 Determination of IC_{50} of IM by MTT assay

Reagents:

- **IM working standards:**

  From 10mM IM stock, different working standards were prepared as mentioned below



- **Vehicle control (DMSO):**

  2 µl DMSO was solubilized in 18 µl cRPMI. From this, 12 µl was added to 188 µl cRPMI (to mimic the dilution volumes used for 60uM IM solution), to prepare DMSO working standard.

- **MTT reagent:**

5mg/ml MTT stock was prepared in sterile PBS. The contents were solubilized by placing the tube on rocker and stored at 4ºC.

- **Acidified SDS (10% SDS with 0.01N HCl):**

  43 µl conc. HCl was added to 50 ml of 10% SDS to get acidified SDS.

Methodology:

On day 1, 0.3 X 10^6 cells were resuspended in 4ml cRPMI and 100 µl was added to each well of 96 well plate, to get a final cell count of 7500 cells/well. 3 wells with 100 µl cRPMI without cells served as blank. Cells were incubated at 37ºC overnight. On day 2, 5 different IM working standards were prepared (A-E) in dark, from which 20 µl was added to the corresponding wells to get a final concentration of 0.1 µM, 0.5 µM, 1 µM, 5 µM and 10 µM IM respectively, with 5 wells (technical replicates) for each concentration. Since IM was solubilized in DMSO, 20 µl of DMSO working standard was added to blank and control wells. Cells were incubated for 48hrs, followed by addition of 30 µl MTT reagent to each well in dark, to get a final concentration of 1mg/ml. Cells were then incubated for 4 – 6 hrs. to allow formazan crystal formation. The crystals were solubilized with 100 µl acidified SDS, incubated overnight at 37ºC. Absorbance at 570nm was then recorded using ELISA plate reader.

4.3.4 OPTIMIZATION OF IM TREATMENT CONDITIONS FOR COMPARING S AND R CELLS

4.3.4.1 Cell viability by Annexin-FITC/PI staining using flow cytometry

One million cells seeded in 60mm tissue culture plates were maintained untreated or treated with 0.75µM IM For K562/S cells and 1 µM IM for KCL22/S and KU812/S cells for 2, 6, 18 and 24 hrs. To obtain annexin and PI positive cells for compensation, 2x10^6 cells were seeded in 60mm tissue culture dish and treated with higher concentration of IM (10 µM) for 4 hrs (termed as ‘comp cells’). At the end of stipulated treatment duration, all cells were harvested, washed with PBS and

reconstituted in 150µl of 1X FACS buffer provided in the kit This was followed by addition of cells and dyes to 5ml round bottom falcon tubes as mentioned in Table 4.6.

Table 4.6: Details of sample sets for flow cytometric analysis

Tube no.	Label	Constituents
1.	Control (only cells)	50 µl control cells (untreated) + 25 µl comp cells
2.	Compensation tube (only annexin)	60 µl comp cells + 5 µl Annexin V FITC
3.	Compensation tube (only PI)	60 µl comp cells + 5 µl PI
4.	Dual stained control	50 µl control + 5 µl Annexin V FITC + 5 µl PI
5.	2 hr. IM treatment (only cells)	25 µl treated cells set 1 + 25 µl treated cells set 2
6.	2 hr. IM treatment (Dual staining set 1)	100 µl treated cells set 1 + 5 µl Annexin V FITC + 5 µl PI
7.	2 hr. IM treatment (Dual staining set 2)	100 µl treated cells set 2 + 5 µl Annexin V FITC + 5 µl PI
Similar set of tubes (tube no. 5 – 7) were maintained for other IM treatment durations.		

Note: All additions were done in dark as the dyes are light sensitive.

The tubes were then incubated in dark at room temperature (RT) for 15 min and 400 µl 1X FACS buffer was added to each tube after incubation. Data acquisition was carried out in BD FACS Calibur™.

4.3.4.2 Bcr-Abl activity by western blotting

Preparation of whole cell lysate

Reagents:

SDS lysis buffer **-** 10% glycerol, 2% SDS, 5% β-mercaptoethanol and 62.5 mM Tris pH 6.8

Methodology:

$1x10^6$ cells (fresh or snap frozen) were resuspended in 100µl SDS buffer, boiled for 10 min, placed in ice for 2 min and centrifuged at 13,000xg for 15min. The supernatant was collected in a fresh tube and stored at -20ºC until use. In samples to be used for SDS-PAGE, 2 µl bromophenol blue (BPB) was added.

SDS-PAGE and Western blotting

Reagents:

30% Acrylamide:

Acrylamide 29.2 g

Bis-acrylamide 0.8 g

The components were solubilized in 50ml distilled water by placing on magnetic stirrer, in dark. After solubilization, the volume was made up to 100 ml, filtered using Whatman filter paper and stored in amber bottle.

1M Tris:

In 750ml distilled water 121.1 gm Tris was dissolved. The pH was adjusted to either 6.8, 7.5 or 8.8 depending on the requirement, with concentrated HCl and the volume was made up to 1L.

20% (SDS):

In about 60ml distilled water 20 gm SDS was solubilized by placing it in 37^{o}C shaker incubator or in normal incubator with intermittent stirring every 1 hr. Upon complete solubilization, the volume was made up to 100ml.

20% (APS)**:**

To prepare 20% APS 0.2gm was dissolved in 1ml distilled water.

1X Electrode buffer:

Glycine - 72 gm

Tris - 15 gm

SDS - 10 gm.

The volume was made up to 5 L with distilled water.

1X transfer buffer:

Glycine - 72 gm

Tris - 15 gm

Methanol - 1 L

Glycine and tris were dissolved in about 3 L distilled water. Methanol was then added and the volume was made up to 5 L with distilled water.

Tris buffered saline (TBS):

NaCl - 146.1 gm

Tris pH 7.5 - 100 ml

The volume was made up to 5 L with distilled water.

Wash buffer (TBST):

TBST was prepared by adding 0.05%, 0.1% or 0.2% Tween-20 to TBS.

Methodology:

Bcr-Abl activity was assessed based on the phosphorylation status of its downstream substrate STAT5, which was detected by western blotting 20 µl lysate corresponding to 0.2 x10^6 cells from untreated and IM treated sensitive cells was boiled for 5 min and loaded on to 8% mini (Bio-Rad Protean III) polyacrylamide gel (Table 4.7). The gel was run initially under a constant voltage of 50V for about 20 min to ensure proper stacking of proteins and was then increased to 150 V.

Table 4.7: Composition of 8% SDS-PAGE

6.5 % Stacking gel		**8% Resolving Gel**	
30% Acrylamide	1.08 ml	30% Acrylamide	2.66 ml
1M Tris pH 6.8	875 µl	1M Tris pH 8.8	3.73 ml
20% SDS	50 µl	20% SDS	67 µl
20% APS	50 µl	20% APS	67 µl
TEMED	2 µl	TEMED	3.33 µl
Distilled water	2.94 ml	Distilled water	3.47 ml
Total Volume	**5 ml**	**Total Volume**	**10 ml**

SDS-PAGE was followed by western blotting wherein proteins were transferred from gel onto the PVDF membrane. Prior to transfer, the PVDF membrane was activated with methanol. A sandwich was prepared by placing sponge and Whatman filter paper no. 1 on either side of gel and PVDF membrane in a cassette which was then placed in the tank containing transfer buffer. The transfer was

carried out at 100V for 1 hr. Following transfer, proteins were visualised by staining the membrane with 0.1% fast green stain and colorimetric image for determining whole lane intensity was obtained in Chemidoc^{TM}MP imaging system (BioRad). The stain was removed by washing in 0.1% TBST. This was followed by immunodetection of STAT5 and phospho STAT5 proteins, the conditions for which is mentioned in Table 4.13. The blot was then developed using ECL Prime® developing reagent and the chemiluminescence was either captured on X-Ray films or imaged using Chemidoc^{TM}MP imaging system (Bio-Rad). The band and lane intensities were calculated using ImageJ software, NIH (for X-ray film-based detection) or Image Lab software, Bio-Rad (for chemidoc based detection). The band intensities were normalized with their corresponding whole lane intensity.

4.3.5 STATUS OF KNOWN MECHANISMS OF RESISTANCE IN IM RESISTANT CELLS

4.3.5.1. *BCR-ABL* gene amplification by fluorescent in-situ hybridization (FISH)

BCR-ABL gene amplification by FISH was carried out at cytogenetics lab, ACTREC. Briefly, LSI B*CR/ABL1* dual colour, dual fusion probe (Zytovision, Germany) containing green fluorophore tagged BCR specific gene sequence and red fluorophore tagged ABL specific gene sequence was used for FISH on interphase and metaphase cells obtained from IM-sensitive and resistant cell lines. About 200 interphase cells and 5 metaphase cells were analyzed and images were captured using Olympus BX61 fluorescence microscope with GenASIs software from Applied Spectral Imaging, Israel.

4.3.5.2. Bcr-Abl overexpression by western blotting

SDS-PAGE and western blotting for Bcr-Abl protein was carried out as mentioned in section 4.3.4.2, using c-Abl antibody which detects c-Abl as well as Bcr-Abl. The immunodetection conditions used are mentioned in Table 4.14.

4.3.5.3 *BCR-ABL* kinase domain mutation by sequencing

BCR-ABL kinase domain mutation analysis of IM sensitive and resistant cells was carried out at Hematopathology lab, ACTREC, using the direct sequencing protocol [102] that is part of routine screening procedure for CML patients. Briefly, cDNA was synthesized from RNA extracted from IM-sensitive and resistant cells. *BCR-ABL* kinase domain was amplified through nested PCR approach, with the initial BCR-ABL1 gene amplification starting from exon 2 of BCR to exon 10 of ABL, followed by amplification of ABL kinase domain (exon 4-10 of ABL1) in a second round of PCR using overlapping primer sets. The resulting amplicon was subjected to Sanger sequencing using Big-Dye v3.1 chemistry (Life technology Ltd.).

4.3.5.4. Status of drug transporters by western blotting

To assess the status of IM import and export proteins hOCT-1 and P-glycoprotein (P-gp), western blotting was carried out as mentioned in section 4.3.4.2. P-glycoprotein was detected in native state as specified in the product datasheet. The immunodetection conditions used are mentioned in Table 4.14.

4.3.5.5. Intracellular IM concentration by targeted LC-MS analysis

IM-sensitive and resistant cells (~$8x10^6$) treated with IM for 12hrs and untreated sensitive cells which serve as negative control, were harvested and washed 5 times with 10 ml PBS by centrifuging at 1500 rpm for 5 min, to remove traces of residual IM from media. The pellet was then resuspended in 1ml PBS, cells were counted and equal number of cells from the comparison sets were used for intracellular IM extraction, using liquid-liquid extraction protocol [103]. Cells were centrifuged at 1500 rpm for 5 min and the pellet was lysed by reconstituting in 150 µl ice cold D/W. The tubes were plunged into beaker containing water placed at -20ºC for 30 sec and 37ºC for 30 sec. Trazodone (60 µl, diluted 1000 times with D/W), which served as internal standard, mixed with 600 µl methanol prechilled at -20ºC was added to the tubes containing cell suspension and vortexed. This was followed by incubation at RT for 5 min and addition of 450 µl chloroform. The tubes were vortexed for 30 sec every 5 min for 30 min, by placing the tube on ice in between vortexing. The tubes were then transferred to RT, 150 µl ice cold D/W was added and centrifuged at 1000 x g for 1 min at 4ºC. The

tubes were incubated at -20ºC for 3 hrs. This resulted in formation of aqueous phase at the top and organic phase at the bottom with lysed cells at the interface. To detect IM, aqueous phase was collected in a fresh tube, equal amount of acetonitrile was added and incubated at 4ºC for 20min to precipitate proteins. The tubes were centrifuged at 10,000 x g for 10 min at 4ºC. The supernatant containing metabolites was transferred to a fresh tube, dried in speed vac, reconstituted in 100 µl of 0.1% formic acid in water and subjected to LC-MS/MS analysis.

Targeted LC-MS/MS analysis for Imatinib was performed using Shimadzu LC-MS-8060 triple quadrupole system (courtesy Shimadzu Analytical India Pvt Ltd, Mumbai) wherein 5 µl sample was injected. HPLC separation was carried out using Shim-pack GIST (4.6 X 75mm, 3µm) C8 column (Shimadzu), maintained at 40^{O}C with methanol as mobile phase A and 0.1% formic acid in water as mobile phase B with a flow rate of 0.4 ml/min and a total run time of 8 min. The autosampler was set at 5^{O}C. The mass spectrometer was operated in the positive ESI mode and the MRM transitions used for quantification of Imatinib was 494.00 > 394.2 while that of trazadone was 372.10 > 148.15. The CID gas pressure was set at 270 kPa with interface voltage of 4kV, nebulizing gas flow of 3L/min, heating gas flow of 10L/min, interface temperature of 300 ^{O}C, desolvation line temperature of 250ºC, heat block temperature of 450ºC and drying gas flow of 10L/min.

4.3.6 PROTEOMIC ANALYSIS OF K562 CELL LYSATES

4.3.6.1. Protein estimation by Bradford's method

Reagents

BSA standards:

1 mg/ml BSA was prepared in distilled water and was serially diluted to obtain a range of BSA standards i.e., 1 mg/ml, 0.5 mg/ml, 0.25 mg/ml, 0.125 mg/ml and 0.0625 mg/ml.

Methodology

Protein estimation was carried out in 96 well plate format. Each sample was maintained in triplicate. Distilled water served as blank. 5 µl D/W or 5 µl BSA standards or 4 µl D/W + 1 µl lysate were added

in their respective wells. This was followed by the addition of 100 µl Bradford's reagent to all wells in dark. The plate was incubated in dark for 5 min and the absorbance was recorded at 595 nm using ELISA plate reader.

4.3.6.2 Sample preparation for LC-MS/MS analysis

Reagents:

- 1M Tris-HCl pH 8:

Tris (121.14 gm) was dissolved in 1L of distilled water. The pH was adjusted to 8 with conc. HCl and stored at 4°C.

All other reagents were prepared fresh before use

- 6M urea, 50mM Tris-HCl pH 8.0:

Urea (180mg) + 25µl Tris-HCl pH 8.0 (1M) + 475 µl D/W.

- 200mM DTT, 50mM Tris-HCl pH 8.0:

DTT (3.08mg) + 5 µl Tris-HCl pH 8.0 (1M) + 95 µl D/W.

- 200mM IAA, 50mM Tris-HCl pH 8.0:

IAA (3.7mg) + 5 µl Tris-HCl pH 8.0 (1M) + 95 µl D/W.

- 1mM CaCl2, 50mM Tris-HCl pH 8.0:

$CaCl_2.2H_2O$ (0.74mg) + 250 µl Tris-HCl pH 8.0 (1M) + 4750 µl D/W.

- Trypsin:

Trypsin vial containing 20 µg lyophilized powder was reconstituted with 100µl of 50mM acetic acid to get a final concentration of 0.2µg/µl. Aliquots of 10 µl were stored at -80°C.

- Equilibration/Rinse solution:

0.1% Formic acid (FA) in D/W

- Wetting/Elution solution:

0.1% Formic acid in 80% ACN

Methodology

In-solution digestion:

Whole cell lysate was prepared as mentioned in section 4.3.4.2. The lysate was subjected to acetone precipitation to remove detergent by adding 1ml chilled acetone, incubating for 1hr and centrifuging at 13,000 rpm for 10min. Protein pellet thus obtained was resuspended in 100 µl 6M urea and acetone precipitation was repeated. The pellet was resuspended in 100 µl of 6M urea and 5 µl of 200mM DTT. The tube was vortexed gently and incubated for 1hr at RT. Protein estimation was carried out by Bradford's method (section 4.3.6.1). Further, 20 µl of 200mM IAA was added to the lysate, vortexed gently and incubated for 1hr in dark. Twenty µl of 200mM DTT was then added and incubated for 1hr in dark. Before trypsin digestion, urea concentration was brought down to 0.6M using 855 µl of 1mM $CaCl_2$. From this, 10 µg protein was subjected to in-solution trypsin digestion by adding proteomic grade trypsin in the ratio of 1:50 trypsin: protein (w/w) and incubating for 16 hrs. at 37^{o}C. The peptides were dried in speed vac and desalted using C18 column.

Desalting of peptides:

The dried peptide sample was reconstituted in 70 µl equilibration solution and mixed well by vortexing. C18 spin tip (Ziptip – Pierce cat no. 84850) was placed in microcentrifuge tube using spin adapter. The ziptip was conditioned by adding 20 µl of wetting solution and centrifuging at 1000 X g for 1 minute. This step was repeated thrice. The tip was then equilibrated by adding 20 µl equilibration solution and centrifuging at 1000 X g for 1 minute, thrice. Ziptip and adapter were transferred to a new microcentrifuge tube, 70 µl sample was added and centrifuged at 1000 X g for 1 minute. The sample flow through was again passed through the ziptip to ensure complete binding of peptides. This process was repeated twice. The tip was transferred to a new tube and washed with 20 µl rinse solution thrice to ensure removal of unbound impurities. The tip was transferred to a new tube and the sample was eluted by adding 15 µl elution solution and centrifuging at 1000 X g for 1 minute. This step was repeated thrice. The sample was then dried in speed vac, reconstituted in 20 µl

0.1% FA so as to obtain a concentration of 0.5 µg/ µl, vortexed and centrifuged at 13,000 rpm for 5min. 19µl sample was transferred to HPLC vial.

4.3.6.3 LC-MS/MS data acquisition and analysis for IDA run

Data acquisition:

Samples reconstituted in 0.1% FA were injected into Eksigent ekspert^{TM} nano-LC 400 with cHiPLC® system, with trap column (200µmX 0.5mm) and analytical column (75µmX 15cm), both packed with 3 µl ChromXp C18 (120A^{0}). On column peptide concentration was maintained as 2µg for IDA runs. For reverse phase HPLC, 0.1% FA in water and 0.1% FA in acetonitrile (ACN) served as solvent A and B respectively. A gradient elution with increasing percentage of mobile phase B was used to elute the peptides at a flow rate of 300nL/min. Eluate from the column was analyzed in an on-line Triple TOF 5600+ (Sciex, USA) mass spectrometer in a positive ion mode. Mass spectra were acquired in IDA mode which involved a survey scan over a mass range of 350-1250 m/z and MS/MS scan over 200-1800 m/z for top 30 precursor ions with rolling collision energy, 50mDa mass tolerance and accumulation time of 250msec for MS and about 50msec for MS/MS.

Data analysis:

The raw spectral file generated in .mgf format was converted to .wiff format using Peakview 1.2 software (Sciex, USA). This software also facilitates visualization of total ion chromatogram (TIC). The .wiff file was then analysed in Protein Pilot 4.5 software (Sciex, USA) with paragon algorithm to obtain protein identities. The parameters used were as follows: Cysteine alkylation – IAA; Digestion – trypsin; no special factor was chosen; Species – Homo sapiens and no specific processing was chosen. The search effort was set to 'thorough ID' and false discovery rate (FDR) analysis was enabled. Proteins identified with 1% FDR were considered reliable. The search was carried out against UniProt database (November 2016 release) containing reviewed human proteins.

4.3.6.4 Label-free quantification by SWATH-MS

Optimization of normalization method and criteria for differentiator identification

The scheme of experiments employed to identify the normalization strategy optimum for SWATH-MS data involves -

A. Inclusion of a quantitatively defined dataset from public domain, generated from hybrid of peptides from three different sources mixed in defined proportions, to serve as a 'reference set'. Generation of datasets using K562 cells for quantitation by SWATH-MS, referred to as 'study set' which includes one set with smaller number of samples and two sets with larger number of samples. Further, inclusion of two datasets from public domain comprising of larger sample size, to serve as 'validation set' to confirm the findings in the study set.

B. SWATH-MS analysis of reference, study and validation set.

C. Normalization of SWATH-MS data obtained from reference, study and validation sets using methods in MS-instrument based Marker view software and statistical tool Normalyzer and identification of optimum method of normalization based on statistical criteria.

D. Identification of differentiators from this normalized data based on criteria of p value, fold-change and both, followed by cluster analysis of these differentiators.

Details of samples for SWATH-MS analysis

The reference set was obtained from data published by Navarro et. al., [104] wherein samples were prepared by mixing known proportions of constituent proteome (i.e. with known fold-difference in quantities). Samples with a hybrid of human, yeast and E. coli peptides referred to as HYE124 had differences in relative proportions of the constituent peptides and served as control (65% w/w human, 30% w/w yeast, 5% w/w *E.coli* peptides) & test (65% w/w human, 15% w/w yeast, 20% w/w *E.coli* peptides). SWATH runs of these samples in technical triplicate and their corresponding spectral ion library deposited in Proteome Xchange consortium (identifier-PXD002952), was used for SWATH data analysis.

The 'study set' was generated in our laboratory using IM-sensitive untreated (S), 12hr. IM treated (S+IM) and IM-resistant (R) K562 cells. The samples were prepared as mentioned in section 4.9.1.2.

Additionally, these samples were spiked with 1pmol/µl of digested *Escherichia coli* β-galactosidase (β-gal) peptides, which served as internal standard, before transferring to HPLC vials. SWATH-MS profiles were generated for four biological replicates of S, S+IM and R, each run-in triplicate.

The 'validation set' constituted SWATH data deposited by Tan et.al.[105] and Guo et.al.[106] in Proteome Xchange consortium with identifiers PXD006106 and PXD000672 respectively. SWATH runs of ten biological replicates of HeLa Kyoto cells untreated (UT) and treated with formaldehyde (FA) were obtained from PXD006106 while duplicate SWATH runs of normal (N) and tumorous (T) kidney tissue samples from nine patients were obtained from PXD000672.

LC-MS/MS data acquisition for the study set

In SWATH-MS, each sample was subjected to 1 IDA run for spectral ion library generation followed by 3 DIA (SWATH) runs, which served as technical replicates. Thus, with four biological replicates, K562 S, S+IM and R cells had 4 IDA runs and 12 SWATH runs each. LC separation with a 225-minute gradient (Table 4.8) and MS data acquisition in IDA mode were carried out as mentioned in section 4.3.6.3. For DIA-SWATH acquisition, the instrument was tuned to a looped product ion mode. A sequential isolation window width of 25m/z (with 1m/z overlap) covering a mass range of 350-1250 m/z was set, resulting in 36 overlapping windows. The accumulation time was 50ms for MS scan and 80ms for MS/MS scan, thereby making a total cycle time of about 3 seconds.

Table 4.8: LC gradient for SWATH analysis

Time (mins)	%A (0.1% FA in water)	%B (0.1%FA in ACN)
0	95	5
12	90	10
92	70	30
112	50	50
127	45	55
147	35	65
167	25	75
187	15	85

207	5	95
225	5	95

The conditions used to generate data by Navarro et al.[104], Guo et.al [106] and that used to generate data experimentally in this study were comparable, while data generated by Tan et.al.[105] used 64 variable wide precursor ion selection window. Further, samples in the reference set and validation set were spiked with indexed retention time (iRT) peptides for retention time calibration while those in the study set were spiked with *E.coli* β-gal peptides.

Generation of spectral ion library for the study set

The reference set from Navarro et.al.[104] was referred to as Dataset A. The data acquired from S, S+IM and R sets were further grouped for comparison into datasets (Table 4.9). Only one out of the four sets of S and S+IM each, was considered as dataset B while all four together as dataset C. All four sets of S+IM and R were included in dataset D. The validation sets from Tan et.al.[105] and Guo et.al [106] were referred to as dataset E and F respectively.

A common spectral ion library was generated for datasets B and C while a separate library was created for dataset D. The spectral ion library for datasets B, C and D was generated by pooling the IDA runs of the corresponding biological replicates and analyzing in Protein Pilot software v4.5 (Sciex, USA), as mentioned in section 4.3.6.3. In SWATH-MS since the samples were spiked with *E. coli* β-gal, the search was carried out against UniProt database containing *E. coli* β-gal protein along with human proteins. The result (.group) file thus generated served as the spectral ion library. For dataset A the spectral ion library deposited by Navarro et al.,[104] generated by pooling individual libraries of constituent human, yeast and *E.coli* peptides, was used. For datasets E and F comprehensive human SWATH library with about 10,000 proteins deposited in SWATH Atlas by Rosenberger et.al.[107] was used.

SWATH data analysis

Spectral alignment and targeted data extraction of the swath runs of all six datasets were carried out in Peak View 2.2 software using MS/MS ALL with SWATH acquisition microapp (Sciex, USA).

Proteins from spectral ion library identified with 1% FDR were first imported into Peak View 2.2 software. Retention time calibration was carried out using iRT peptides for datasets A, E and F and *E. coli* β- gal peptides for datasets B-D. Processing settings were used to filter the ion library, where up to 6 peptides per protein and 6 transitions per peptide with peptide confidence threshold of 99% and FDR of 1%, were chosen for quantification. Modified peptides were excluded from extraction. Extracted ion chromatogram (XIC) window was set to 5 min for datasets A, B, C, E, F and 15 min for dataset D with XIC width of 50ppm. The MS/MS extracted peak areas from the filtered results were exported to Marker View software v1.3 (Sciex, USA) for quantification. The marker view output raw data file with list of proteins and their peak areas were used for further analysis.

Normalization of SWATH Data

The raw data of all datasets was processed and analyzed in a statistical tool Normalyzer, wherein it was log2 transformed and then normalized globally (G) or locally (R) using 10 statistical methods. Global normalization is carried out without consideration of affiliation of the sample such as replicate, control group, test group, etc. [108]. In SWATH-MS since each sample is run individually, errors can arise irrespective of their origin. Thus, in the present study global normalization methods were included. However, since the study focuses on identification of normalization method conducive to biomarker identification, retention of distinguishing features of the comparison groups was necessary while normalizing the data. This was achieved by including local normalization methods for analysis [108]. The normalization methods include locally estimated scatterplot smoothing (Loess-R, Loess-G) which assumes non-linear relationship between the bias in the data and magnitude of protein intensity; robust linear regression (RLR-R, RLR-G) which assumes that the bias in data is linearly dependent of the magnitude of the measured protein intensity; variance stabilization normalization (VSN-R, VSN-G) which aims at making the sample variances nondependent from their mean intensities and bringing the samples onto the same scale; quantile normalization which forces the distribution of the samples to be the same; total intensity (TI-G), average intensity (AI-G) and median

intensity (MedI-G) normalization methods wherein intensity of each variable is divided by sum of intensities, mean of sum of intensities, median intensities of all variables respectively [108, 109].

Table 4.9. Details of datasets:

Datasets	**Source**	**Constituents**	**Purpose – in this study**
Dataset A	Pride ID - PXD002952	3 samples of 65% human, 30% yeast, 5% *E.coli* peptides (Control). 3 samples of 65% human, 15% yeast, 20% *E.coli* peptides (Test)	Reference set – a well-defined dataset with predictable quantification.
Dataset B	In vitro experiments carried out in this study	3 samples of K562/S cells (Control) 3 samples K562/S+IM cells (Test)	Study set to check comparability of observations in defined (A) versus undefined (B) datasets
Dataset C	In vitro experiments carried out in this study	12 samples of K562/S cells (Control) 12 samples K562/S+IM cells (Test)	Larger dataset (C) to check the application of observations from small dataset (B).
Dataset D	In vitro experiments carried out in this study	12 samples of K562/S+IM cells (Control) 12 samples K562/R cells (Test)	Larger dataset (D) to check the consistency of observations in independent large datasets
Dataset E	Pride ID-PXD006106	10 samples of untreated HeLa Kyoto cells (Control) 10 samples of formaldehyde treated HeLa Kyoto cells (Test)	Validation set to check the consistency of observations in independent large datasets
Dataset F	Pride ID – PXD000672	18 non-tumorous kidney tissue samples (Control) 18 tumorous kidney tissue samples (Test)	Validation set to check the consistency of observations in independent large datasets

Marker view v1.3 along with quantitation also provides options for sample normalization using either total area sums (TAS) wherein total area of all peaks in a sample is considered or using area of the selected peaks or internal standard (IS). In this study spiked iRT peptides and trypsin digest of *E.coli* β- gal served as an internal standard for dataset A, E, F and datasets B-D respectively. In TAS as well as IS normalization, the peak areas of each sample were normalized by multiplying with its corresponding scale factor. The scale factor for TAS method was obtained by dividing the average of total area of all samples by the total area of each sample while for IS method the average area of internal standard of all samples was divided by the area of internal standard of each sample. Data normalized by the above two methods i.e. TAS and IS was log2 transformed before running through Normalyzer, to generate the evaluation report.

The normalization efficiency of all 12 methods was assessed through 'Normalyzer' quantitatively by pooled intragroup coefficient of variation (PCV) and qualitatively by relative log expression (RLE) plot as reported in earlier studies [108, 109]

Identification of differentiators from normalized data and cluster analysis

Differentiators were identified from the data of all datasets normalized by 12 methods based on the criteria of p-value, fold-change and a combination of both. To obtain p-value, log2 transformed data, normalized by different normalization methods from comparison groups were assessed by Student's t-test using IBM SPSS statistics 21. Differences in protein intensities with p-value ≤ 0.05 were considered statistically significant and chosen as differentiators. The fold change difference in protein levels was calculated from the peak area values and a cut-off of 1.5-fold change was applied. Further, the efficiency of differentiators obtained from data normalized using the 12 methods to segregate the comparison groups was assessed by cluster analysis. The peak areas of differentiators identified using p-value (≤ 0.05), fold change (1.5 fold) and combination of both were used as inputs for cluster analysis (Fig.1D) in Genesis software v.1.8.1. Hierarchical clustering was performed with the following parameters: Agglomeration rule – Average linkage WPGMA & Calculation parameters – Cluster experiments.

4.3.6.5 Labelled quantification by iTRAQ

Sample preparation

Reagents:

- Conditioning buffer (0.2M $NaH_2PO_4.2H_2O$ + 0.3M CH_3COO-Na, pH 3.0-6.5)

312mg of $NaH_2PO_4.2H_2O$ and 246 mg of CH_3COO-Na were dissolved in 9ml of distilled water. pH was adjusted using formic acid and the volume was made up to 10ml.

- Loading buffer (8mM Ammonium formate in 25% ACN, pH 3.0)

5mg Ammonium formate was dissolved in 5 ml distilled water. 2.5ml of 100% ACN was added and pH was adjusted to 3.0 using formic acid. The volume was finally made up to 10ml using distilled water.

- Elution Buffer (500mM Ammonium formate in 25% ACN, pH 3.0)

320mg Ammonium formate was dissolved in 5 ml distilled water. 2.5ml of 100% ACN was added and pH was adjusted to 3.0 using formic acid. The volume was finally made up to 10ml using distilled water.

Methodology

Two biological replicates of K562 S, S+IM and R cells were used. Whole cell lysate was prepared as mentioned in section 4.3.4.2. To get rid of SDS, buffer exchange with TEAB was carried out in 3KDa cut off spin column (Millipore cat no. UFC5003). The column was equilibrated with 400 µl 0.5M TEAB by centrifuging at 13,000 rpm for 10 min. To 100 µl whole cell lysate 300 µl 0.5M TEAB was added and passed through the Amicon® 3KDa cut off filter cup by centrifuging at 13,000 rpm for 10 min. The flow through was discarded. 300 µl TEAB buffer was added to the filter cup and centrifuged. This was repeated thrice. The buffer exchanged sample in the filter cup was then collected by placing the cup upside down in a clean microcentrifuge and centrifuging at 1000 x g for 2 min. Protein estimation of this sample was carried out by Bradford's method and 30 µg of protein was used for further processing.

In-solution digestion:

To each sample tube containing 30 µg of protein, dissolution buffer (0.5M TEAB) was added to make up the volume to 20 µl. 1 µl denaturant (2% SDS) and 2 µl reducing agent (TCEP) were added to each tube, vortexed to mix, spun down and incubated at 60ºC for 1 hr. After incubation, 1 µl cysteine blocking reagent (MMTS) was added, vortexed, spun down and the tubes were incubated at room temperature for 10 min. Trypsin was added to the tubes in the ratio 1:50 trypsin: protein (w/w), vortexed, spun down and incubated overnight for about 16hrs at 37 ºC.

iTRAQ labelling:

iTRAQ labelling of peptides was carried out using iTRAQ® reagent multiplex kit (Sigma - cat.no. 4352125). The peptide samples were spun down and the iTRAQ labels (114, 115, 116 & 117) were brought to room temperature. Half the contents of each label (~10 µl) were transferred to a fresh tube and 35 µl ethanol was added to each tube, vortexed to mix and spun down. Each label was then transferred to the corresponding sample tube wherein for K562 S (control) vs S+IM (test) comparison group, the first biological replicate of K562 S and S+IM were labelled with 114 and 115 while second replicate with 116 and 117 respectively. In K562 S+IM (control) vs R (test) comparison group, the first set of K562 R and S+IM were labelled with 114 and 115 while the second with 116 and 117 respectively. The peptides and labels were mixed well, spun down and incubated at RT for 1hr. The reaction was stopped by adding 25 µl distilled water, to hydrolyse unbound labels, vortexed and spun down. Labelled peptides from all samples were pooled together and dried in speed vac.

Fractionation using strong cation exchange (SCX) column:

The dried sample was reconstituted in 150 µl loading buffer. pH of the sample was checked using pH strip and formic acid was added to bring down the pH to 3.0, if necessary. The sample was then fractionated using SCX macro spin column (Nest- Cat.no. SMM HIL-SCX).

Conditioning of the column – The column was wetted by adding 500 µl wetting solvent and centrifuging at 300 x g for 2 min. This was followed by 2 washes with 500 µl distilled water. The column was conditioned by adding 500 µl conditioning buffer, spinning it for 10 sec and letting it

stand in the tube for 1 hr. prior to initial use. The solution was then drained by centrifuging and washed with 500 µl distilled water.

Equilibration - 500 µl loading buffer was added to the column and centrifuged at 300 x g for 2 min. This process was repeated thrice.

Sample processing – 150 µl labelled peptide sample was added to the column and centrifuged at 300 x g for 1 min. This was followed by 4 washes with 100 µl loading buffer to remove impurities and traces of detergents if any. Peptides remain bound to the column.

Peptide elution and fractionation - A gradient elution with increasing salt concentration was carried out to elute peptides. The gradient concentration was prepared as mentioned in Table 4.10. The peptides were eluted by adding 100 µl of each gradient concentration and centrifuging at 300 x g for 1 min. Each fraction was collected in a fresh tube and dried in speed vac.

Table 4.10: Gradient concentrations for elution

Gradient concentration	Volume of elution buffer	Volume of loading buffer
50 mM	20 µl	180 µl
100 mM	40 µl	160 µl
150 mM	60 µl	140 µl
200 mM	80 µl	120 µl
250 mM	100 µl	100 µl
300 mM	120 µl	80 µl
400 mM	160 µl	40 µl

Desalting – Ammonium formate used for peptide elution had to be removed before injecting the sample into LC column. Since ammonium formate is a volatile salt, it was removed by repeated reconstitution of each peptide fraction in 50 µl of 0.1 FA and drying it in speed vac. After 3 rounds of this process, each fraction was reconstituted in 16 µl of 0.1% FA. From this 8 µl sample was transferred to HPLC vial from which 6 µl was injected for one LC-MS/MS run.

LC-MS/MS data acquisition for iTRAQ

iTRAQ data acquisition was carried out in IDA mode as mentioned in section 4.3.6.3, using 146 min gradient (Table 4.11) for LC elution. For iTRAQ, the MS/MS scan range was modified to 100-1800 m/z for top 20 precursor ions. The samples were subjected to duplicate LC-MS runs, which were pooled and analysed.

Table 4.11: LC gradient for iTRAQ

Time (min)	%A (0.1% FA in water)	%B (0.1%FA in ACN)
0	95	5
12	90	10
92	70	30
112	50	50
113	20	80
126	20	80
127	95	5
146	95	5

iTRAQ data analysis

The raw spectral file was converted from .mgf to .wiff format using Peakview 1.2 software (Sciex, USA) and was followed by protein identification and quantification in Protein Pilot 4.5 software (Sciex, USA), by pooling the duplicate IDA runs of the sample. The parameters used for this analysis were: Sample type – iTRAQ 4 plex (peptide labelling), Cysteine alkylation – MMTS, Digestion – Trypsin, no special factor was chosen, Species – Homo sapiens, Specify Processing – Quantitate, bias correction and background correction were chosen, ID focus – Biological modification. The search effort was set to 'thorough ID' and false discovery rate (FDR) analysis was enabled. The search was carried out against UniProt database (November 2016 release) containing reviewed human proteins. Proteins identified with 1% FDR were considered for further analysis. Quantitative changes in proteins were obtained as ratio of iTRAQ labels (fold change) with their corresponding p-value. The ratio was obtained for each test and control samples in both the biological replicates, resulting in four

iTRAQ label ratios for each comparison group. Proteins that are differentially expressed with p-value ≤ 0.05 in all ratios in all four groups were qualified as differentiators.

4.3.7. STRING ANALYSIS OF DIFFERENTIATORS

Differentiators identified from SWATH and iTRAQ were pooled and were assessed for their functional association by STRING analysis. The input for STRING was either the gene name or uniport accession number. The settings used for analysis are as follows –

- Organism – Homo sapiens
- Meaning of network edges - confidence
- Active interaction sources – Experiments and databases
- Minimum required interaction score – high confidence (0.7)
- Network display mode – interactive svg
- Display simplification – hide disconnected node in network.

4.3.8. VALIDATION OF DIFFERENTIAL EXPRESSION OF KEY DIFFERENTIATORS BY WESTERN BLOTTING

For detection of 14-3-3 family proteins, p38 and phospho p-38 MAPK, 12% SDS-PAGE (Table 4.12) was used and western blotting carried out as mentioned in section 4.3.4.2 The immunodetection conditions used are mentioned in Table 4.14.

Table 4.12: Composition of 12% SDS-PAGE

6.5 % Stacking gel		**8% Resolving Gel**	
30% Acrylamide	1.08 ml	30% Acrylamide	4 ml
1M Tris pH 6.8	875 µl	1M Tris pH 8.8	3.73 ml
20% SDS	50 µl	20% SDS	67 µl
20% APS	50 µl	20% APS	67 µl
TEMED	2 µl	TEMED	3.33 µl
Distilled water	2.94 ml	Distilled water	2.13 ml
Total Volume	**5 ml**	**Total Volume**	**10 ml**

4.3.9 FUNCTIONAL VALIDATION OF KEY DIFFERENTIATORS

4.3.9.1. 14-3-3 ε knock-out in K562/S cells

LentiCRISPRv1 vector with puromycin selection marker, containing guide RNA for 14-3-3 ε and the empty vector were a kind gift from Dr. Sorab Dalal, ACTREC, India. Prior to initiation of transfection puromycin killing curve was plotted to identify the concentration wherein there was 90% cell death.

Puromycin killing curve – MTT assay

Reagents

- Puromycin working standards:

From 1mg/ml stock, 6 different working standards 1.5 µg/ml , 3 µg/ml, 4.5 µg/ml, 6 µg/ml, 9 µg/ml and 12 µg/ml were prepared.

- MTT reagent:

MTT stock (5mg/ml) was prepared in sterile PBS. The contents were solubilized by placing the tube on rocker and stored at 4ºC.

- Acidified SDS (10% SDS with 0.01N HCl):

43 µl conc. HCl was added to 50 ml of 10% SDS to get acidified SDS.

Methodology

On day 1, 0.3 X 10^6 cells were resuspended in 4ml cRPMI and 100 µl was added to each well of 96 well plate, to get a final cell count of 7500 cells/well. Three wells with 100 µl cRPMI without cells served as blank. Cells were incubated at 37ºC overnight. On day 2, 6 different puromycin working standards (1.5µg/ml, 3µg/ml, 4.5µg/ml, 6µg/ml, 9µg/ml, 12µg/ml) were prepared, from which 20 µl was added to the corresponding wells to get a final concentration of 0.25µg/ml, 0.5µg/ml, 0.75µg/ml, 1µg/ml, 1.5µg/ml and 2µg/ml respectively, with 5 wells (technical replicates) for each concentration. 20 µl of cRPMI was added to blank and control wells. Cells were incubated for 72hrs and cell viability was checked by MTT assay as described in section 4.3.3.1.

Generation of 14-3-3 ε knock-out clones

Transfection:

For transfection, 0.2 X 10^6 K562/S cells were seeded in 35mm tissue culture dish containing 800µl cRPMI without antibiotic and incubated for minimum of 4hrs to overnight. Transfection mix was prepared by adding 3µg (vector control or 14-3-3 ε knockout) plasmid and 3µl Xtreme gene HP transfection reagent (1:1, plasmid: reagent) to 200 µl plain RPMI (devoid of FBS and antibiotic) and incubated for 30 min at RT, after gentle mixing. The transfection mix was then added to the cells dropwise and incubated for 72hrs during which, after 24 hrs 1ml cRPMI without antibiotic was added to the plate.

Generation of stable transfects:

After 72 hrs of incubation, cells were harvested, washed with PBS and subjected to puromycin selection, by growing cells in cRPMI containing 1µg/ml puromycin (cRPMI + puro). The cells were passaged every 72hrs and maintained for at least 6-7 passages before clonal selection.

Single cell sorting for clonal selection:

Once the viability of stably transfected cells reaches about 95%, 0.1 X 10^6 cells were resuspended in 300µl cRPMI + puromycin and subjected to single cell sorting wherein the sorted cells were collected into each well of 96 well plate (U-bottom) containing 150 µl of cRPMI + puromycin. The plate was incubated at 37^{o}C until single cells grow to form visible cell cluster. The clones were then transferred to 6 well plate and scaled up further by growing in 60mm plates.

Screening and validation of 14-3-3 ε knockout clones:

The vector control and knockout clones were screened by assessing the protein level of 14-3-3 ε by western blotting as mentioned in section 4.3.8. Three vector control and knockout clones were chosen for further analysis. To validate the knockout of 14-3-3ε, DNA from two representative vector control and 14-3-3ε KO clones was isolated. PCR amplification for 14-3-3ε was carried our using the primers listed in table 4.13. The PCR products were then cloned in pTZ57R/T (Thermo Scientific-K-1214). Positive clones were sequenced and validated using DNASTAR Lasergene software.

Table 4.13: Primers for PCR amplification of 14-3-3ε

14-3-3ε Fwd	-TTGCCATAGAGCTGAGCAGT-3'
14-3-3ε Rev	-TCACATTCCAGGGCATAGAGC-3'

Determination of effect of 14-3-3 ε knockout on IC_{50} of IM

To assess the effect of 14-3-3 ε knockout on IC_{50} of IM, MTT assay was carried out as mentioned in section 4.3.3.1, with three 14-3-3 ε knockout and vector control clones.

4.3.9.2 Inhibition of p38-MAPK activity in K562/R

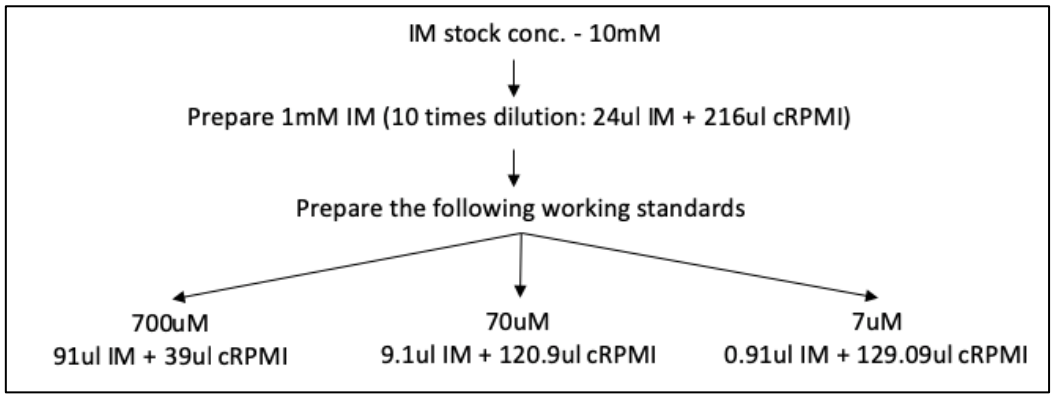
The Ser/Thr kinase activity of p38-MAPK was inhibited by maintaining cells in media containing 10 µM inhibitor SB20358 (calbiochem - cat no. 559389) for 1hr. The inhibition was confirmed by assessing the phosphorylation status of its downstream substrate msk by western blotting. The immunodetection conditions used are mentioned in Table 4.14

Effect of inhibition of p38 activity on IC_{50} of IM

Reagents

- IM working standards:

From 10mM IM stock, different working standards were prepared as mentioned below.



- p38 inhibitor:

From 25mM stock, 70µM working standard was prepared.

- Vehicle control (DMSO):

DMSO was diluted 10 times with cRPMI and from this 140 µl was added to 60 µl cRPMI (to mimic the dilution volumes used for 700uM IM solution), to prepare DMSO working standard.

- MTT reagent:

5mg/ml MTT stock was prepared in sterile PBS. The contents were solubilized by placing the tube on rocker and stored at 4ºC.

- Acidified SDS (10% SDS with 0.01N HCl):

43 µl conc. HCl was added to 50 ml of 10% SDS to get acidified SDS.

Methodology:

On day 1, two set of tubes with 3ml cRPMI containing 0.54 X 10^6 K562/R cells were prepared. 100 µl cell suspension from each tube was added to 20 wells of 96 well plate to get a final cell count of 18000 cells/well, creating 2 panel (20 wells each) of cells, in which cells of panel 1 would be treated only with IM and those of panel 2 with IM as well as p38 inhibitor (p38i). 3 wells with 100 µl cRPMI without cells served as blank and 5 wells in each panel with untreated K562/R cells served as control. Cells were incubated at 37ºC overnight. On day 2, from 70µM p38i working standard 20µl was added to all wells of panel 2 except control, to get a final concentration of 10µM. This was followed by addition of 20 µl of IM working standards (7µM, 70µM and 700µM) to the corresponding wells in both panels to get a final concentration of 1 µM, 10 µM and 100 µM IM respectively, with 5 wells (technical replicates) for each concentration. Since IM and p38i were solubilized in DMSO, 20 µl of DMSO working standard was added to blank and control wells of panel 1 while 40 µl was added to control wells of panel 2. Cells were incubated for 1hr and 30 µl MTT reagent was added to each well of panel 1 and 35 µl MTT reagent to each well of panel 2 (since it has additional 20 µl volume from p38i), to get a final concentration of 1mg/ml. Cells were then incubated for 4 – 6 hrs. to allow formazan crystal formation. The crystals were solubilized with 100 µl acidified SDS, incubated overnight at 37ºC. Absorbance at 570nm was then recorded using ELISA plate reader.

4.3.10. STATISTICAL ANALYSIS

Statistical analysis was carried out using IBM SPSS statistics software 21 and Graphpad Prism 5. All experiments were carried out in biological triplicate. Student's t-test was used to compare two groups of data and p-value ≤ 0.05 was considered statistically significant.

Table 4.14: Immunoblotting conditions for detecting proteins

Antibody	Blocking	Primary antibody condition	Washes	Secondary antibody condition	Washes
STAT-5 (Life Technologies–335900)	5% BSA in 0.1% TBST – 1 hr RT	1:1000 - 5% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.M. – 1:1000 - 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
pSTAT-5 (Abcam-ab32364)	5% BSA in 0.1% TBST – O/N, 4° C	1:1000 - 5% BSA in 0.1% TBST – 1 hr (Pre-probed)	4 x 10 mins – 0.1% TBST	A.R. – 1:2000 - 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
c-abl (Abcam-ab85947)	5% BSA in 0.1% TBST – 1 hr RT	1:500 - 5% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.R. – 1:2000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
* P-Glycoprotein (Abcam-ab170904)	5% BSA in 0.1% TBST – 1 hr RT	1:250 - 5% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.R. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
OCT-1 (Abcam-ab181022)	5% BSA in 0.1% TBST – 1 hr RT	1:1000 - 5% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.R. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
14-3-3 ε (Santacruz-sc-23957)	5% BSA in 0.05% TBST – 1 hr RT	1:250 - 5% BSA in 0.05% TBST – O/N, 4° C	4 x 10 mins – 0.05% TBST	A.M. – 1:1000 – 1% BSA in 0.05% TBST – 1hr RT	4 x 10 mins – 0.05% TBST

14-3-3 γ (Santacruz-sc-398423)	5% BSA in 0.1% TBST – 1 hr RT	1:1000 - 5% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.M. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
p38-MAPK (CST-9212)	5% milk in 0.1% TBST – 1 hr RT	1:2000 - 1% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.R. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
p-p38 MAPK (CST-9211)	5% milk in 0.1% TBST – 1 hr RT	1:2000 - 1% BSA in 0.1% TBST – O/N, 4° C	4 x 10 mins – 0.1% TBST	A.R. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
Msk (Abcam–ab99412)	5% BSA in 0.1% TBST – 1 hr RT	1:1000 - 1% BSA in 0.1% TBST – O/N, 4°C	4 x 10 mins – 0.1% TBST	A.R. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST
p-msk (Millipore-04-384)	5% BSA in 0.1% TBST – 1 hr RT	1:4000 - 1% BSA in 0.1% TBST – O/N, 4°C	4 x 10 mins – 0.1% TBST	A.R. – 1:1000 – 1% BSA in 0.1% TBST – 1hr RT	4 x 10 mins – 0.1% TBST

* - For P-Glycoprotein detection, the whole cell lysate prepared was loaded on to the gel without boiling as per antibody datasheet guidelines.

5. RESULTS

5.1. SYSTEMATIC REVIEW AND META-ANALYSIS OF PROTEOMICS -BASED CML STUDIES

5.1.1. SYSTEMATIC REVIEW

The search strategy employed in Pubmed's advanced search tool identified 284 studies (Table 5.1). Abstract level selection of these articles based on the inclusion criteria (mentioned in section 4.2.1 of materials & methods) resulted in 22 studies. This was followed by full-text screening wherein studies that (a) did not involve global profiling (b) used murine cell line transfected with wildtype or mutant Bcr-Abl (c) used cells resistant to drugs other than imatinib, were eliminated thereby resulting in 9 most relevant studies [87-90, 93-95, 97, 99] (Fig 5.1).

Table 5.1: Number of studies identified using the search strategy

Search	Query	Items found
#6	Search (#4 OR #5)	**284**
#5	Search (#1 AND #3)	175
#4	Search (#1 AND #2)	163
#3	Search Mass spectrometry	328917
#2	Search (Proteome) OR Proteomic	119569
#1	Search ((Chronic Myeloid leukemia [MeSH Terms]) OR BCR-ABL) OR BCR/AB	23789

Studies with similar comparison groups were grouped together as: (a) those that involve comparison of cell lines with and without IM treatment (+/- IM) (Table 5.2) and (b) those that involve comparison of imatinib sensitive and resistant cells (S vs R) (Table 5.3), which help in understanding Bcr-Abl downstream signalling and molecules altered in resistance respectively, thereby contributing to the identification of therapeutic targets other than Bcr-Abl to overcome IM resistance in CML. Of these, only one study from +/- IM comparison group has reported an alternate therapeutic strategy for CML-BC, i.e. inhibition of hypusination of eIF5A [90]. Similarly, of the five studies included in S vs R comparison group, two studies have identified either a key molecule - HSP70 [95] or biological

process - calcium homeostasis [97] contributing to IM resistance. The proteomic profile generated in other studies did not identify or report novel alterations which can improve the understanding of Bcr-Abl downstream signalling or pathways associated with resistance. In order to explore the data from these studies in depth, we carried out meta-analysis.

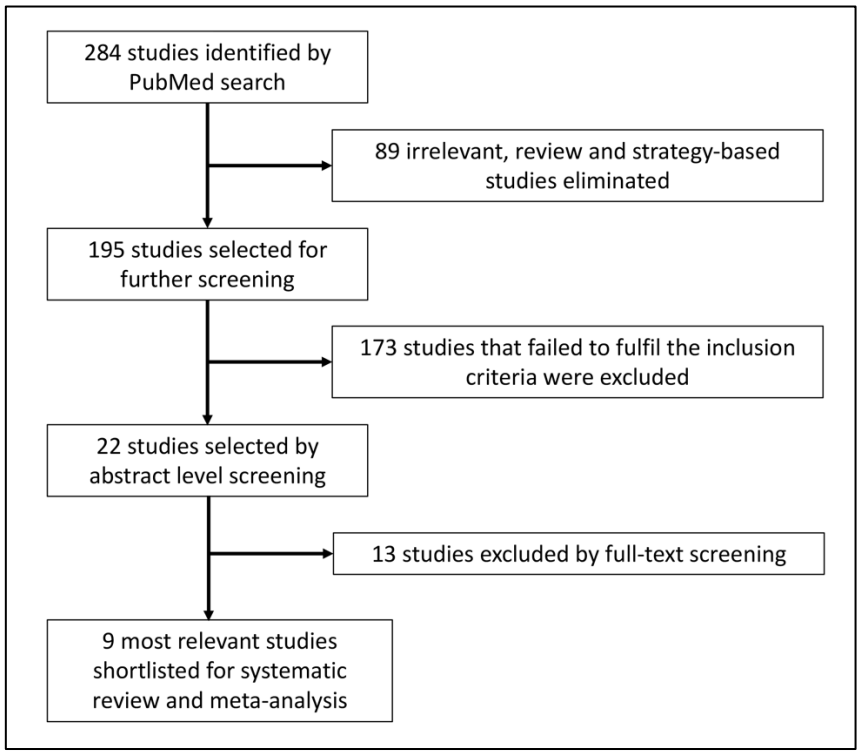


Fig.5.1. Study selection flowchart for systematic review and meta-analysis: *9 most relevant studies finalized by full-text screening were chosen for further analysis.*

5.1.2. META-ANALYSIS

In comparison group with cells +/- IM (Table 5.2), all five studies have been carried out using K562 cell line thereby enabling comparison of these studies, three of which have employed 2D-MS approach and two have employed LC-MS based SILAC approach. In case of S vs R comparison group of the five studies (Table 5.3), consistency with respect to cell lines used was observed only in two studies where K562 has been used. Hence only these two studies were considered for analysis. A data extraction sheet with differentiators identified from each study was prepared for both the comparison groups (Appendix 1: Table A-1 & A-2).

Table 5.2: Studies that involve comparison of cells with and without IM treatment (+/- IM comparison group)

S.No.	Article	Year	Cell line	Proteomic analysis used	Differentiators identified
1.	Identification of differentially expressed proteins in imatinib mesylate-resistant chronic myelogenous cells.	2005	K562	2D - MS	38
2.	Hypusination of eukaryotic initiation factor 5A (eIF5A): a novel therapeutic target in BCR-ABL-positive leukemias identified by a proteomics approach	2007	K562	2D MS	19
3.	A systems biology understanding of the synergistic effects of arsenic sulfide and Imatinib in BCR/ABL-associated leukemia.	2008	K562	2D-MS	46
4.	Global proteome quantification for discovering imatinib-induced perturbation of multiple biological pathways in K562 human chronic myeloid leukemia cells.	2010	K562	LC-MS(SILAC)	73
5.	Proteome Changes Induced by Imatinib and Novel Imatinib Derivatives in K562 Human Chronic Myeloid Leukemia Cells.	2014	K562	LC-MS (SILAC)	7

Table 5.3: Studies that involve comparison of IM-sensitive and resistant cells (S vs R comparison group)

S.NO.	Article title	Year	Cell line	Proteomic analysis used	No. of differentiators
1.	Identification of differentially expressed proteins in imatinib mesylate-resistant chronic myelogenous cells.	2005	K562	2D - MS	34
2.	Comparative proteomic analysis of chronic myelogenous leukemia cells: inside the mechanism of imatinib resistance.	2006	LAMA 84	2D-MS	44
3.	Proteomic analysis of an imatinib-resistant K562 cell line highlights opposing roles of heat shock cognate 70 and heat shock 70 proteins in resistance.	2008	K562	2D-MS	24
4.	Gaining insights into the Bcr-Abl activity-independent mechanisms of resistance to imatinib mesylate in KCL22 cells: a comparative proteomic approach.	2010	KCL22	2D- MS	34
5.	Proteomic analysis of imatinib-resistant CML-T1 cells reveals calcium homeostasis as a potential therapeutic target.	2016	CML-T1	2D-MS	8

5.1.2.1. Identification of common differentiators among studies

To identify which molecules are differentially expressed consistently across studies upon IM treatment or IM-resistance in K562, the differentiators from studies within each comparison group were assessed for their overlap using Venn diagram. In K562 +/- IM comparison group, Venn diagram (Fig 5.2.A) revealed Rho GDP dissociation inhibitor as common to studies 1 and 2 while eukaryotic

translation initiation factor 2 subunit 1 as common to studies 3 and 4. In K562 S vs R comparison group on the other hand, only heterogeneous nuclear ribonucleoprotein A/B (hnRNP-A/B) was found to be common to both the studies (Fig 5.2.B).

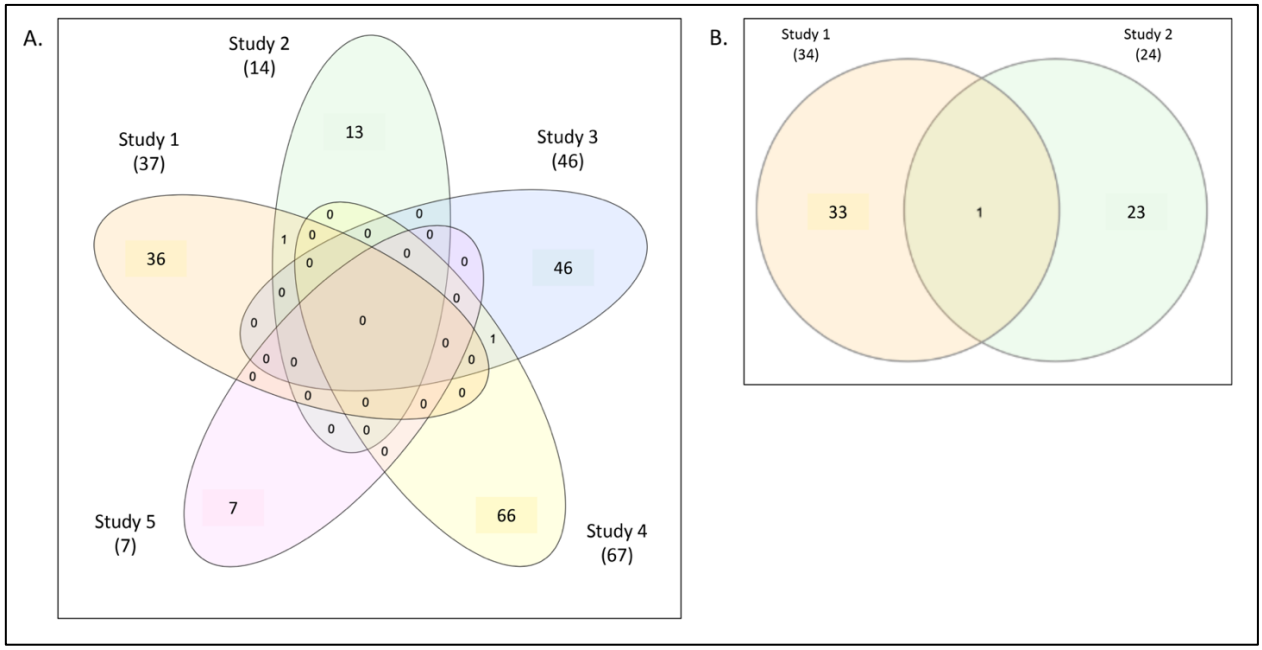


Fig.5.2. Venn diagram to assess overlap of differentiators among studies: *(A) Studies comparing K562 cell line with and without IM treatment. No differentiator was common to all the studies (B) Studies comparing K562 cells sensitive and resistant to IM. Only hnRNP-A/B was common to both the studies.*

5.1.2.2. STRING analysis of differentiators

Further clarity on the components of Bcr-Abl pathway was obtained when differentiators identified in studies within each comparison group were pooled and subjected to STRING analysis. STRING analysis of differentiators from K562 +/- IM comparison group revealed multiple network clusters of proteins belonging to functional groups like cell cycle, nucleosome formation, mRNA splicing and stability, ribosome biogenesis and protein synthesis, mRNA/protein nuclear import and export, protein folding, protein degradation, apoptosis, cytoskeletal organization, energy metabolism, erythroid differentiation and signalling (Fig.5.3). Further to understand which of these altered proteins functionally associated with Bcr-Abl, Abl was introduced as one of the inputs in STRING (as Bcr/Abl is not an annotated SWISS-PROT entry) and Abl was found to be associated majorly with cell cycle

proteins. The data from meta-analysis contributes moderately towards understanding the Bcr-Abl downstream pathways. However, more in-depth analysis would help essential to identify novel therapeutic targets.

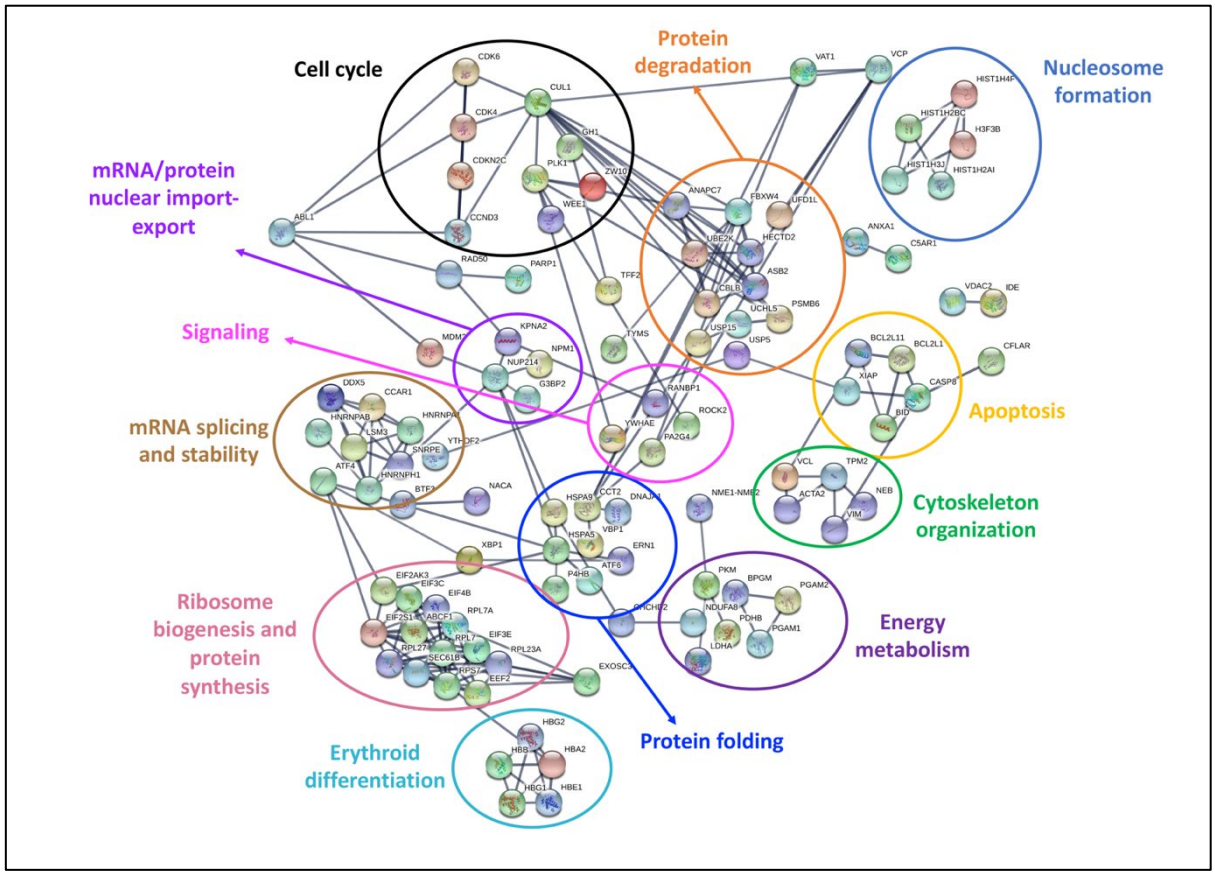


Fig.5.3. STRING analysis of differentiators from K562 +/- IM comparison group. *Functional association among proteins provided some information on pathways altered by Bcr-Abl.*

In STRING analysis of differentiators from K562 S vs R comparison group (Fig.5.4.), very few functional associations were observed among proteins. The pathways uniquely altered in resistant cells are yet to be explored and meta-analysis does not contribute to this knowledge.

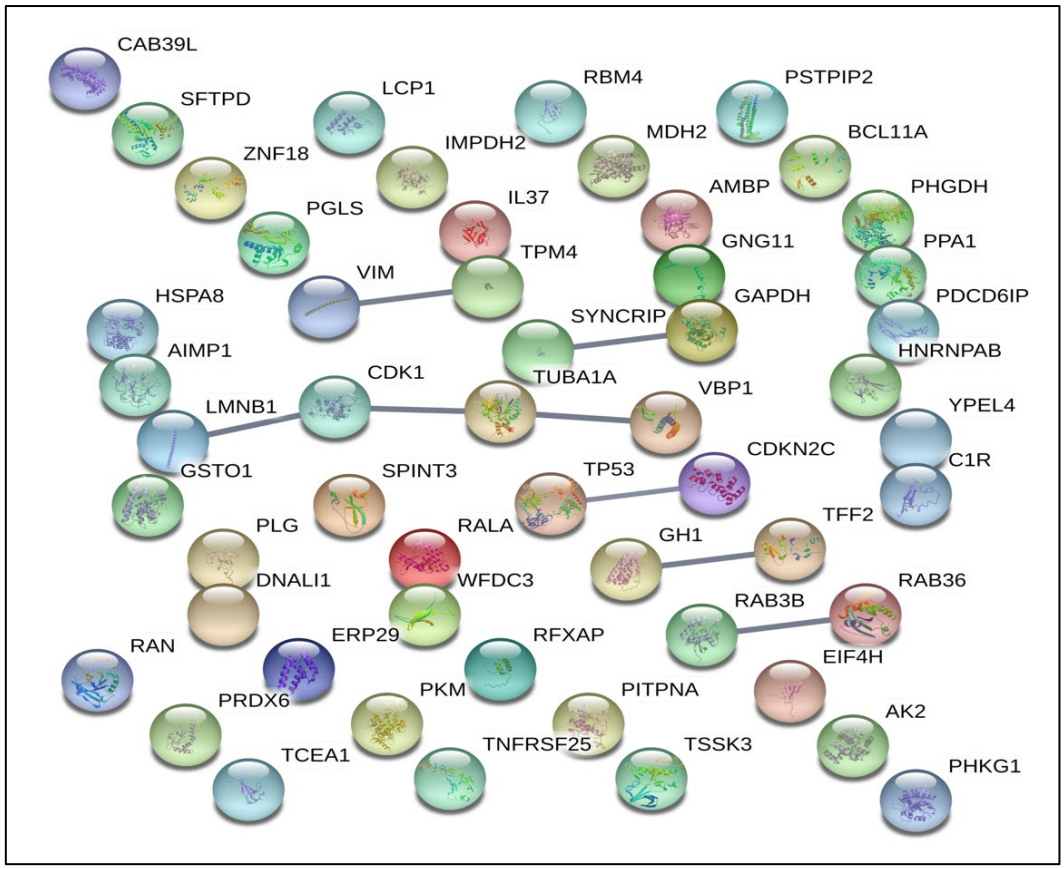


Fig.5.4. STRING analysis of differentiators from K562 S vs R comparison group. *Very few proteins were functionally associated and most differentiators remained disconnected providing no information regarding pathways contributing to IM-resistance.*

In summary, systematic review and meta-analysis of proteomic studies on CML-BC cell lines revealed paucity of information to delineate either components of Bcr-Abl downstream signalling pathway or non-Bcr-Abl mediated signalling pathway that contribute to IM-resistance, so as to identify alternate therapeutic targets for IM-resistant CML.

5.2. ESTABLISHMENT OF APPROPRIATE BIOLOGICAL SYSTEM AND CONDITIONS FOR THE STUDY

5.2.1. BASIC CHARACTERIZATION OF CELL LINES

The morphology of all the three CML-BC cell lines – K562, KCL22 and KU812 is shown in Fig.5.5 and their doubling time was determined as 20 hrs, 23 hrs and 21 hrs respectively. The time point at which cells remain in the mid log phase with more than 90% viability, was chosen as the optimal harvest time. Based on growth curve and viability, it was identified as 48hrs for K562 and 72 hrs for KCL22 and KU812 cell lines (Fig.5.6). Cells were harvested at these time points for all further biological assays.

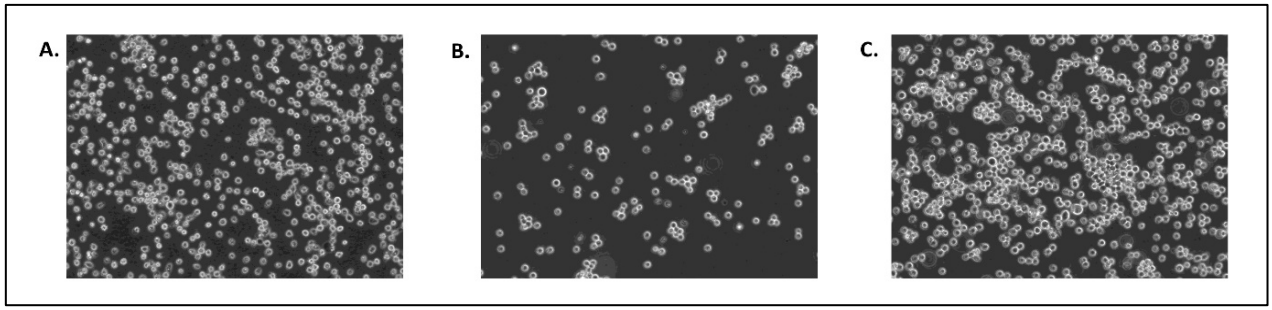


Fig.5.5. Morphology of CML-BC cell lines*: (A) K562 (B) KCL22 (C) KU812 (10X magnification).*

These wild type CML-BC cell lines were termed sensitive to Imatinib (S) and designated as K562/S, KCL22/S and KU812/S respectively. Their IC_{50} of imatinib was determined based on MTT assay, which was found to be 0.5µM for K562/S and KU812/S while 0.3 µM for KCL22/S with 48hrs. treatment (Fig. 5.7).

5.2.2. DEVELOPMENT OF IM-RESISTANT CELL LINES

IM-Resistant (R) cell lines K562/R, KCL22/R and KU812/R were developed from their sensitive counterpart by gradual IM dose escalation. Their resistance to IM was confirmed by MTT assay (Fig.5.8), which indicated that at IC_{50} of sensitive cells, the corresponding resistant cells showed ~80% viability. Further, IC_{50} of IM for K562/R, KCL22/R and KU812/R were found to be 5 µM, 2

µM and 4 µM respectively upon 48hr treatment, which is about 6-10-fold higher than that of sensitive cells. Once developed, the resistant cells were always maintained under a constant drug pressure of 0.75µM IM for K562/R and 1µM IM for KCL22/R and KU812/R cells, the concentration at which cells could be cultured for long term without loss of viability and resistance.

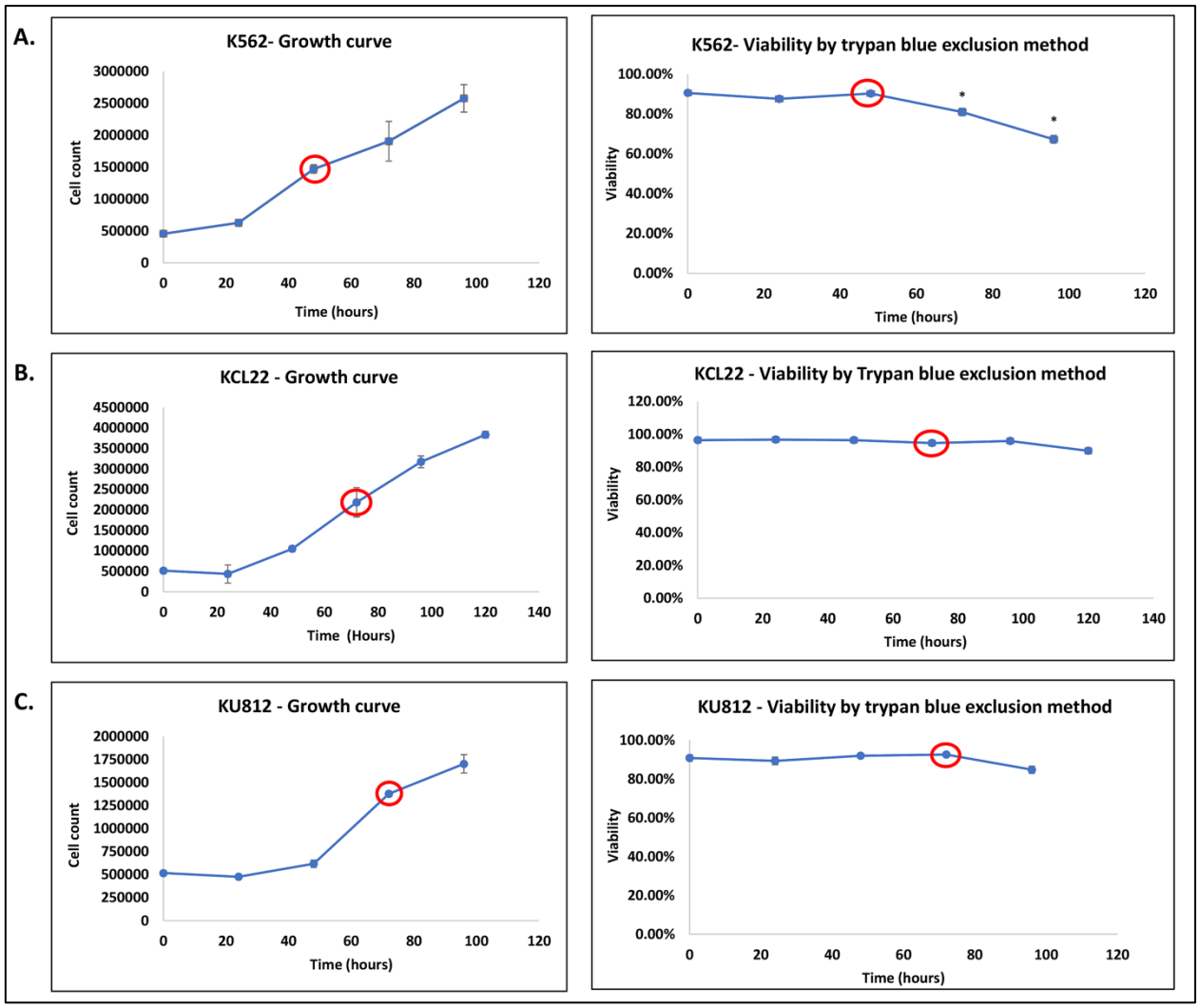


Fig.5.6. Optimal harvest time: *Based on growth curve and viability the optimal harvest time was identified as 48hrs for (A) K562 and 72hrs for (B) KCL22 (C) KU812 cell lines. The experiment was carried out in biological triplicate. The graph represents mean ± SEM. Statistical significance was assessed by ANOVA, (*- p value <0.05).*

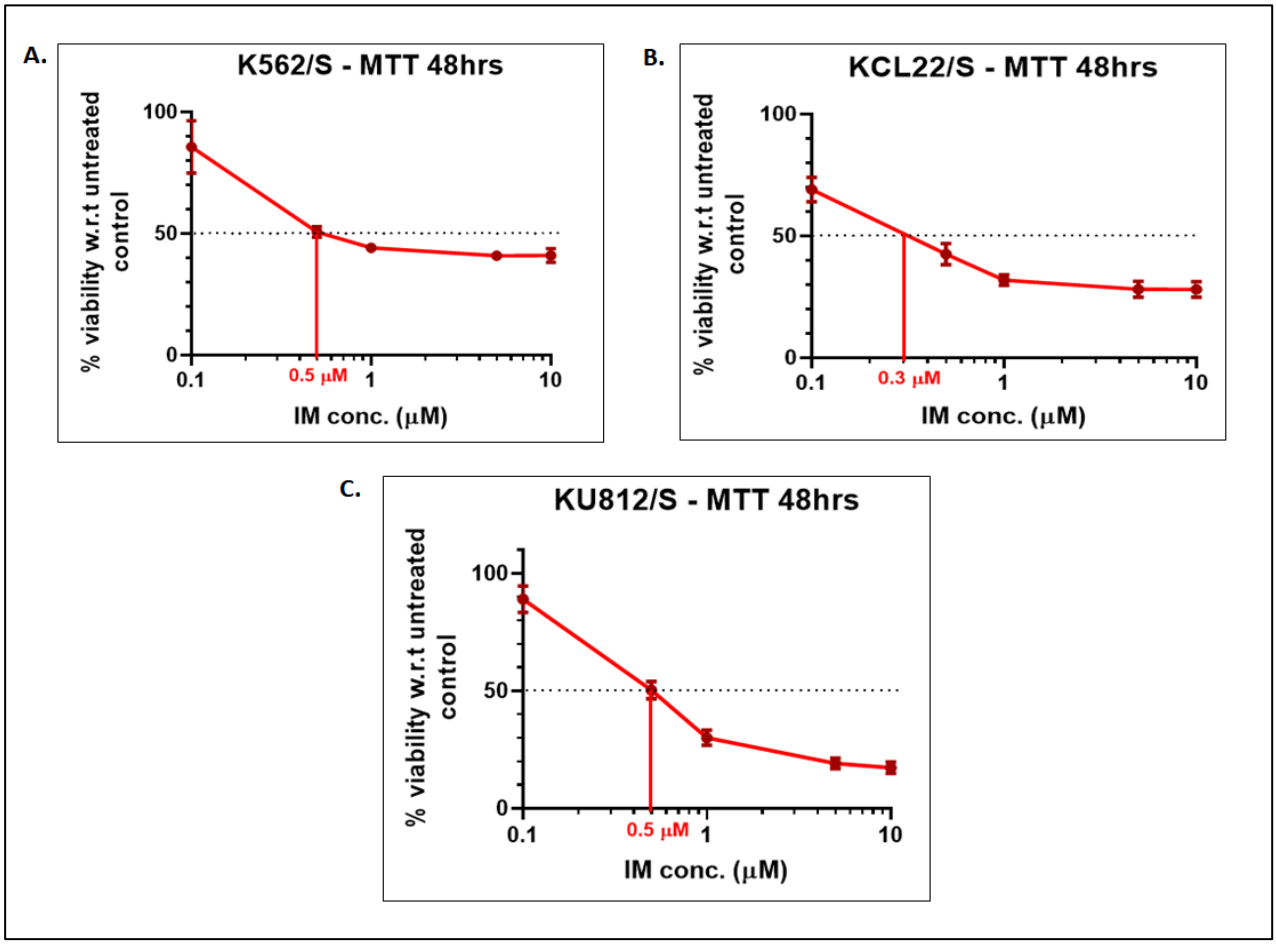


Fig. 5.7. Determination of IC_{50} of IM by MTT assay: *48hr IM treatment resulted in IC_{50} of 0.5 µM for (A) K562/S and (C) KU812/S; 0.3 µM for (B) KCL22/S cell line. All experiments were carried out in biological triplicates. The graph represents mean ± SEM.*

5.2.3. OPTIMIZATION OF IM-TREATMENT CONDITION FOR COMPARATIVE EVALUATION OF IM- SENSITIVE AND RESISTANT CELLS

An optimal IM treatment condition is that IM concentration and treatment duration at which Bcr-Abl activity is maximally inhibited without compromising on cell viability. Since the study demands comparison of IM-treated sensitive (S) and resistant (R) cells, it necessitates treating sensitive cells with the same IM concentration used to maintain the corresponding resistant cells. So, K562/S should be treated with 0.75 µM IM, while KCL22/S and KU812/S with 1 µM IM. However, from Fig.5.7 it is evident that treating sensitive cells with these IM concentration for 48hrs would result in >50% cell

death. Hence, a treatment duration less than 48hrs, where cell viability is maintained as well as Bcr-Abl activity is inhibited, had to be identified.

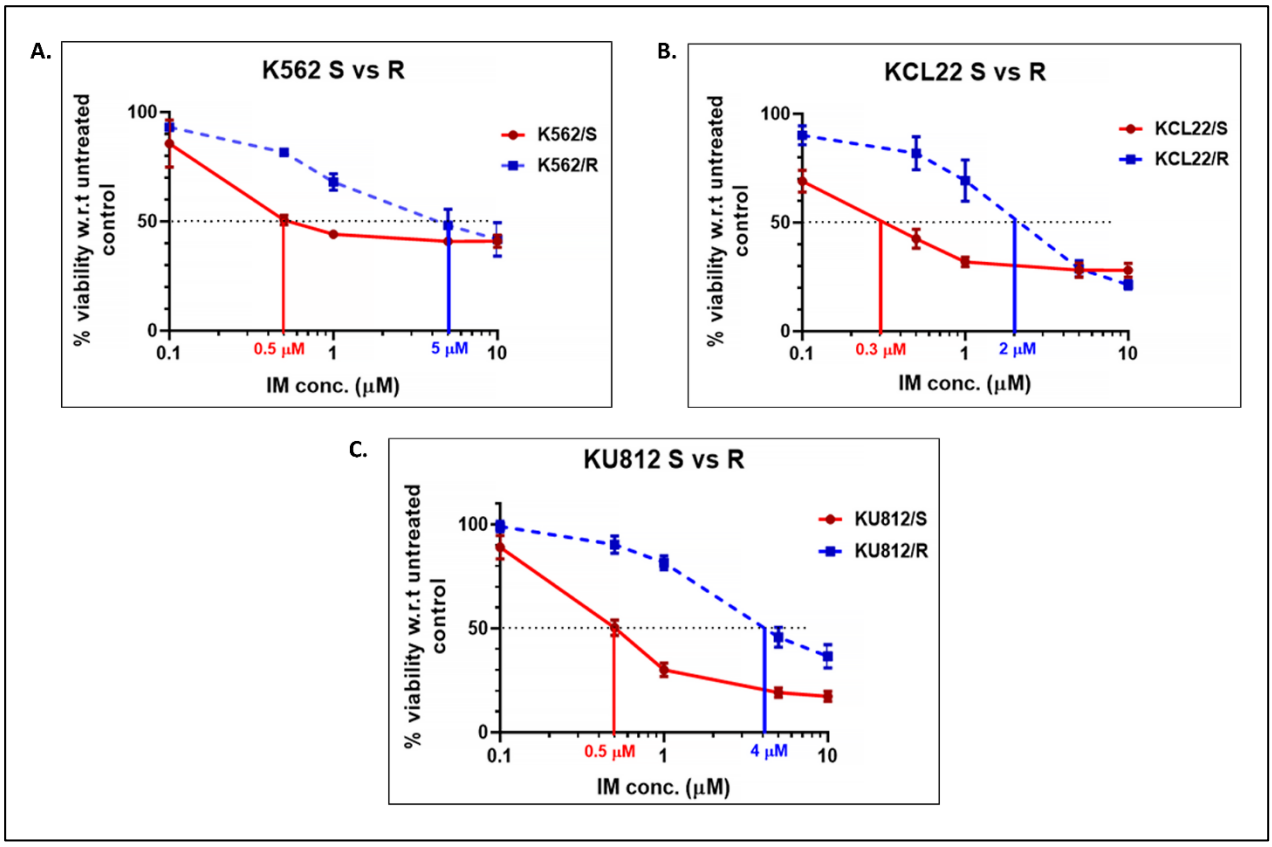


Fig.5.8. MTT assay to confirm resistance: *(A) K562/R (B) KCL22/R (C) KU812/R cells show about 6-10-fold increase in IC50 compared to their sensitive counterpart confirming resistance to IM. All experiments were carried out in biological triplicates. The graph represents mean ± SEM.*

Viability of K562/S cells treated with 0.75 µM IM, KCL22/S and KU812/S cells treated with 1 µM IM for 2, 6, 12, 18 and 24 hrs. were found to be unaffected up to 24 hr. treatment in all the 3 cell lines (Fig.5.9), indicating that any treatment duration up to 24 hrs with inhibition of Bcr-Abl activity could be considered optimal. Bcr-Abl tyrosine kinase activity in all the three cell lines showed maximum inhibition at 12 hr. treatment with IM (Fig.5.10).

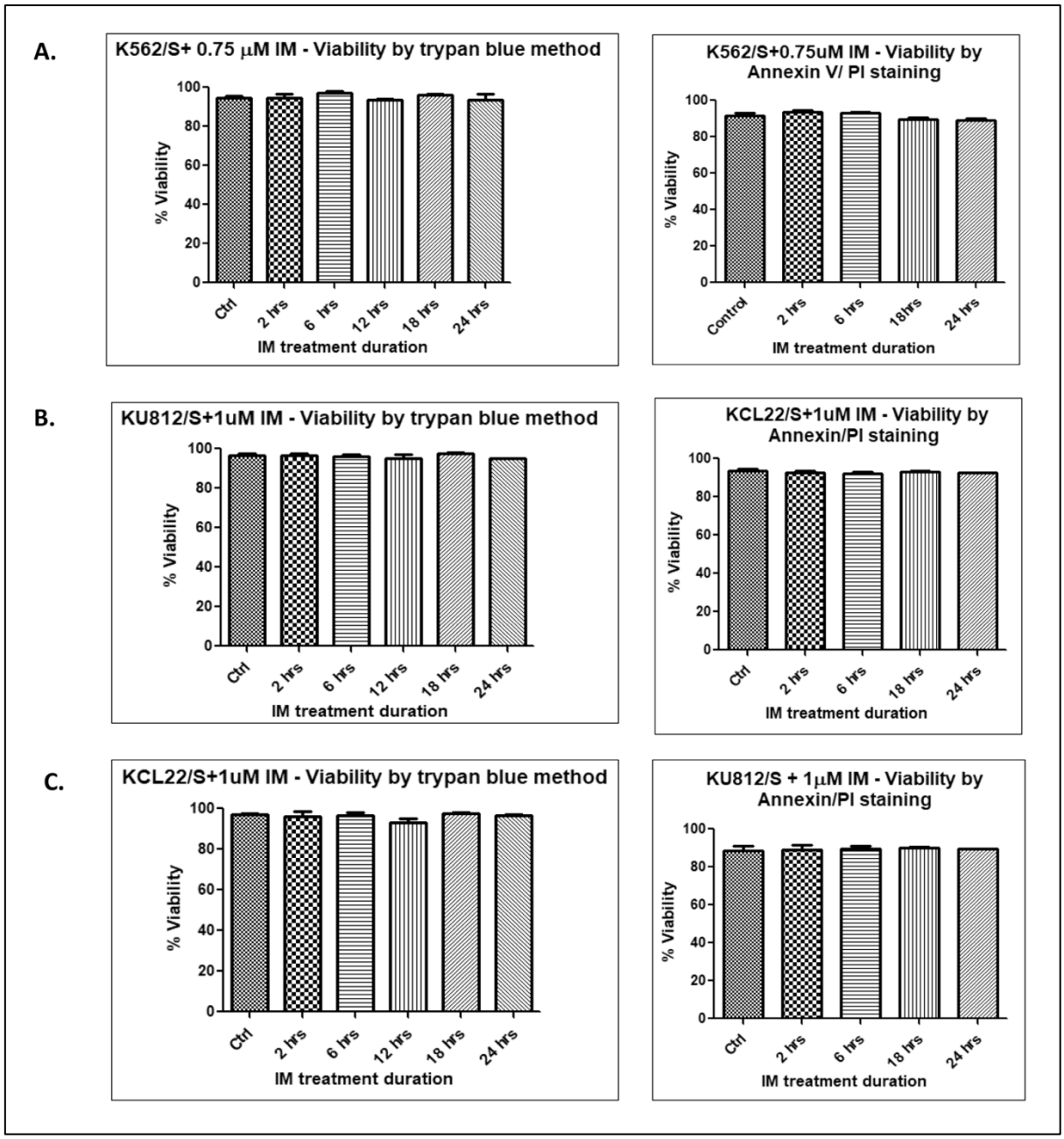


Fig. 5.9. Assessment of cell viability upon IM treatment by trypan blue dye exclusion method and Annexin V-FITC/PI staining: *No significant reduction in viability was observed in (A) K562/S (B) KCL22/S and (C) KU812/S up to 24 hr IM treatment. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by ANOVA.*

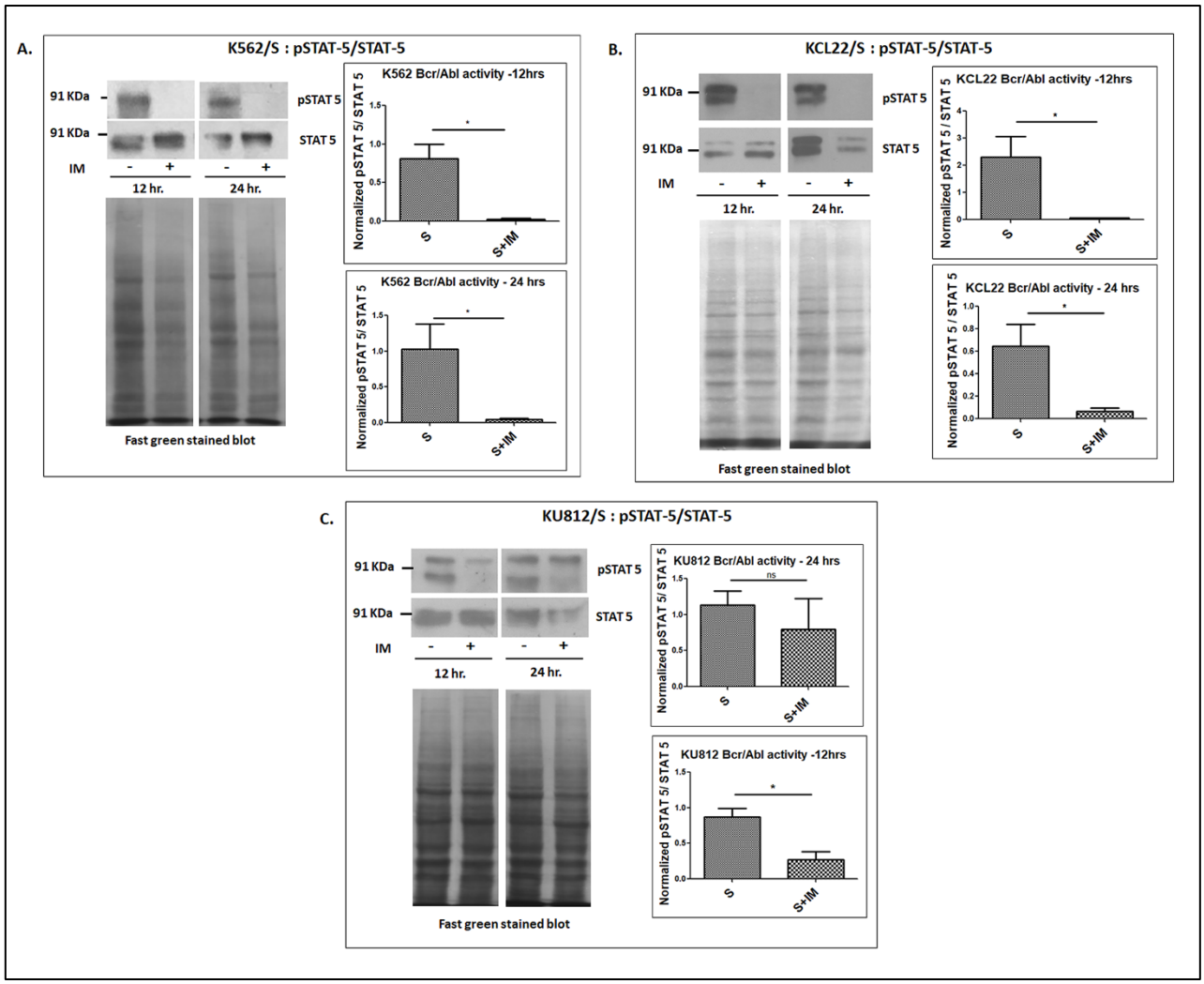


Fig. 5.10. Assessment of Bcr-Abl tyrosine kinase activity by western blotting: *The ratio of p-STAT5/STAT5 revealed that Bcr-Abl activity was significantly inhibited at 12hr as well as 24hr IM treatment in (A) K562/S and (B) KCL22/S cells while only 12hr treatment showed significant reduction in (C) KU812/S cells. STAT5 and p-STAT5 band intensities were normalized with their corresponding whole lane intensity. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05).*

In summary for comparative evaluation of S and R cells, treatment of K562/SR cells with 0.75 µM IM and KCL22/SR, KU812/SR with 1 µM, for a duration of 12 hrs. was found to be optimal.

5.3. STATUS OF KNOWN MECHANISMS OF IM RESISTANCE IN CML-BC CELL LINES

To understand the cause for IM-resistance, IM-sensitive and resistant K562, KCL22 and KU812 cells were screened for the presence of established mechanisms of resistance which includes (a) Bcr-Abl gene amplification and overexpression (b) Bcr-Abl kinase domain mutation and (c) altered drug transport.

5.3.1. *BCR-ABL* GENE AMPLIFICATION

Interphase FISH of K562/S and R as well as KU812/S and R cells revealed *BCR-ABL* gene amplification, with 8-10 copies of the fusion gene in both cell types, which was further confirmed by metaphase FISH (Fig 5.11.A & C). In case of KCL22/S and R interphase and metaphase FISH indicated presence of three copies of BCR-ABL fusion in both cell types (Fig. 5.11.B). Thus, with respect to *BCR-ABL,* the chromosomal abnormalities detected by FISH remained consistent between the sensitive and resistant counterpart of each cell line, with no further gene amplification in resistant cells. However, western blotting to assess the level of Bcr-Abl protein in untreated sensitive (S), sensitive treated with IM (S+IM) and resistant cells (R) revealed a significant increase in K562/R cells compared to that of K562/S and S+IM (Fig. 5.12.A). However, no significant increase was observed in KCL22/R and KU812/R cells (Fig.5.12. B-C).

5.3.2 KINASE DOMAIN MUTATIONS

The resistant cells were screened for the presence of kinase domain mutations using a sequencing-based approach. A G250E mutation from glycine to glutamate was identified in KCL22/R cells (Fig.5.12. D). No kinase domain mutation was detected in K562/R and KU812/R cells.

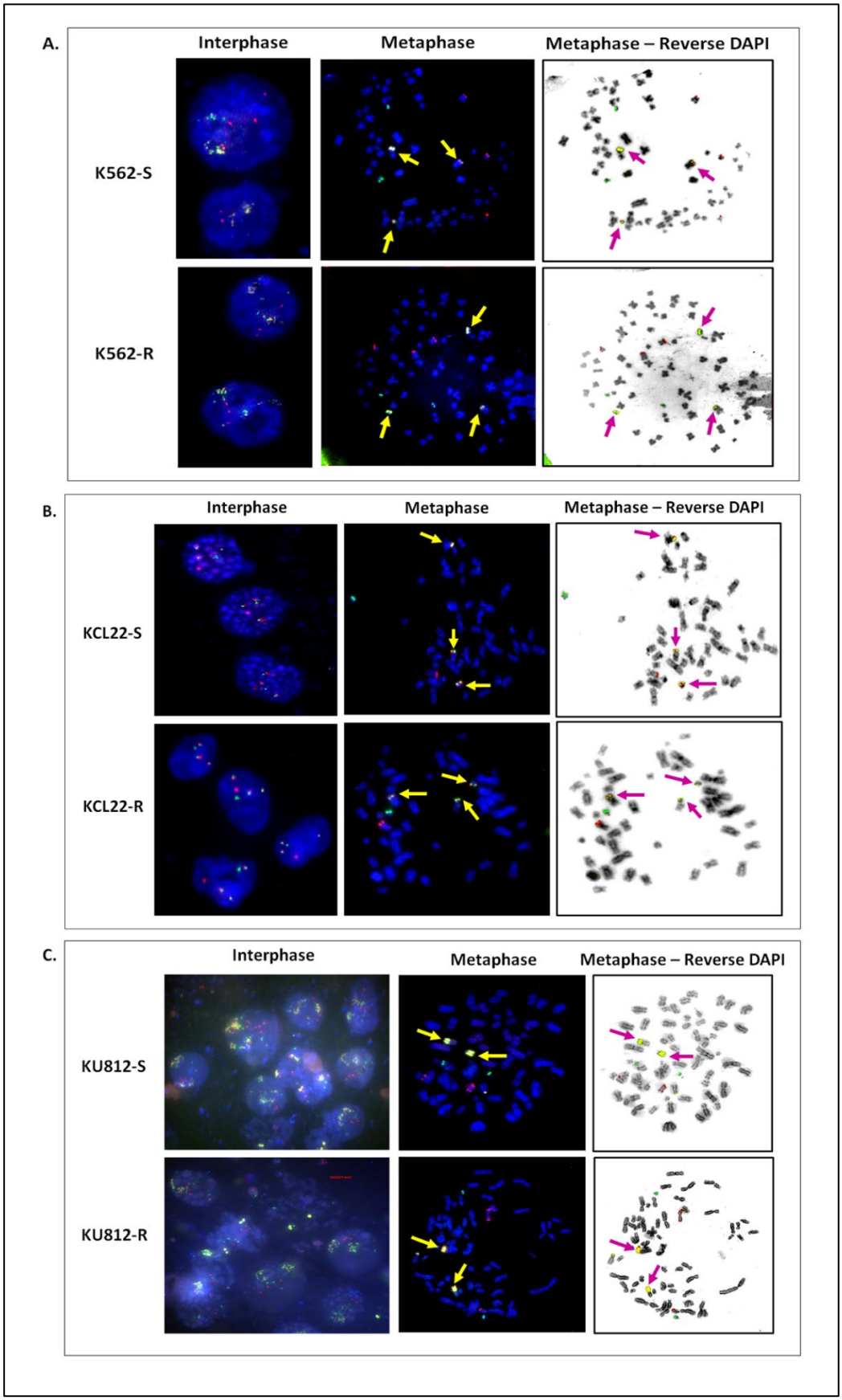


Fig.5.11. Assessment of BCR-ABL gene amplification by FISH: *(A) Interphase cells of K562/S and R showing 8-10F3R2G indicating 8-10 copies of BCR-ABL1 fusion, 3 copies of ABL1 (Red signals)*

and 2 copies of BCR (Green signals). Metaphase cell showing amplified copies of BCR-ABL1 fusion on 2 copies of der(22), BCR-ABL1 fusion on chromosome 9. Reverse-DAPI image of the same metaphase which confirms complex rearrangements. (B) Interphase cells of KCL22/S and R showing 3F1R1G indicating 3 copies of BCR-ABL1 fusion, 1 copy of ABL1 and 1 copy of BCR. Metaphase showing BCR-ABL1 fusion on three copies of der(22) indicating duplication of Ph chromosome. Reverse-DAPI image of the same metaphase which confirms complex rearrangements. (C) Interphase cells of KU812/S and R showing 8-10F3R2G indicating 8-10 copies of BCR-ABL1 fusion, 3 copies of ABL1 and 2 copies of BCR. Metaphase showing amplified copies of BCR-ABL1 fusion on der(9) and der(22). Reverse-DAPI image of the same metaphase which confirms complex rearrangements. All images were acquired at 100X magnification.

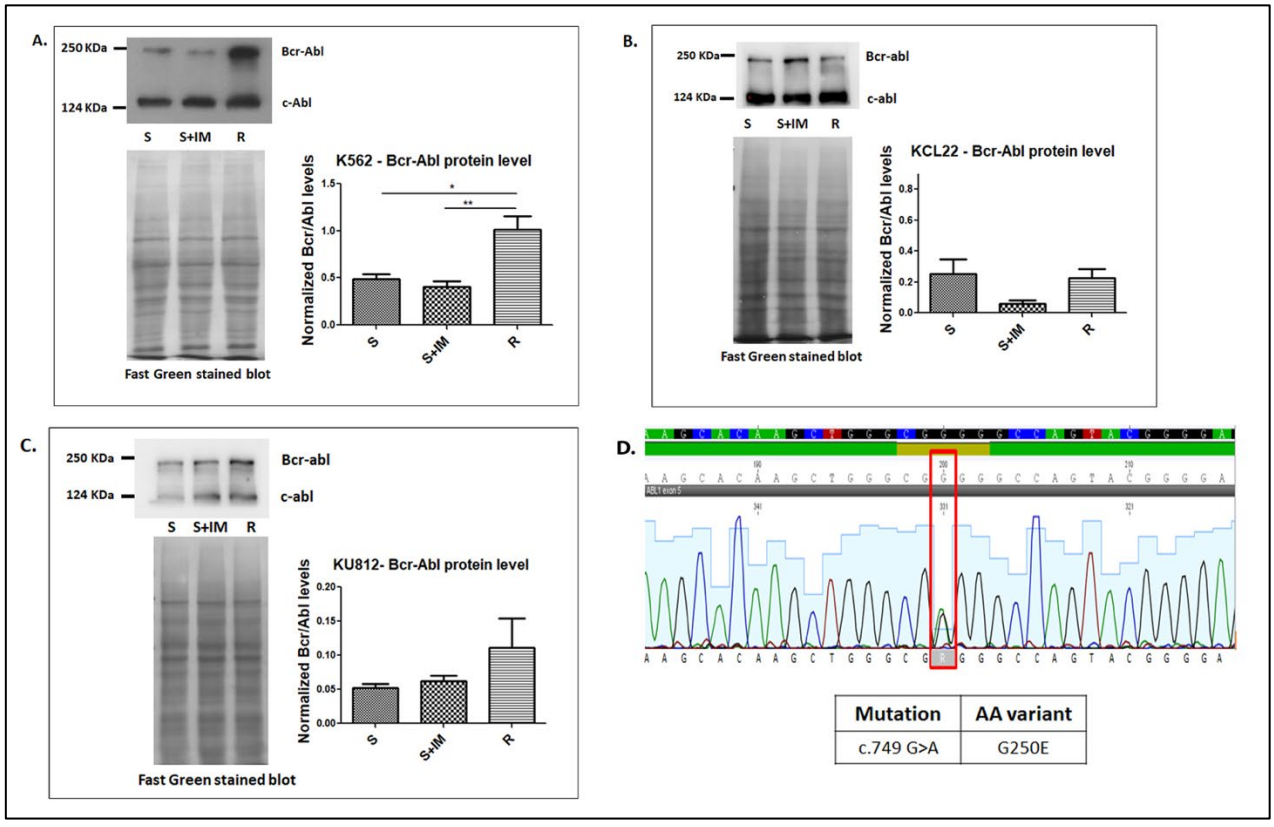


Fig. 5.12. Assessment of Bcr-Abl overexpression and kinase domain mutation: *(A-C) Western blotting analysis of Bcr-Abl in (A) K562, (B) KCL22, (C) KU812 – S, S+IM and R cells indicating a significant increase in K562/R cells. No significant difference was observed in KCL22 and KU812*

cells. Bcr-Abl band intensities were normalized with their respective whole lane intensity. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (-p value <0.05, **-p value < 0.01). (D) Sequencing chromatograph of KCL22/R cells representing a heterozygous point mutation in the nucleotide at position 749 from G to A, resulting in amino acid glutamate at position 250 instead of glycine (G250E).*

5.3.3 ALTERED DRUG TRANSPORT

Altered IM transport was evaluated by (i) assessing the status of IM influx and efflux proteins OCT-1 and P-glycoprotein (P-gp) and (ii) determining the intracellular imatinib concentration. In K562/R, a significant downregulation of influx protein OCT-1 and about 60-fold upregulation of the efflux protein P-gp was observed. Reduction of intracellular imatinib level by about 50% was observed in K562/R compared to that of K562/S (Fig.5.13.A,D). Like in K562/R, KU812/R cells showed a significant upregulation of P-gp and about 50% reduction in intracellular imatinib levels (Fig.5.13.C, D). However, despite about 80% reduction in intracellular IM in KCL22/R compared to KCL22/S (Fig.5.13.D), no significant change in the levels of OCT-1 and P-gp was observed (Fig.5.13.B).

The status of known resistant mechanisms in all the three cell lines is summarized in Table 5.4. It is evident that in all resistant cell lines at least one of the established mechanisms of resistance exists, which in turn could result in a sub-optimal imatinib response i.e., reduced inhibition of tyrosine kinase activity of Bcr-Abl in resistant cells compared to that of their sensitive counterpart. To confirm this, the status of Bcr-Abl tyrosine kinase activity upon IM treatment in sensitive and resistant counterparts of all cell lines, was evaluated.

Table 5.4. Summary of status of known resistant mechanisms

Cell line	Bcr-Abl amplification	Bcr-Abl overexpression	Kinase domain mutation	Altered drug transport & reduced Intracellular imatinib
K562/R	X	✓	X	✓
KCL22/R	X	X	✓	✓
KU812/R	X	X	X	✓

5.3.4. ASSESSMENT OF STATUS OF BCR-ABL TYROSINE KINASE ACTIVITY

To study the effect of IM treatment on Bcr-Abl tyrosine kinase activity, the phosphorylation status of STAT5, the downstream substrate of Bcr-Abl, was assessed in untreated sensitive (S), IM-treated sensitive (S+IM) and IM-treated resistant cells (R). A partial inhibition of Bcr-Abl tyrosine kinase activity was observed in KCL22/R cells compared to its sensitive counterpart (Fig. 5.14.B), but the extent of Bcr-Abl activity inhibition in both K562/R and KU812/R cells was found to be same as that of their sensitive counterpart treated with IM (S+IM) (Fig. 5.14.A,C). This indicates that while resistance in KCL22/R cells is probably mediated by Bcr-Abl, resistance in K562/R and KU812/R cells is clearly independent of Bcr-Abl tyrosine kinase activity.

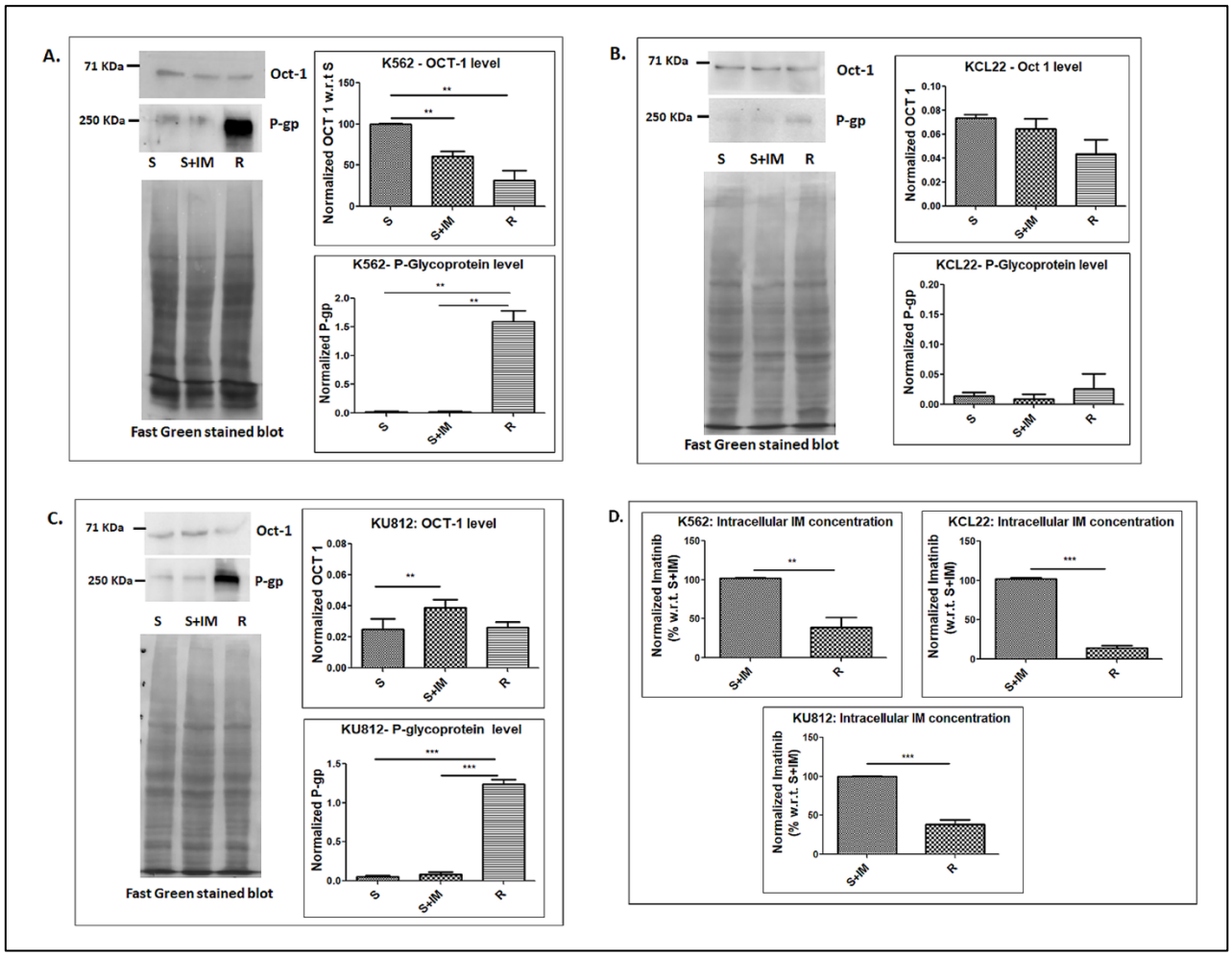


Fig. 5.13. Assessment of status of IM-drug transporters and intracellular IM concentration in IM-sensitive and resistant cells *(A-C) Western blotting for IM-importer OCT-1 and exporter P-gp in (A) K562 (B) KCL22 (C) KU812 – S, S+IM, R cells revealed a significant downregulation of OCT-1 in K562/R and upregulation of P-gp in K562/R as well as KU812/R. No change was observed in KCL22/R. Band intensities of OCT-1 and P-gp were normalized with their corresponding whole lane intensity. (D) Intracellular IM concentration determined by targeted MS approach was found to significantly reduced in all 3 resistant cells. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05; **-p value <0.01; ***-p value <0.001).*

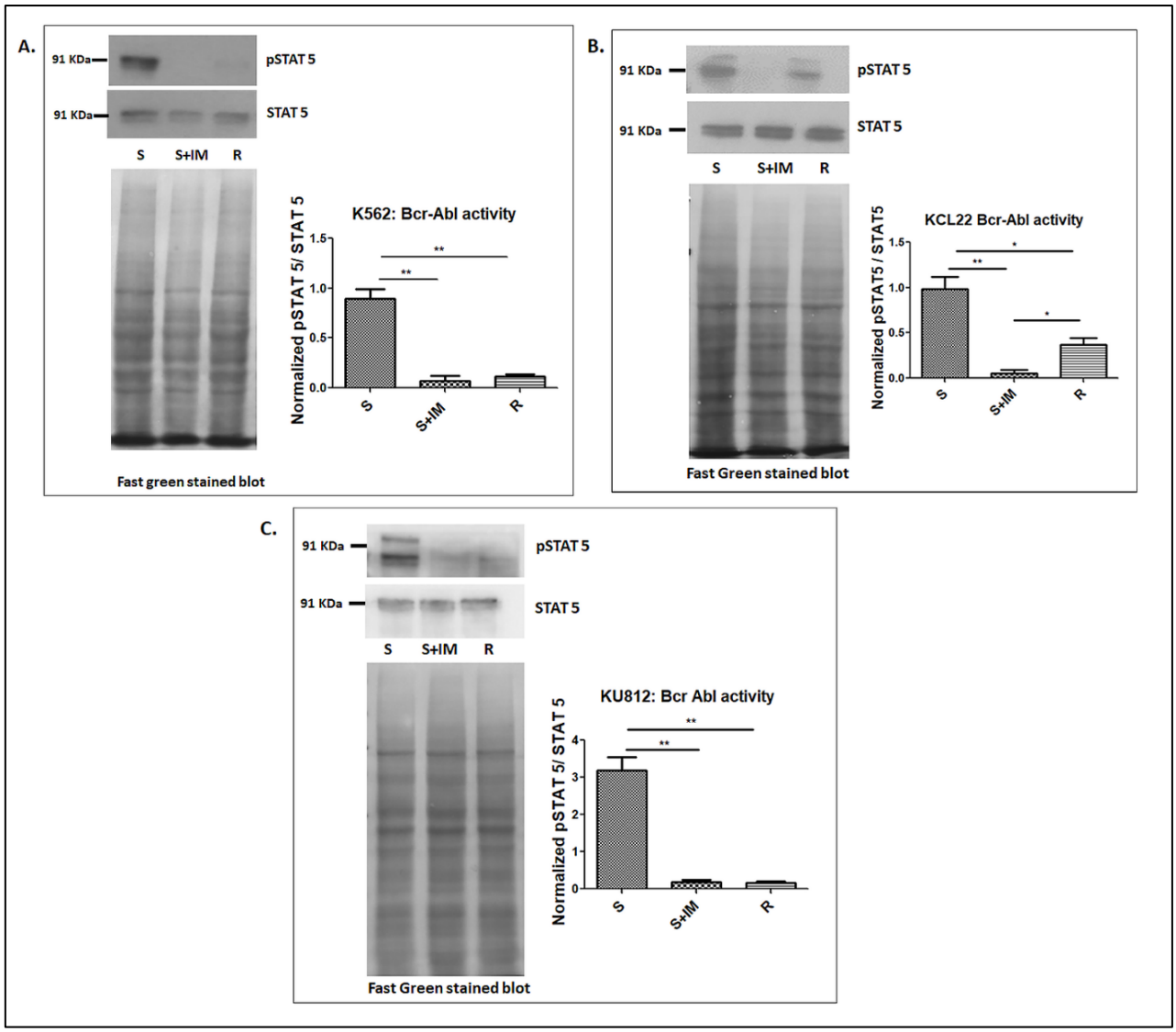


Fig. 5.14. Comparison of Bcr-Abl tyrosine kinase activity between IM-sensitive and resistant cells: *(A-C) Western blotting for STAT5 and phospho-STAT5 in (A) K562 (B) KCL22 (C) KU812 – S, S+IM, R cells. The ratio of p-STAT5/STAT5 indicated that Bcr-Abl activity in 'R' was inhibited to the same extent as 'S+IM' in K562 and KU812 cells. In KCL22/R only a partial inhibition of Bcr-Abl activity was observed. STAT5 and p-STAT5 band intensities were normalized with their corresponding whole lane intensity. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05; **-p value <0.01).*

In summary, despite reduced intracellular IM and overexpression of Bcr-Abl protein, IM could efficiently inhibit Bcr-Abl activity in K562/R and KU812/R cells indicating that IM resistance in both these cell lines is not mediated by tyrosine kinase activity of Bcr-Abl. However, in KCL22/R cells the sub-optimal inhibition of tyrosine kinase activity by IM could be attributed to the presence of G250E mutation in kinase domain of Bcr-Abl, implying the existence of Bcr-Abl dependent resistance.

5.4. PROTEOMIC ANALYSIS TO IDENTIFY ALTERATIONS IN BCR/ABL AND ALTERNATE PATHWAY IN IM-RESISTANT CELLS

Proteomic analysis was carried out only in K562 cell line and the alterations observed were validated in KU812 cell line. Based on Bcr-Abl activity and the extent of its inhibition by IM, the comparison groups for proteomic analysis were K562/S vs S+IM to identify components of Bcr-Abl downstream signalling pathway and K562/S+IM vs R to identify components that confer resistance. Quantitative proteomic analysis was carried out using label-free (SWATH) as well as labelled (iTRAQ) approach.

5.4.1. IDENTIFICATION OF OPTIMAL NORMALIZATION METHOD AND CRITERIA FOR DIFFERENTIATOR IDENTIFICATION IN SWATH-MS BASED LABEL-FREE QUANTITATIVE PROTEOMICS

SWATH-MS based label-free quantitative proteomic approach is being widely used for clinical biomarker discovery and therapeutic target identification, due to its ability to (a) provide in-depth profiling by employing DIA approach and (b) re-interrogate of data, when technical advancements enable generating more comprehensive spectral ion library, resulting in identification of more proteins [66]. A testimony to this is the wide use of SWATH-MS in clinical proteomics after its discovery in 2012. PubMed results show that 44% (20/45) of the SWATH-MS studies on clinical samples published till date are aimed at biomarker discovery or therapeutic target identification.

However, a feature in quantification by SWATH-MS, if overlooked, can hinder biomarker or target identification. Unlike labelled quantification by IDA wherein all samples for relative quantification are run together, in label-free quantification by SWATH, each sample from the comparison group is run individually in MS. This increases the probability of both systematic and random error. Intervention to reduce these variations by 'normalization' is thus a prerequisite to subsequent analysis of SWATH data for identification of differentiators. The data from reported SWATH-MS studies is normalized using either methods provided by the MS instrument-based software or those used to normalize microarray data [109-112]. As the source of systematic bias differs between MS and

microarray, it is essential to experimentally validate the appropriate normalization strategy for SWATH data. Hence in this study, before moving on to differentiator identification to understand the IM-resistant mechanism, we undertook a systematic approach using study sets involving comparing between K562 S vs S+IM and K562 S+IM vs R along with reference and validation sets obtained from Proteome Xchange consortium, to experimentally identify an appropriate normalization method for SWATH-MS data.

To achieve this, the statistical tool 'Normalyzer', which compares the efficiency of diverse methods to normalize 'omics' data based on statistical criteria [108], was used. Further, considering the wide application of SWATH-MS in biomarker identification, in this study we have supplemented the statistical evaluation with biologically relevant criteria of precise stratification of comparison groups by cluster analysis. Towards this (a) Normalization of data was assessed using 'Normalyzer' to identify the optimal method of normalization based on statistical criteria (b) from the data normalized by different methods in Normalyzer, differentiators between comparison groups were identified based on p-value, fold change and combination of both. The potential of these differentiators to segregate comparison groups distinctly, was assessed by cluster analysis. The details of samples used and experimental workflow are summarized in Fig. 5.15.

5.4.1.1. Identification and quantitation of proteins

In this study, each of the four biological replicates of K562 S, S+IM and R, underwent one IDA run for the generation of spectral ion library followed by three DIA runs for SWATH-MS analysis, thereby resulting in a total of 4 IDA and 12 DIA runs for K562 S, S+IM and R each. Samples with improper chromatogram were eliminated from analysis sets leaving 11 runs each in S and R in datasets C and D respectively (Fig 5.15.B). In dataset F, there were 2 technical replicates for each sample. Upon spectral alignment and filtering of ion library, 4404, 1450, 1757, 1808, 7057 and 5316 proteins that fulfilled the criteria (described in section 4.3.6.4) were further used for quantification of

datasets A, B, C, D, E and F respectively. Quantities of the identified proteins were further assessed for variation.

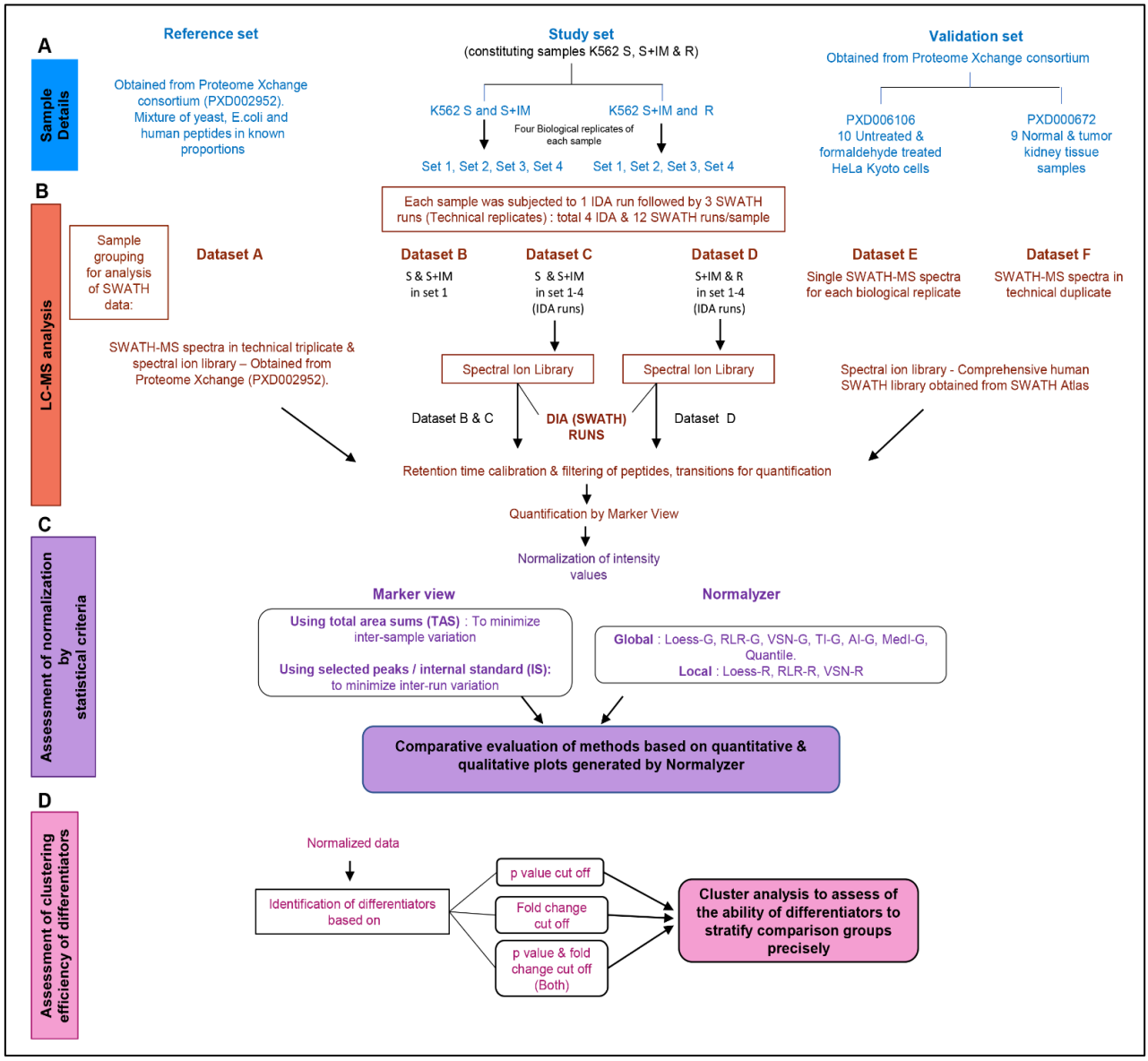


Fig.5.15. Scheme of experiments: *It describes (A) Samples used in this study which include IM-sensitive K562 cells (S) untreated or treated with imatinib (S+IM), IM-resistant K562 cells (R) and a3 datasets from public domain. (B) Generation of spectral ion library for all comparison groups in A from information dependent acquisition (IDA) data and generation of quantitative proteomic profile by data independent acquisition (DIA) using Sequential window acquisition of all theoretical fragment-ion spectra (SWATH). (C) Normalization of SWATH data using different methods. (D)*

Identification of differentiators based on p-value, fold change and combination of both followed by cluster analysis of the identified differentiators.

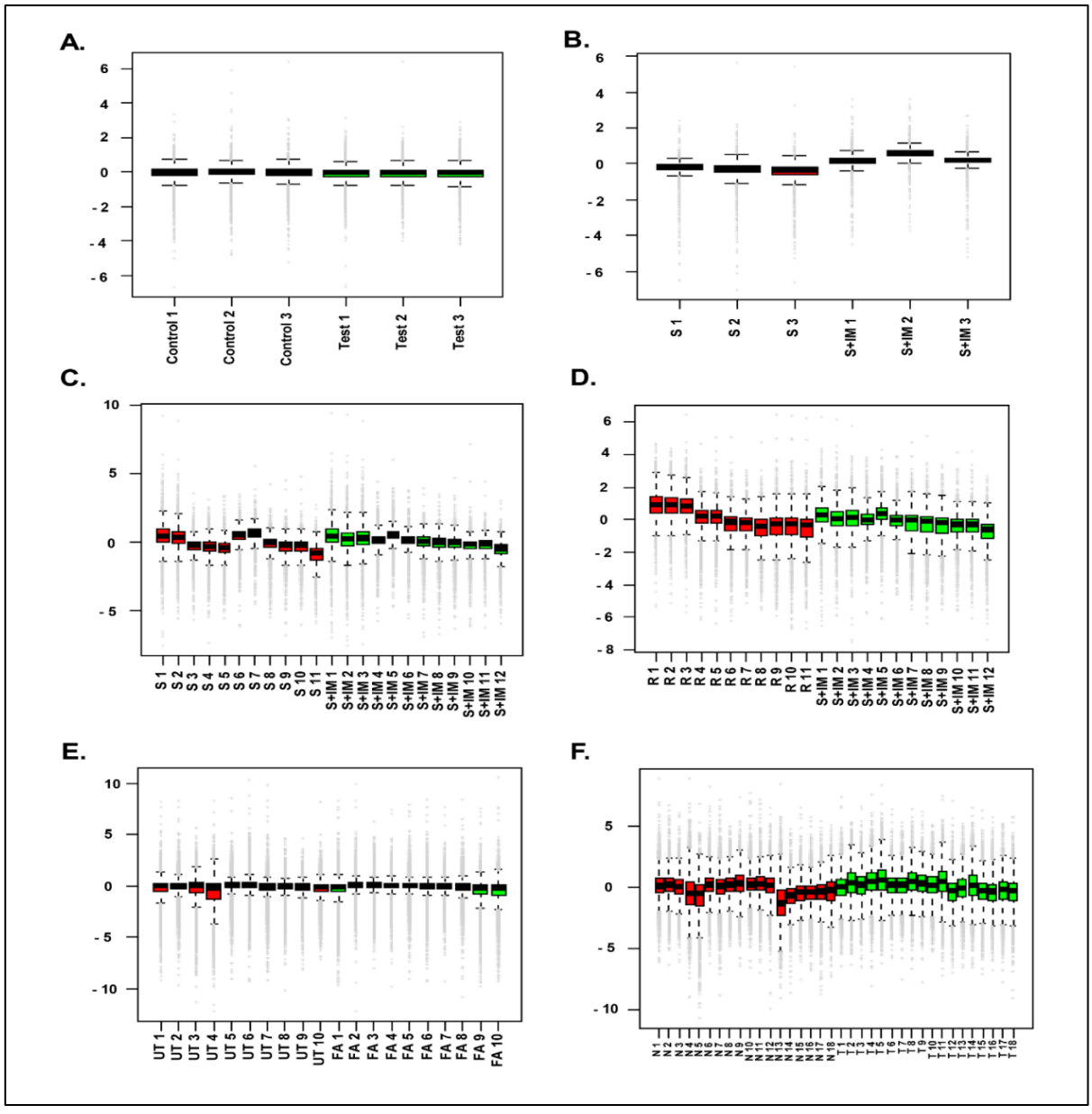


Fig.5.16. Analysis of unnormalized SWATH data for datasets A-F by RLE plot: *Qualitative assessment of the spread of data shows that the test and control groups vary in their spread of values in all datasets except A and E.*

5.4.1.2. Assessment of variation in un-normalized data

The quantified log2 transformed 'un-normalized' data of each dataset was evaluated based on RLE plot, which assesses the inter- and intra-group alignment of the replicates qualitatively. In RLE plot, samples should be aligned around zero. Any deviation would indicate discrepancies in the data[108]. Among the datasets constituted of single set of samples, alignment around zero was seen in all the representative samples of dataset A (Fig.5.16.A) and 50% of those in dataset B (Fig.5.16.B) . Datasets C (Fig. 5.16.C), D (Fig, 5.16.D) and F (Fig. 5.16.F) comprising of multiple sets, showed considerable deviation from zero in replicates as well as between groups in RLE plots, indicating the need for normalization of SWATH-MS data.

5.4.1.3. Identification of optimum method for normalization using 'Normalyzer'

The efficiency of 12 different normalization methods to normalize datasets A-F, was assessed quantitatively and qualitatively in 'Normalyzer' using PCV and RLE plots respectively. PCV reflects the ability of a normalization method to decrease intragroup variation between technical and/or biological replicates [109]. The results indicated that, VSN-G- normalized data consistently showed lesser intra-group variation in all datasets compared to data normalized by other methods (Fig. 5.17-I). Additionally, in datasets B-F VSN-R normalized data also reduced intra group variation. Further, qualitative assessment of the normalization methods with lowest PCV (VSN-G and VSN-R) by RLE plot indicated that only VSN-G showed good inter and intra group alignment among the replicates in all datasets (Fig. 5.17-II). Thus, VSN-G was identified as the optimal normalization method using 'Normalyzer' based evaluation.

5.4.1.4. Assessment of VSN-G normalized data by cluster analysis

Differentiators identified from data normalized by VSN-G method based on p-value, fold change and a combination of both were subjected to cluster analysis. Differentiators identified by all three criteria could segregate the comparison groups appropriately in datasets A, B and D but not in dataset C, E

and F (Fig.5.18). Though VSN-G was identified as optimal normalization method based on PCV and RLE plots, the differentiators identified did not show consistent efficiency in clustering. In order to understand the contribution of VSN-G normalization to improper clustering of datasets C, E and F, differentiators identified by all three criteria, from data normalized with the remaining eleven other methods were assessed for their clustering ability. The aim was to detect if any other normalization method could improve segregation of datasets C, E and F while retaining the efficient segregation of datasets A, B and D in VSN-G normalized data.

5.4.1.5. Assessment of data normalized by methods other than VSN-G by cluster analysis

As observed in VSN-G normalized data, clusters obtained from data normalized with the remaining eleven methods yielded improper clustering in datasets E and F. Thus the improper features of clusters i.e. formation of separate cluster by a few normal samples in datasets E and F; segregation of a pair of normal samples (N9 and N18) with tumor samples in dataset F was taken as a consistent feature across normalized data for these two datasets and was not applied to eliminate a cluster as imprecise. While retaining these features, clear segregation of the remaining control and test samples was considered as acceptable clustering efficiency of datasets E and F. Based on this relaxed criteria, it is seen in Fig.5.19 (Detailed dendrograms for cluster analysis is given in Appendix 2. Fig. A1 – A18) & Table 5.5 that differentiators identified based on p-value efficiently segregate the comparison groups for data normalized by majority of methods. On the other hand, differentiators identified based on fold change could not segregate the comparison groups in majority of the datasets. The ability of differentiators obtained from the combination of p-value and fold-change to segregate sets therefore could be attributed to the influence of p-value. Based on the above experimental evidence p-value is chosen as the criteria for differentiator identification in this study.

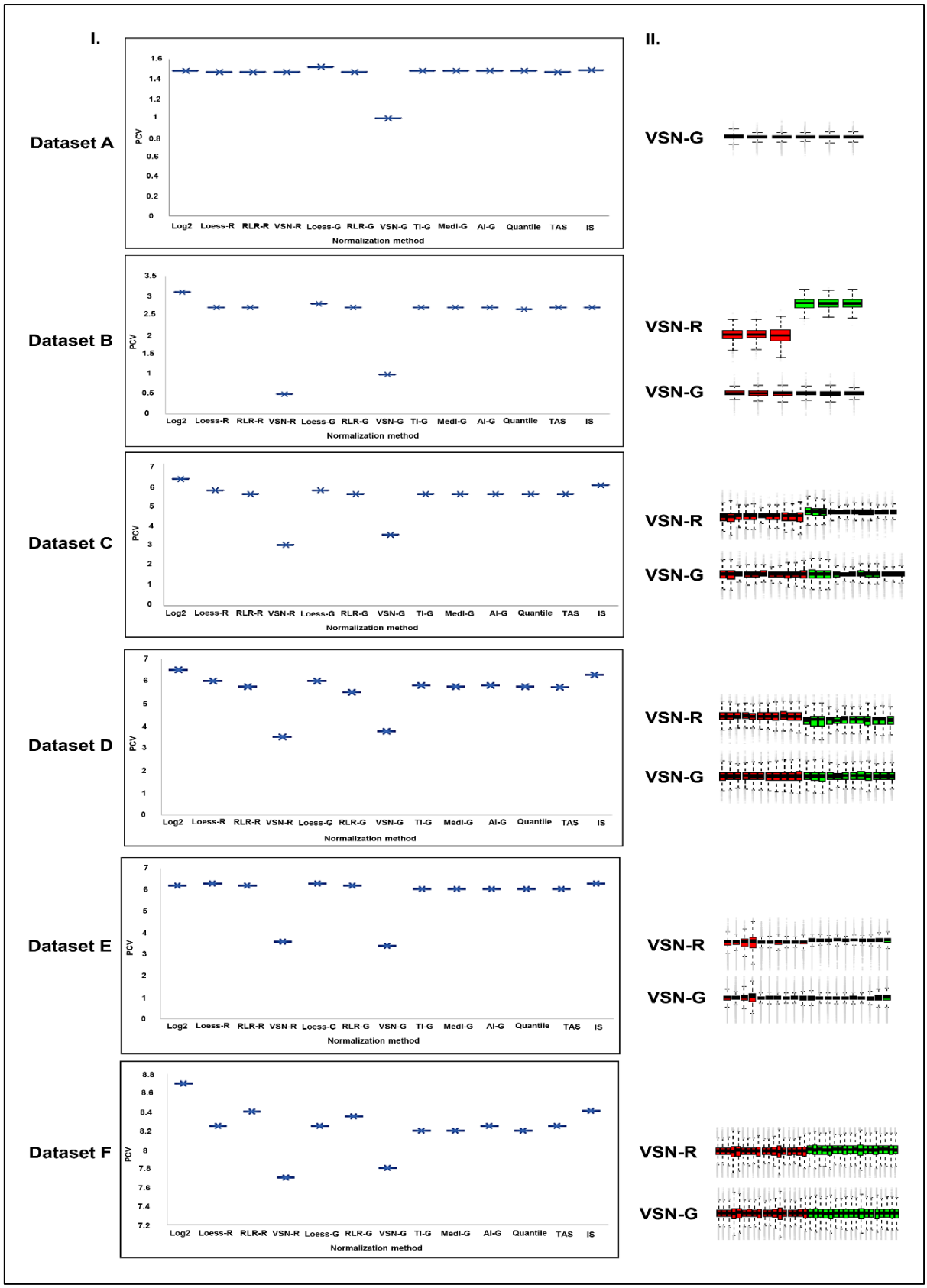


Fig. 5.17. (I) PCV plot: *Quantitative assessment of twelve normalization methods indicates that VSN-G has less PCV in all datasets along with VSN-R in datasets B-F, C and D.* ***(II) RLE plot:*** *Qualitative*

analysis of methods with less PCV, by RLE plot revealed good inter group alignment only in VSN-G in all datasets.

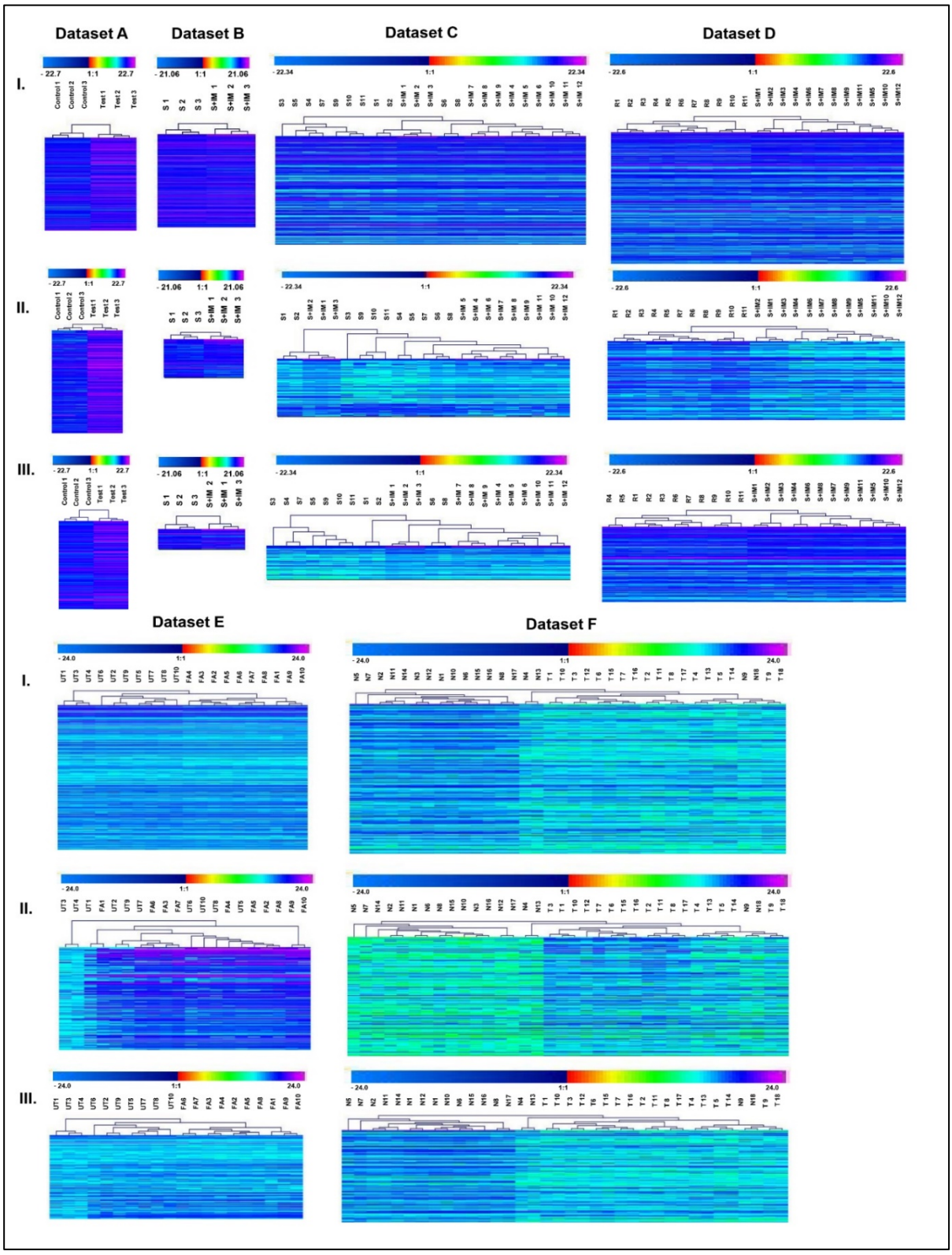


Fig. 5.18. Hierarchical clustering of differentiators obtained from VSN-G normalized data: *based on I- p-value, II- fold change, III- p- value together with fold change.*

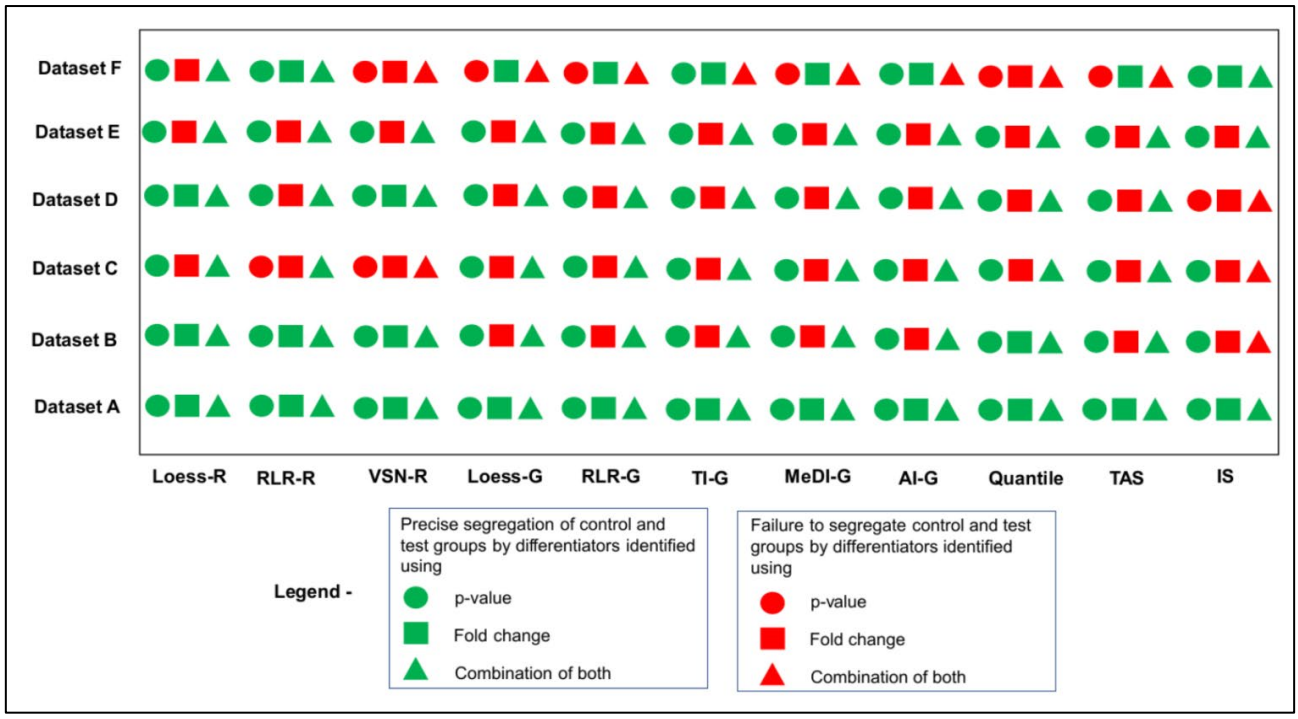


Fig. 5.19. Ability of differentiators to cluster the study groups distinctly.

Table 5.5. Clustering efficiency of differentiators identified based on p-value, fold change and combination of both, from data normalized by 12 methods

	Clustering efficiency		
Datasets	**p-value**	**Fold change**	**Both**
A	100% (12/12 methods)	100% (12/12 methods)	100% (12/12 methods)
B	100% (12/12 methods)	42% (5/12 methods)	92% (11/12 methods)
C	75% (9/12 methods)	0% (0/12 methods)	75% (9/12 methods)
D	92% (11/12 methods)	25% (3/12 methods)	92% (11/12 methods)
E	100% (12/12 methods)	0% (0/12 methods)	100% (12/12 methods)
F	42% (5/12 methods)	66.6% (8/12 methods)	25 (3/12 methods)

Of the 11 normalization methods assessed, differentiators identified based on p-value from data normalized by 3 methods (Loess-R, TI-G and AI-G) segregated the comparison groups precisely in all datasets (Fig 5.19). These were further evaluated using more stringent criteria to identify the most optimal method for biomarker discovery. The criteria was to sub-cluster the technical replicates, of control and test groups, belonging to each biological replicate precisely in datasets C, D and F. Dataset E was not subclustered as each sample was run only once [105]. A scoring system was used to achieve this, wherein the ability to segregate control and test groups was given a score of 2. In dataset F, for every control which segregated separately from the major control cluster, a negative score of 1 was given. Thereafter for every correct subgrouping of the technical replicates of control and test, a score of 1 was given. The total score was calculated as score for precise clustering (2) + score of -1 for each control which clustered separately from the major control cluster in dataset F (not applicable to other datasets) + score for co-segregation of technical replicates in test and control (1) (Fig.5.20).

As mentioned earlier, the efficiency of biomarkers lies in their ability to accurately stratify the heterogenous groups in a given population. It is evident from Fig. 5.20 that differentiators obtained from Loess-R normalized data could not only stratify the comparison groups precisely, but also had maximum sub-stratification score in the three large datasets assessed, thereby indicating its suitability for biomarker discovery by SWATH-MS.

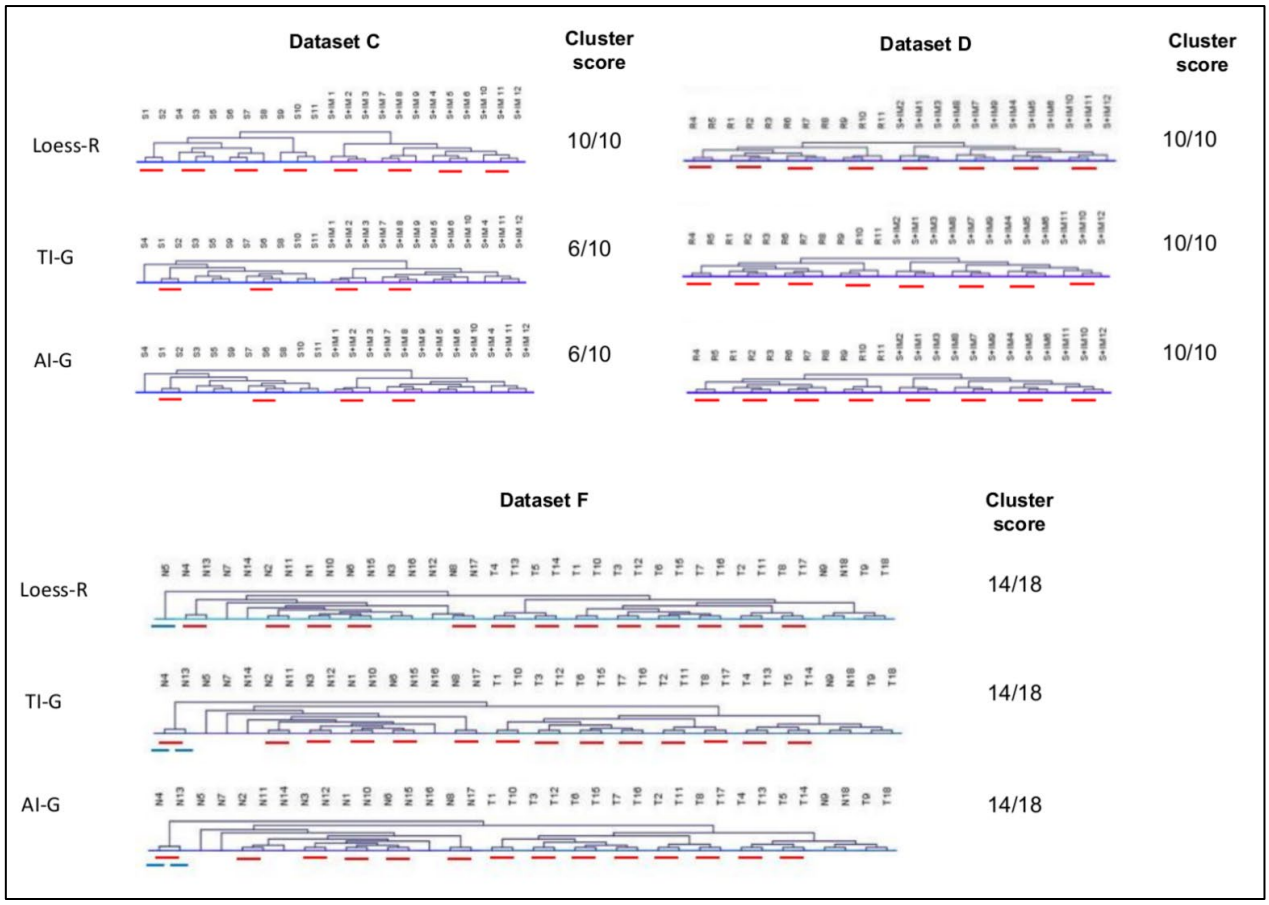


Fig.5.20. Evaluation of efficiency of p-value based differentiators to sub-stratify the technical replicates. *Proper clustering of test and control groups is given a score of 2 and proper sub-clustering of technical replicates of each set indicated by red line, is given a score of 1. In dataset F, a score of -1 is given to each control which formed a cluster outside the major control or test cluster, indicated by blue line.*

In summary by employing a systematic experimental approach, data normalized by LOESS-R method and differentiators identified on the basis of p-value were found to be optimal for SWATH-MS data, as this could yield differentiators that segregate the control and test group precisely. These parameters were employed to process SWATH data in this study.

5.4.2. IDENTIFICATION OF KEY PROTEINS INVOLVED IN IM-RESISTANCE

5.4.2.1. Label free quantification by SWATH-MS

Four biological replicates of K562-S vs S+IM and K562- S+IM vs R were subjected to quantitative proteomic profiling by SWATH-MS approach which resulted in quantification of a total of 1757 and 1808 proteins respectively. Upon normalizing the data using Loess-R method and applying a p-value cut-off of 0.05 for differentiator identification, 386 and 712 differentiators were identified from K562-S vs S+IM and K562- S+IM vs R comparison groups respectively. The list of differentiators is mentioned in Appendix 1 -Table A-3 and A-4.

5.4.2.2. Labelled quantification by iTRAQ

Quantitative proteomic profiling by iTRAQ for K562 S (control) vs S+IM (test) and K562 S+IM (control) vs R (test) was carried out using two biological replicates. The iTRAQ labels used for each sample is mentioned in section 4.3.6.5. Protein quantification results were obtained as ratio of control and test iTRAQ labels. In each comparison group ratio was generated between control and test samples of both the biological replicates as test 1/control 1, test 2/control 1, test 1/control 2 and test 2/control 2, resulting in 4 combination of iTRAQ ratios. Those proteins which show a significant change (p-value < 0.05) consistently in all 4 combinations were considered as differentiators. K562 S vs S+IM comparison resulted in identification and quantification of 2755 proteins of which 32 were found to be differentially expressed and 2462 proteins were quantified in K562 S+IM vs R comparison of which 41 were differentially expressed. The list of differentiators is mentioned in Appendix 1 - Table A-5 and A-6.

5.4.2.3. Identification of key differentiators

The differentiators identified from SWATH and iTRAQ were pooled and were categorized into various functional groups as shown in table 5.6. The signalling molecules identified were further categorized into those that are –

A) Unique to S vs S+IM comparison group – This includes proteins which are components of Bcr-Abl downstream signalling pathway (Fig. 5.21-A).

(B) Common to S vs S+IM as well as S+IM vs R comparison groups – This includes Bcr-Abl downstream components which are altered in resistance (Fig. 5.21-B).

(C) Unique to S+IM vs R comparison group - This includes proteins altered in resistance which are not part of Bcr-Abl pathway (Fig. 5.21-C).

Table 5.6. Categorization of differentiators into functional groups

S.NO.	FUNCTIONAL GROUP		NO. OF PROTEINS	
			S VS S+IM (Status upon IM treatment)	S+IM VS R (Status in resistance)
1	Nucleotide biosynthesis	up reg	5	13
		Down reg	4	3
2	DNA replication & repair	up reg	12	8
		Down reg	-	7
3	Chromatin Remodelling	up reg	12	8
		Down reg	1	4
4	mRNA synthesis	up reg	9	8
		Down reg	-	4
5	mRNA splicing & stability	up reg	46	53
		Down reg	2	10
6	Ribosome biogenesis	up reg	2	42
		Down reg	17	2
7	Amino acid biosynthesis/degradation	up reg	6	6
		Down reg	-	7
8	Protein synthesis	up reg	10	18
		Down reg	10	9
9	Protein folding	up reg	4	18
		Down reg	2	3

10	mRNA/protein nuclear import/export	up reg	7	16
		Down reg	2	4
11	Protein degradation	up reg	16	23
		Down reg	1	9
12	Vesicle transport & microtubule assembly	up reg	27	31
		Down reg	3	9
13	Cell cycle, mitosis	up reg	-	4
		Down reg	-	1
14	Cytoskeleton / actin related proteins	up reg	16	23
		Down reg	1	9
15	Signalling	up reg	14	26
		Down reg	2	4
16	Mitochondria structure & function	up reg	9	15
		Down reg	3	15
17	Energy metabolism	up reg	37	29
		Down reg	1	30
18	Erythroid differentiation	up reg	7	3
		Down reg	-	3
19	Oxidative stress & ROS detoxification	up reg	10	10
		Down reg	-	4
20	Apoptosis	up reg	-	3
		Down reg	1	1

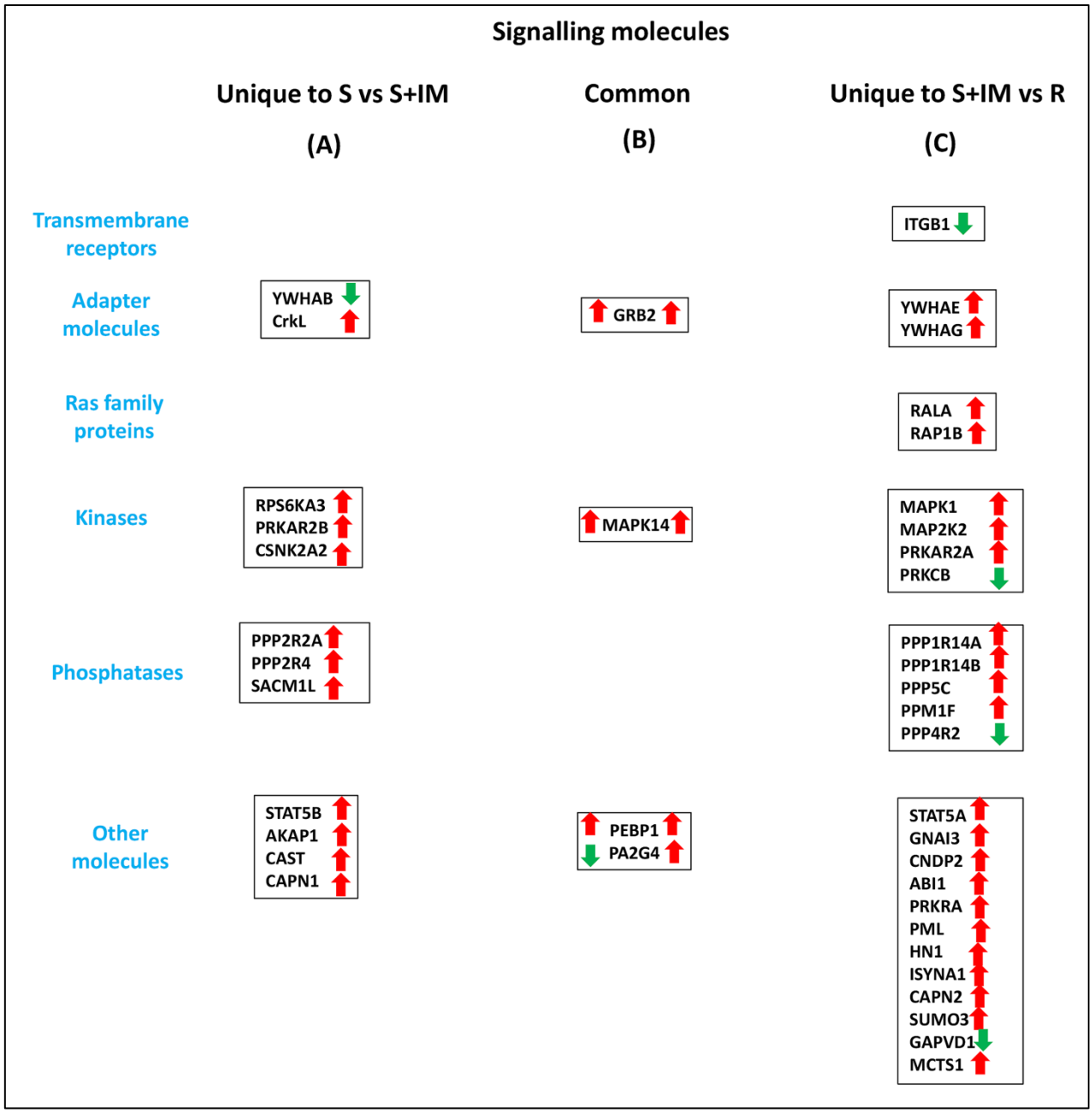


Fig.5.21. Sub-categorization of signalling molecules: *(A) Proteins unique to S vs S+IM comparison group represent components of Bcr-Abl signalling pathway. (B) Proteins common to both groups represent components of Bcr-Abl altered in resistance. (C) Proteins unique to S+IM vs R are altered in resistance but are not components of Bcr-Abl pathway.*

This categorization revealed that some signalling molecules altered upon IM treatment in sensitive cells (S+IM), signifying their involvement in Bcr-Abl pathway, were found to be further modulated in resistant cells. Considering that the extent of Bcr-Abl activity inhibition is comparable in

K562/S+IM and R cells, further alteration in Bcr-Abl pathway components indicate the probable modulation of their levels in a Bcr-Abl independent manner in resistant cells.

To further understand which components of Bcr-Abl pathway are altered in resistance and how are they functionally associated with Bcr-Abl, a string analysis was carried out for both the comparison groups by including Abl as one of the inputs [as Bcr/Abl is not an annotated SWISS-PROT entry] (Fig 5.22-A,B). With Abl as the starting point in string analysis, a downstream functional association network was charted, wherein those with direct functional association with Abl were termed first level interactors, the proteins with which they in-turn associate were termed second level interactors and so on. The association was charted up to third level interactors (Fig.5.23-A,B). A simplified version of this interaction is shown in Fig. 5.23-C,D which indicates that (i) in both the comparison groups 14-3-3 family proteins (YWHAB, YWHAG, YWHAE) were functionally associated with multiple proteins, thereby becoming a hub molecule and (ii) p38-α MAPK (MAPK-14) which was found to be associated with Abl was further altered in resistance.

Thus, from string analysis 14-3-3 family proteins and p38 MAPK were identified as key proteins of Bcr-Abl pathway modulated in resistance.

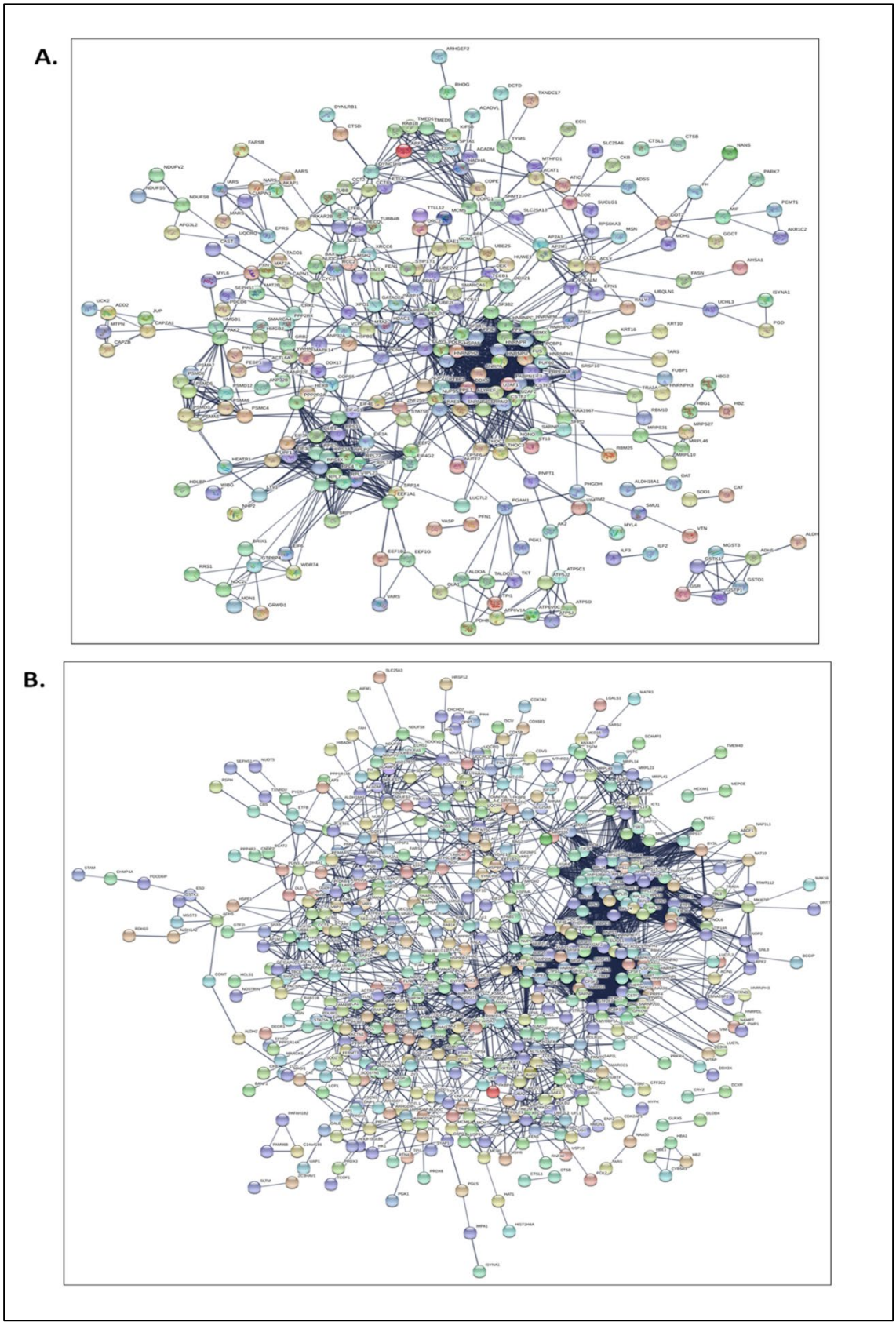


Fig.5.22. STRING analysis of differentiators: *(A) Differentiators identified from K562/S vs S+IM comparison (B) Differentiators identified from K562/S+IM vs R comparison. To understand the functional association of these differentiators with Bcr-Abl, Abl was included in the STRING.*

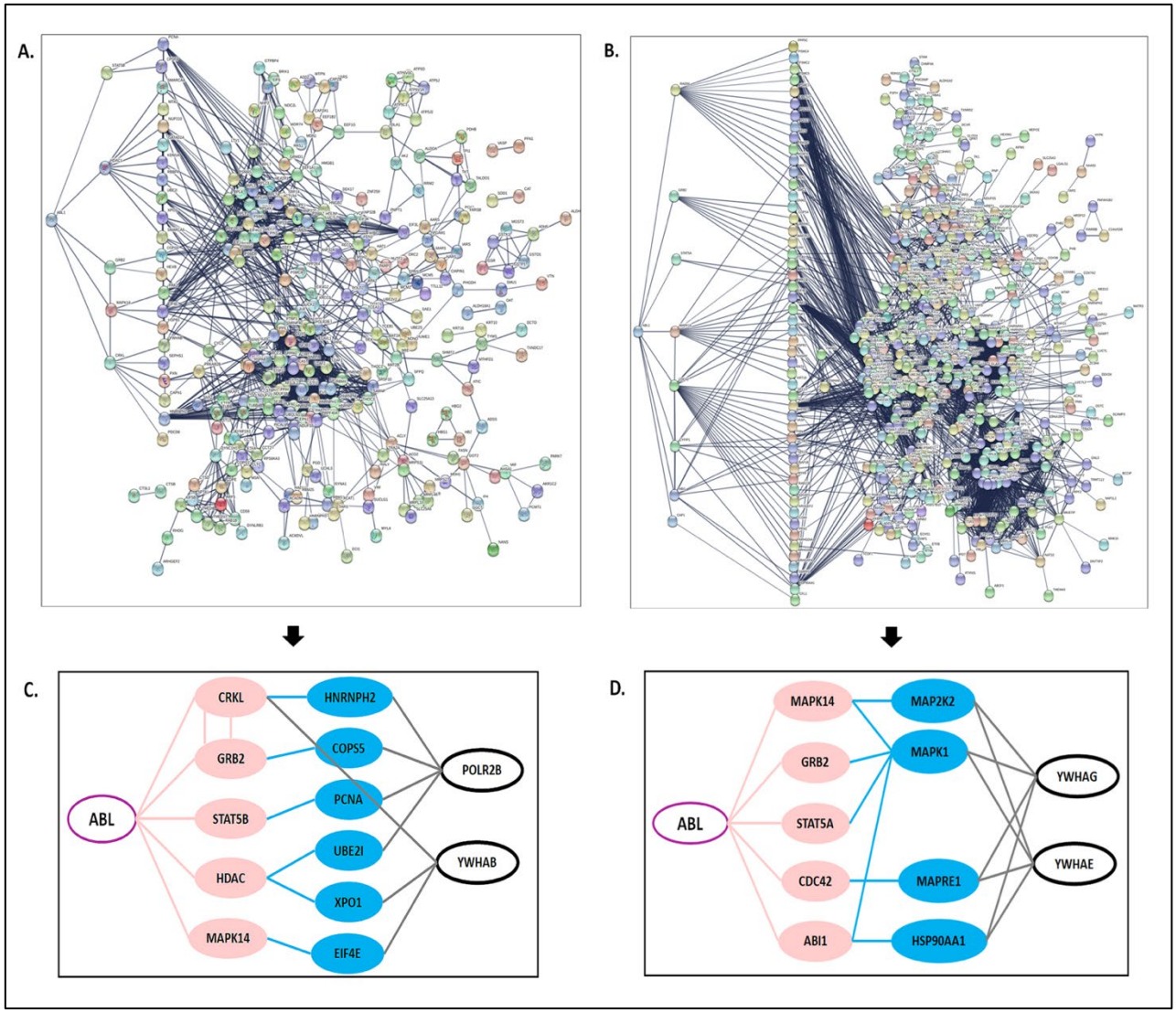


Fig. 5.23. Functional association hierarchy of Abl: *The association was charted up to third level interactors with Abl as the starting point for (A) Differentiators identified from K562/S vs S+IM comparison and (B) Differentiators identified from K562/S+IM vs R comparison, a simplified version of which is shown in (C) and (D) respectively. (C-D) Proteins that directly functionally associate with Abl were termed first level interactors (colour coded in pink), proteins with which they in turn associate were termed second level interactors (colour coded in blue) and the last being third level interactors (in black and white).*

5.4.3. VALIDATION OF KEY DIFFERENTIATORS BY WESTERN BLOTTING

Validation of key differentiators by western blotting revealed that among the 14-3-3 family members (Fig.5.24-A & B), both 14-3-3 γ (YWHAG) and 14-3-3 ε (YWHAE) were found to be significantly altered upon treating K562/S with IM (S+IM). However, only 14-3-3 ε was found to be altered in resistant cells with a significant downregulation, in comparison to untreated as well as IM treated sensitive cells. In case of p38-MAPK protein, western blotting was carried out using antibody specific to p38-α/β/γ MAPK. No change in p38 level was observed in K562/S, S+IM or R cells (Fig5.25-A). However, upon inhibition of Bcr-Abl activity in sensitive cells by IM treatment, phosphorylation of p38-MAPK was significantly reduced. In K562/R, on the other hand where Bcr-Abl activity is inhibited to the same extent as S+IM, a 2-fold increase in phosphorylation of p38-MAPK was observed (Fig 5.25-B).

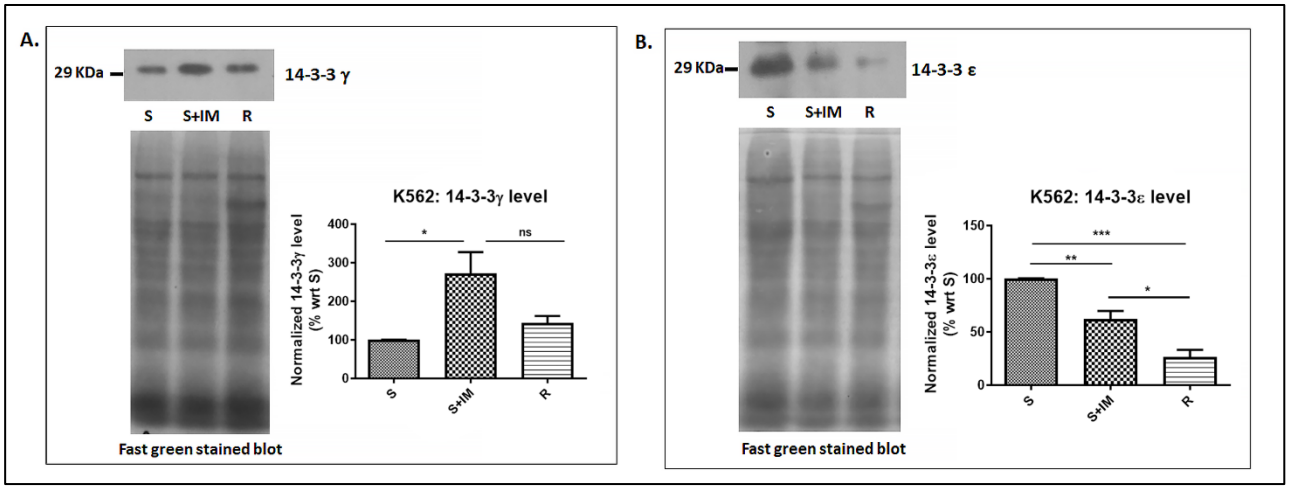


Fig. 5.24. Validation of 14-3-3 family proteins by western blotting: *(A) 14-3-3 γ was found to be significantly upregulated in sensitive cells upon IM treatment but no change was observed in resistant cells (B) 14-3-3 ε was significantly downregulated in sensitive cells upon IM treatment which was further downregulated in resistant cells. Band intensities of 14-3-3 γ and ε were normalized with their corresponding whole lane intensity. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05; **-p value <0.01; ***-p value <0.001).*

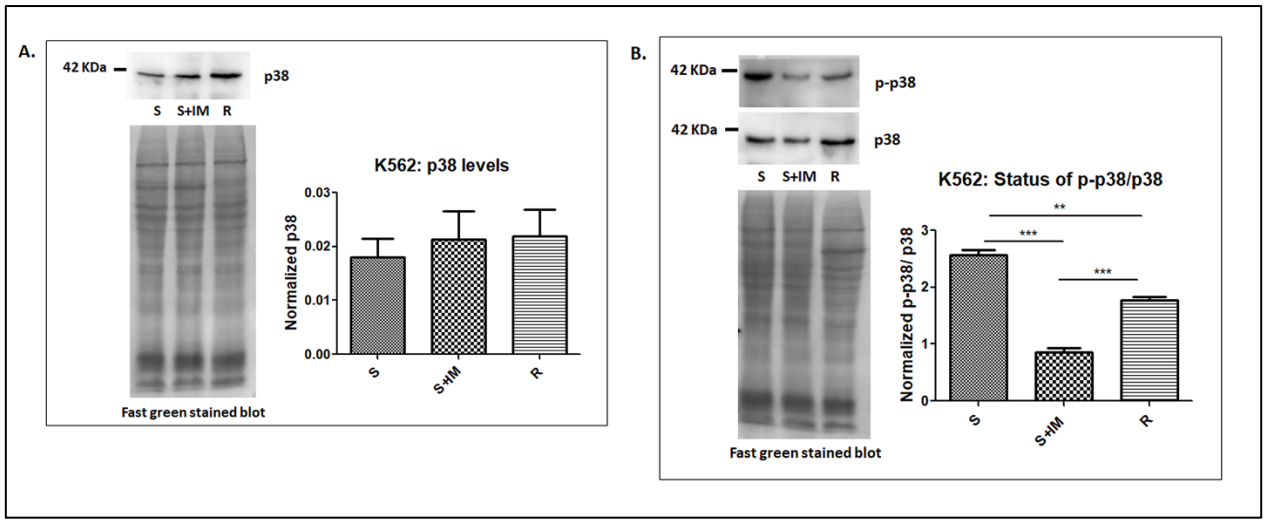


Fig.5.25. Validation of p38-MAPK by western blotting: *(A) No change in levels of p38-MAPK was observed. (B) A significant reduction in phosphorylation of p38-MAPK was observed in IM treated sensitive cells while a significant increase was observed in resistant cells with respect to sensitive cells treated with IM.* Band intensities of p38 and phospho p38 were normalized *with their corresponding whole lane intensity. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05; **-p value <0.01; ***-p value <0.001).*

5.4.3.1. Validation of observations from K562/R cells in KU812/R cells

The status of key differentiators 14-3-3 ε and p-38 MAPK were assessed in KU812/R cells, in which like K562/R resistance is not totally mediated by Bcr-Abl. Western blotting of KU812/S and R cells demonstrated that like in K562/R, 14-3-3 ε was found to be significantly downregulated (Fig.5.26-A) while phosphorylation of p38-MAPK was significantly upregulated in KU812/R cells (Fig.5.26-B), corroborating the importance of these molecules in IM-resistance.

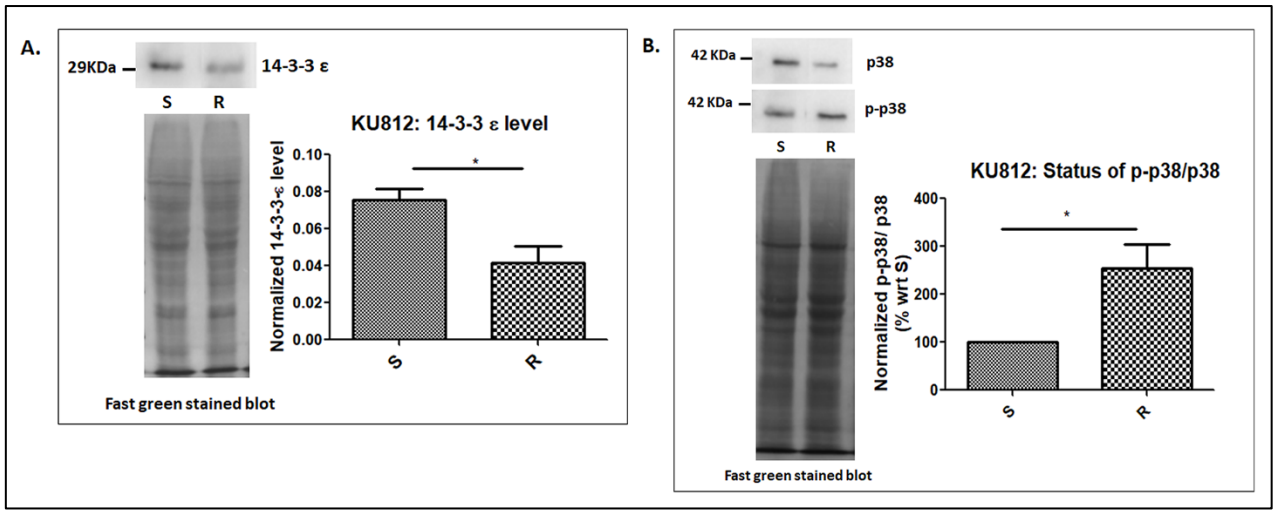


Fig.5.26. Assessment of status of key differentiators in KU812/R cells: *Western blotting of KU812/S vs R for (A) 14-3-3ε revealed a significant down-regulation in KU812/R (B) p38 and phospho-p38 MAPK revealed a significant increase in phosphorylation of p38 MAPK in KU812/R cells, both being consistent with that observed in K562/R cells. Band intensities of 14-3-3ε, p38 and phospho p38 were normalized with their corresponding whole lane intensity. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05).*

In summary, proteomic analysis by SWATH and iTRAQ resulted in identification of 14-3-3 family proteins and p38-MAPK as key differentiators. Validation of these proteins by western blotting revealed a significant downregulation of 14-3-3 ε and increase in phosphorylation of p38-MAPK in resistant cells, which has been further assessed to understand their contribution towards IM-resistance.

5.5. FUNCTIONAL VALIDATION OF KEY DIFFERENTIATORS

5.5.1. EFFECT OF MODULATION OF 14-3-3 ε LEVEL ON RESPONSE OF CELLS TO IM TREATMENT

To study the effect of reduction of 14-3-3 ε level in CML-BC cells, 14-3-3 ε was knocked out in K562/S cells using crispr-cas9 system.

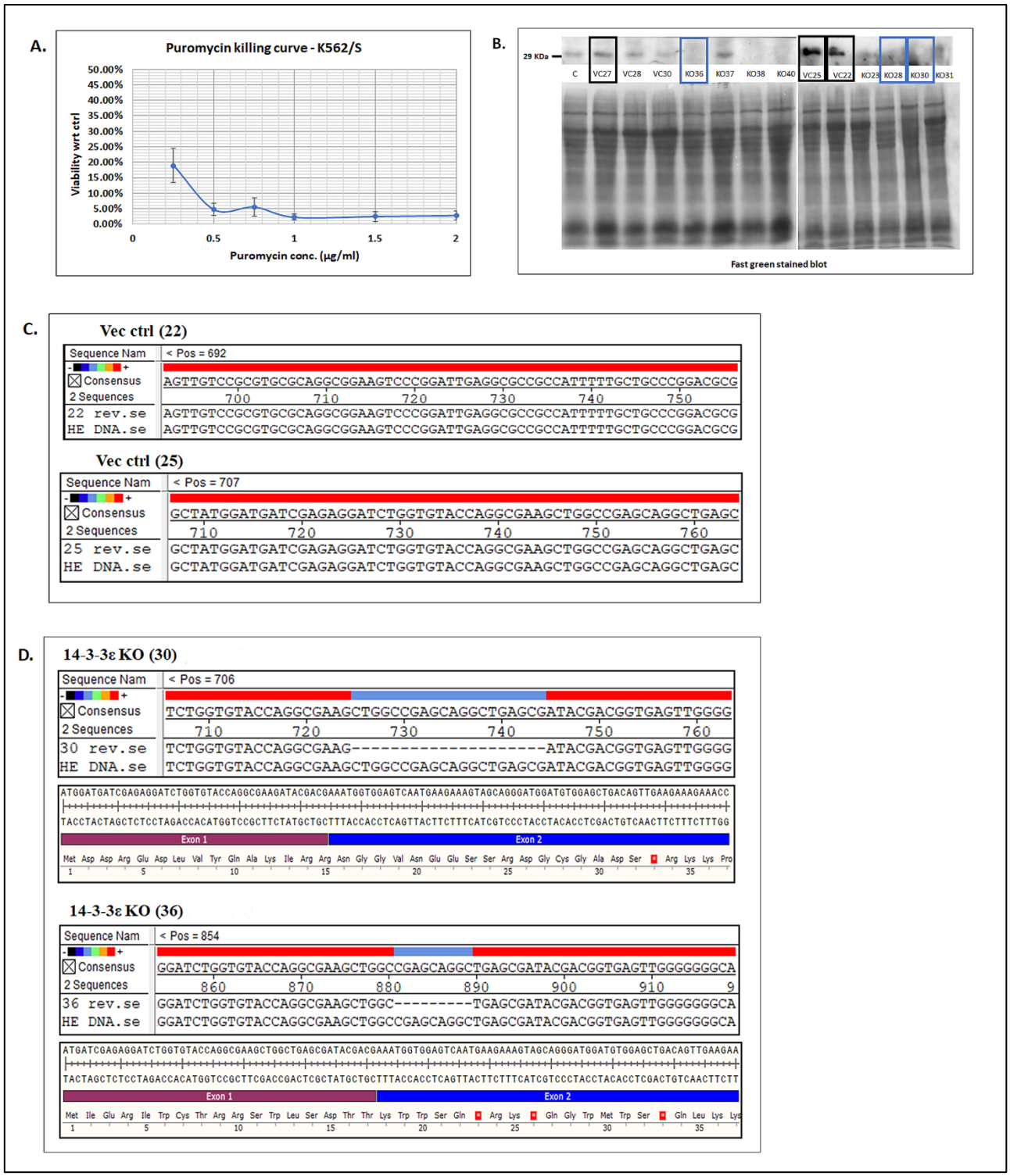


Fig.5.27. Knockout of 14-3-3 ε in K562/S cells: *(A) Puromycin killing curve of K562/S cells. Since 1μg/ml puromycin results in death of >90% untransfected cells, the concentration was chosen for selection of stably transfected cells. (B) Western blotting of 14-3-3 ε knockout and vector control clones. Based on the expression level, 3 vector control (boxed in black) and 3 knockout clones (boxed in blue) were chosen for further analysis. (C) Two vector control clones (22 and 25) were sequenced and aligned with WT 14-3-3ε sequence using DNASTAR Lasergene software. Red indicates a complete match with the WT 14-3-3ε sequence. (D) Two 14-3-3ε KO clones (30 and 36) were sequenced and aligned with WT 14-3-3ε sequence using DNASTAR Lasergene software. Red indicates a complete match with the WT 14-3-3ε sequence whereas blue indicates deletion. The amino acid sequence for the respective DNA sequence is shown where the asterisk indicates a stop codon.*

Stable clones were generated by maintaining transfected cells under 1μg/ml puromycin selection pressure, determined based on puromycin killing curve (Fig.5.27-A). 14-3-3 ε knockout and vector control clones obtained from single cell sorting of stably transfected cells, were screened for 14-3-3 ε protein level by western blotting (Fig.5.27-B), from which three vector control and knockout clones each with or without expression of 14-3-3 ε respectively were selected for further study. Two representative knockout and vector control clones were sequenced to confirm 14-3-3 ε knockout. The results revealed deletion of sequences in both the knockout clones leading to frame shift mutation resulting in premature stop codon (Fig.5.27-C,D), thereby confirming knockout. The cells were then assessed for the effect of 14-3-3 ε knockout on IC_{50} of IM using MTT assay. IC_{50} of IM was found to be significantly increased in 14-3-3 ε knockout clones compared to that of vector control (Fig.5.28).

Thus, downregulation of 14-3-3 ε, a key component of Bcr-Abl pathway was found to mediate development of IM-resistant phenotype.

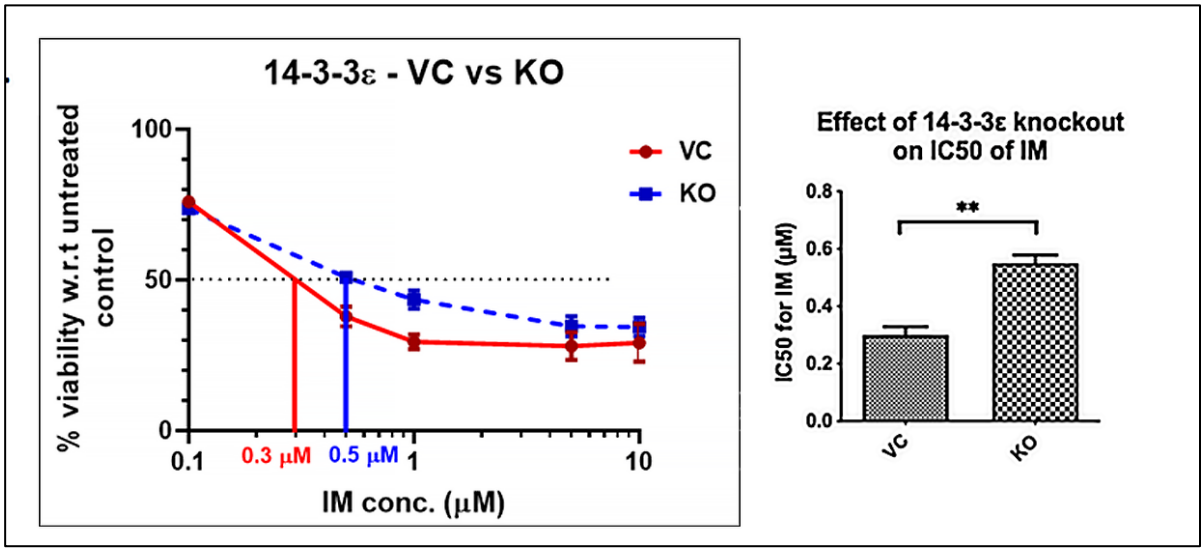


Fig.5.28. MTT assay to determine the effect of 14-3-3 ε knockout on IM treatment: *14-3-3 ε knockout lead to increase in IC_{50} of IM compared to that of vector control, implying its contribution towards development of IM- resistance. The experiment was carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (**-p value <0.01).*

5.5.2. EFFECT OF INHIBITION OF p38-MAPK ACTIVITY ON IM-RESISTANT CELLS

To check if increased p38 phosphorylation (active form) has a role to play in IM-resistance, the Ser/Thr kinase activity of p38 was inhibited in K562/R cells using the inhibitor SB203580 and its effect on viability of K562/R cells as well as their response to IM was assessed. SB203580 is known to inhibit p38 activity by binding to ATP binding pocket thus preventing phosphorylation of its downstream substrates [113]. Hence inhibition of p38 activity was assessed based on the phosphorylation status of its downstream target Msk1 [114, 115]. Western blotting of K562/R cells untreated and treated with 10µM p38i for 1hr (Fig.5.29-A), using msk and p-msk antibodies revealed a significant reduction in phosphorylation of msk confirming inhibition of p38-MAPK.

5.5.2.1. Inhibition of p38-MAPK and its effect on cell viability

To assess the effect of inhibition of p38 activity on K562/R cells, an MTT assay was carried out by treating K562/R cells with 10µM p38 inhibitor (p38i) for 1hr. This resulted in about 40% reduction in viability of K562/R cells compared to that of untreated control (Fig.5.30-B).

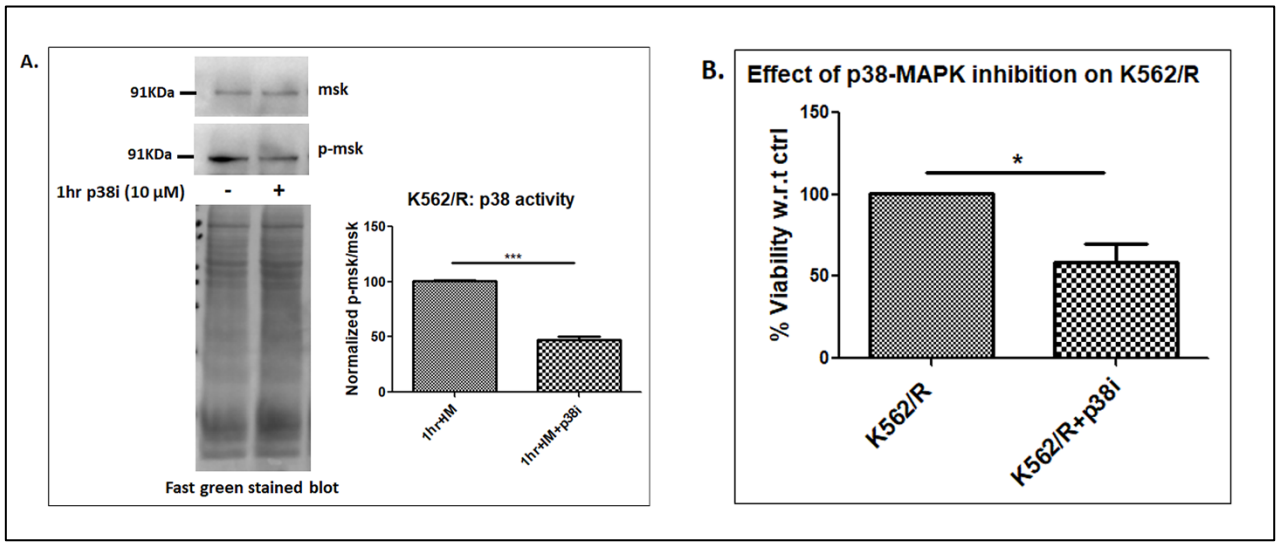


Fig. 5.29. Effect of inhibition of p38-MAPK activity in K562/R cells: *(A) Western blotting to confirm inhibition of p38 activity. The ratio of p-Msk1/Msk1 revealed a significant reduction in its phosphorylation reflecting inhibition of p38-MAPK activity. Band intensities of msk and phospho msk were normalized with their corresponding whole lane intensity. (B) MTT assay to assess the effect of p38-MAPK inhibition, demonstrated a significant reduction in viability of K562/R cells. All experiments were carried out in biological triplicates. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05, ***-p value <0.001).*

5.5.2.2. Inhibition of p38-MAPK and response to IM treatment

Further, to understand whether inhibition of p38-MAPK activity modulates the response of cells to IM, an MTT assay carried out in K562/R cells treated with range of IM concentrations with or without 10µM p38i for 1hr. While IC_{50} of IM with 1 hr. treatment was found to be 50 µM, additional treatment with p38i drastically reduced the IC_{50} of IM to 7 µM (Fig.5.30) thereby sensitizing the cells to IM.

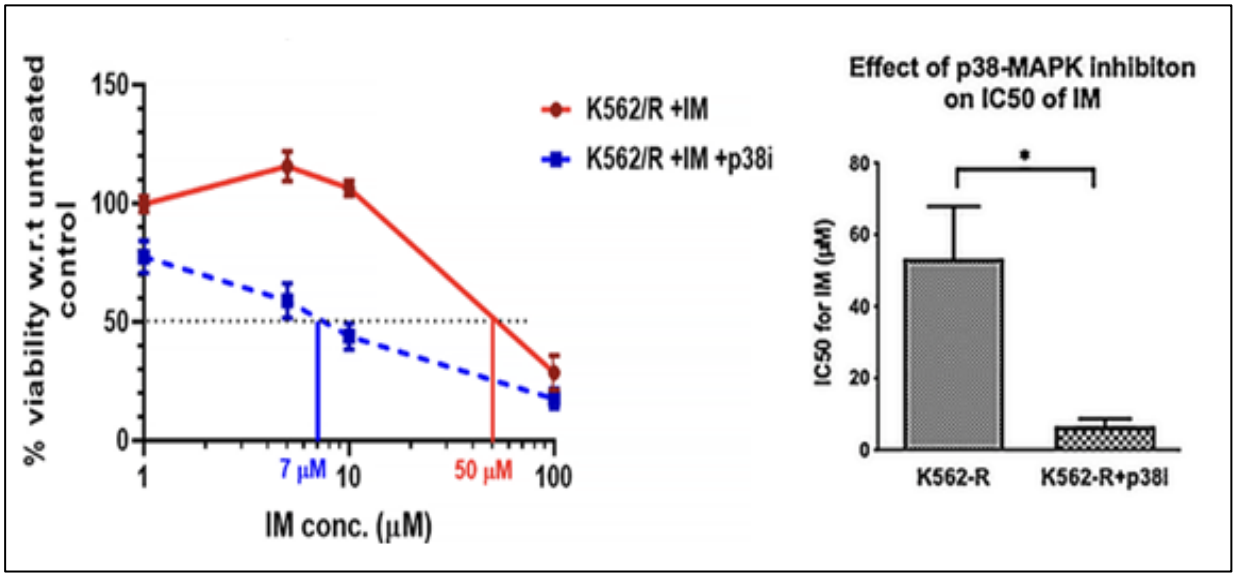


Fig. 5.30. Effect of inhibition of p38-MAPK activity on IM treatment. *MTT assay demonstrated a significant 7-fold reduction in IC_{50} of IM upon inhibition of Ser/Thr kinase activity of p38-MAPK. The experiment was carried out in biological triplicate. The graph represents mean ± SEM. Statistical significance was assessed by student's t-test (*-p value <0.05).*

Thus, p38 inhibitor either alone or in combination with IM could inhibit the growth of K562/R cells, thereby endorsing its potential as therapeutic target for IM resistant cells.

6.DISCUSSION

Introduction of Imatinib in 2001, brought about a paradigm shift in CML therapy by drastically reducing the annual mortality rate [13, 116, 117]. However, this also increased the prevalence of CML by about 4-fold after 2001 and has been expected to reach a plateau in 2030 with about 3 million people affected globally. Imatinib being a maintenance therapy, needs to be administered continuously. As a result, this surge in prevalence apart from increasing the burden of drug availability would also have an alarming consequence of increase in the rate of acquired resistance to TKIs. Patients unresponsive to TKIs would inevitably progress to BC wherein the survival is only up to six – eleven months [13, 15, 28]. This is of even more serious concern in India, where the age of incidence is a about a decade earlier than west [58, 59]. It is therefore essential to discern the pathophysiological events downstream of Bcr-Abl and molecular alterations that provide survival advantage to the cells in blast crisis in a Bcr-Abl independent manner, to identify targetable key mediators, which would help in overcoming IM resistance.

Proteomic approach has been extensively used to identify therapeutic targets in different cancers [74, 77]. In CML too, proteomic studies have been carried out wherein profiles of CML-BC cell line with and without IM treatment (+/- IM) or sensitive and resistant to IM (S vs R) have been generated. However, majority of these studies were limited to identification of differentiators and their functional categorization, with no functional validation to further assess the involvement of differentiators in Bcr-Abl signalling pathway, their role in resistance or as therapeutic target. While hypusination of eIF5A [90] and calcium homeostasis [97] have been identified as the key processes and HSP70 [95] as the key molecule that could be targeted to overcome IM-resistance in CML-BC, their consistency among heterogenous CML-BC patients and translational potential is yet unknown. Thus, independently these studies provide a limited insight to Bcr-Abl downstream or Bcr-Abl independent signalling pathways leading to resistance. A meta-analysis carried out to collectively investigate the outcome of these studies revealed inconsistency among the differentiators identified by various studies within +/- IM and S vs R comparison groups (Fig.5.2).Due to lack of overlap, differentiators belonging to K562 +/- IM comparison group and K562 S vs R group were pooled and subjected to

STRING analysis (Fig.5.3 and Fig.5.4), from which it was evident that differentiators in these studies majorly constitute high abundant proteins involved in cellular processes like splicing, translation, energy metabolism, cell cycle, protein degradation etc. while low abundant signalling molecules were barely detected. As a result, the differentiators from studies in K562 +/- IM group and K562 S vs R group were inadequate to establish any functional association with Abl to delineate Bcr-Abl downstream pathway or identify molecules that contribute to resistance respectively. This thus clearly emphasized the need for more comprehensive proteomic studies to gain deeper insights on Bcr-Abl downstream signalling pathway as well as pathways that confer resistance, which in turn would help identifying alternate therapeutic targets for IM-resistant CML-BC.

WHAT DID WE LEARN FROM THE PREVIOUS STUDIES?

We assessed the reasons for failure of the previous analyses to identify robust experimental system for identification of therapeutic targets. We concluded that:

1. Selection of appropriate biological system to carry out analysis is crucial. In CML-BC, about 65% patients have myeloid blast crisis, 30% have lymphoid blast crisis and 5% have blasts of mixed lineage [6, 10]. Further, among myeloid BC patients, a predominance of blasts from one of the following lineages namely erythroid, neutrophilic, megakaryocytic, basophilic, eosinophilic and monocytic or a combination of these was observed [118]. Hence, the cell lines in this study (Table 4.4) were chosen so as to represent this heterogeneity observed in patients. The cell lines were characterized by their morphology and doubling time, which were consistent with the reported literature [119-121]. Further, resistant cell lines were developed for all three cell lines, which provided a system that mimicked the occurrence of resistance in blasts of different lineages, as detected in CML-BC patients. The wild type cells were termed sensitive and designated as K562/S, KCL22/S, KU812/S cells while their resistant counterpart developed by IM dose escalation were termed K562/R, KCL22/R and KU812/R cells. The resistance of these cells to

IM was confirmed by detection of increase in IC_{50} compared to the sensitive cells in MTT assay (Fig.5.8).

Since several mechanisms were known to mediate IM resistance, it was essential to examine which of the those contributed to resistance in our system. K562/S, KCL22/S and KU812/S being CML-BC cell lines, multiple copies of *BCR-ABL* fusion gene were detected by interphase and metaphase FISH (Fig. 5.11). Consistent with that reported by Virgili *et.al.*[122], intrachromosomal amplification of *BCR-ABL* was observed in K562/S and KU812/S while duplication of Ph-chromosome was observed in KCL22/S. The resistant cells just retained the abnormalities observed in their sensitive counterpart without bearing any further *BCR-ABL* amplification or duplication, thereby ruling out *BCR-ABL* gene amplification as the cause for resistance in these cell lines. However, Bcr-Abl protein was significantly upregulated in K562/R cells (Fig.5.12.A). In KCL22/R cells G250E kinase domain mutation, which is common in IM-resistant patients, was detected (Fig.5.12.D). This mutation occurs in the ATP binding site (P-loop) which prevents imatinib binding, thereby conferring resistance [123, 124]. Assessment of the status of IM drug transporters hOCT-1 and P-glycoprotein revealed a significant upregulation of P-glycoprotein in K562/R and KU812/R cells (Fig. 5.13-A,C) which in turn reflected in about 50% reduction of intracellular IM (Fig.5.13-D). KCL22/R on the other hand showed no significant difference in level of drug transporters yet the concentration of intracellular IM was reduced by 80% with respect to KCL22/S+IM cells (Fig.5.13-B,D). This reduction could be attributed to the changes associated with other imatinib transporters like ABCG2 or OCTN2 [125-127], not evaluated in this study. From the above results, IM-resistance in K562/R cells seem to be mediated by Bcr-Abl overexpression and altered drug transport, in KCL22/R cells by hindrance to IM binding due to mutation in kinase domain along with reduced intracellular IM and in KU812/R cells solely by altered drug transport resulting in reduced intracellular IM (Table 5.4). Hence, it was anticipated that inhibition of Bcr-Abl activity by IM in K562/R, KCL22/R and KU812/R cells would not be as efficient as it was in their sensitive counterparts, which would

thereby mediate resistance. However, assessment of Bcr-Abl activity by western blotting (Fig.5.14) revealed that in contrast to the anticipated outcome, Bcr-Abl activity in K562/R and KU812/R cells were inhibited to the same extent as that of their sensitive counterpart. However, as expected, a sub-optimal inhibition of Bcr-Abl tyrosine kinase activity was observed in KCL22/R cells. This implied that (a) resistance in KCL22/R is likely mediated by Bcr-Abl dependent mechanism, due to residual Bcr-Abl activity in these cells (b) despite effective inhibition of Bcr-Abl activity by IM, the oncogenic potential is retained in K562/R and KU812/R cells, thereby indicating the involvement of molecules other than Bcr-Abl in providing survival advantage to these cells.

As compared to the studies reported in literature, we attempted to generate a more robust assay system which took care of diversity of blasts observed in CML-BC and was well characterized for the known mechanisms of resistance to IM.

2. A global proteomic analysis would help in delineating the molecular alterations or pathways other than Bcr-Abl, involved in resistance. Of the 2 cell lines that harbour Bcr-Abl independent resistant mechanism (K562/R and KU812/R), only K562 cell line was subjected to extensive proteomic analysis and the outcome was validated in KU812 cell line. One of the major lacunae in the reported proteomic studies on CML-BC cell lines is the limited proteome coverage due to use of restricted approach, which failed to yield adequate information on low abundant biologically important proteins. Studies that employed 2D-MS approach could only yield a maximum of 46 differentiators (Table 5.2, 5.3). It is known [128, 129] that better pre-fractionation of complex mixtures of protein was required to improve identities of proteins.

We therefore employed nano- LC-MS based quantitative proteomic profiling by label-free SWATH-MS method and labelled iTRAQ method to improve resolution and sensitivity of protein identification.

3. In addition to taking care of the lacunae which we observed in previous studies on CML cell lines, we also addressed a key issue in MS-data analysis which we believed could improve the assessment of translational potential of the data. In case of SWATH-MS data analysis, lack of experimental demonstration for the choice of (a) normalization method and (b) criteria for differentiator identification were the two major unattended issues, addressed in this study. Towards this, a quantitatively well-defined dataset A obtained from public domain served as reference set while datasets B, C and D containing wild type K562-S cells along with those manipulated extraneously with IM treatment (K562-S+IM & R), thereby depicting the heterogeneity observed in biological samples, constituted the study set. Datasets E and F comprised the validation set, which was included to substantiate the findings from the above datasets. The standards used for retention time calibration, depth of spectral ion library and the peak intensities varied among the study, reference and validation sets representing the diverse methodologies employed by SWATH studies in literature. Inclusion of such diverse sample sets in the study helped in identifying a strategy for SWATH-MS data normalization with universal applicability.

RLE plot of un-normalized datasets (Fig.5.16) revealed that while quantitatively defined dataset A showed no variation, it increased gradually from undefined dataset B to datasets C,D and F, which could be ascribed to increasing sample heterogeneity with increase in sample size. Dataset E, an undefined large dataset, however showed minimal variation which could probably be attributed to experimental precision but not routinely observed due to experimental errors. Overall, these observations highlight the need for SWATH-MS data normalization.

In most of the previously reported SWATH-MS studies, data has been normalized by TI [130-133], TAS [134-142], quantile [143-145], median [105, 146-148] and IS [149, 150] methods. In this study, to identify the optimum normalization method, datasets A-F were normalized using 10 normalization methods from Normalyzer and 2 methods from Marker View software, which included the above-mentioned methods used in previous studies. Thus, assessment of efficiency of these twelve different methods to normalize the data based on the statistical criteria of PCV and

RLE plot, identified VSN-G as the method that could effectively normalize all datasets irrespective of their associated heterogeneity (Fig.5.17), implying its wide applicability. Valikangas et.al., have also reported that, for normalizing label-free proteomic data generated using LTQ-orbitrap, VSN method was found to be optimal [109]. The applicability of this method (VSN-G) in biomarker discovery, for which SWATH-MS is widely used, was assessed by its potential to yield differentiators that could precisely segregate comparison groups, using cluster analysis.

With respect to differentiator identification, PubMed search for SWATH-MS studies revealed that of 45 studies, 31.1% studies have used statistical significance i.e. p-value, 13.3.% studies have used fold change while 55.6% have used a combination of p-value and fold change as criteria for choosing differentiators. While for transcriptomic data, studies have experimentally evaluated the choice of differentiator identification criteria [151, 152] or their cut-off values [153] , no such reports are available for MS data. This study for the first time provides experimental evidence for the choice of criteria for differentiator identification from SWATH-MS data for biomarker discovery, using cluster analysis, based on their potential to segregate control and test groups precisely which is an essential feature of biomarkers. From cluster analysis (Fig.5.19) it was evident that differentiators identified from data normalized by twelve methods, using the criteria of p-value, could efficiently segregate the comparison groups in five of out of six datasets used. Hence for differentiator identification in this study, p-value was chosen as the optimal criteria .

VSN-G, the method identified as optimal for normalizing SWATH-MS data based on statistical criteria however, failed to yield differentiators that precisely segregated comparison groups, groups in all six datasets (Fig.5.18), thereby challenging its utility in biomarker discovery. On the other hand, Loess-R which was not top rated by PCV and RLE plots, was the only method to yield differentiators which could not only precisely cluster comparison groups (Fig. 5.19) but also sub-cluster the technical and biological replicates in with utmost efficiency in all datasets evaluated (Fig.5.20), thereby demonstrating its suitability for biomarker discovery using SWATH-MS. This could probably be attributed to the distinction in assumptions made by these normalization

methods. VSN-G normalization bring the samples onto the same scale by making the sample variances non-dependent on their mean intensity. Based on this assumption in order to achieve optimal normalization, the differences in intensities between the samples are reduced, which is not conducive towards identification of differentiators that could segregate the comparison groups. Loess normalization on the other hand, probably retains the differences between intragroup protein intensities by assuming non-linear relationship between biases in the data and the magnitude of protein intensity - a feature essential for segregation of comparison groups [154].

We thus demonstrated that while choosing a normalization method for biomarker discovery using SWATH-MS data, apart from statistically recommended criteria, it is essential to evaluate the method using, a biologically relevant criteria like precise stratification of comparison groups. This study thus for the first time identified Loess-R based normalization and p-value based differentiator identification as optimal for SWATH-MS data.

4. There are two possible ways by which IM resistance could be induced in spite of Bcr-Abl tyrosine kinase activity inhibition as seen in our system. (i) By activation of a novel signalling pathway, not mediated by Bcr-Abl, that provides survival advantage to the cells (Fig.6.1.A). (ii) Due to presence of molecular alterations that restore Bcr-Abl downstream oncogenic signalling by activating the pathway components in a Bcr-Abl independent manner (Fig. 6.1.B). Thus, in this study, based on Bcr-Abl activity and the extent of its inhibition by IM, the comparison groups for proteomic analysis were K562/S vs S+IM to identify components of Bcr-Abl downstream signalling pathway and K562/S+IM vs R to identify components that confer resistance. This would help in gaining insight on whether the observed alterations contribute to resistance by activating a novel pathway that provides survival advantage or by modulating components of Bcr-Abl signalling pathway.

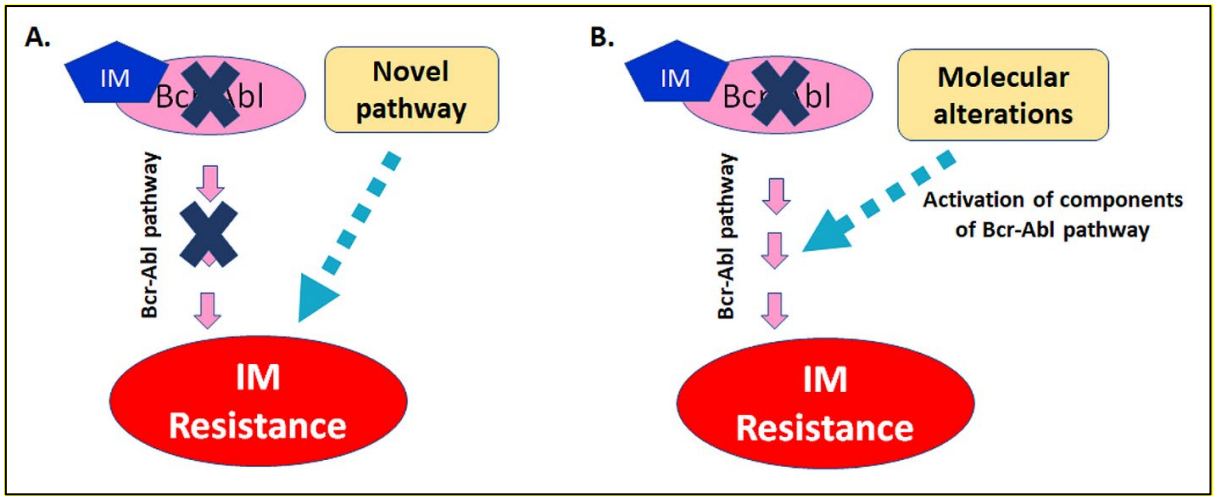


Fig. 6.1. Plausible Bcr-Abl independent IM-resistant mechanisms: *(A) Activation of novel signalling pathway that provides survival advantage, which is not regulated by Bcr-Abl. (B) Presence of molecular alterations that activate components downstream of Bcr-Abl signalling pathway, thereby maintaining the pathway active, despite inhibition of Bcr-Abl activity.*

WHAT DID OUR PROTEOMIC ANALYSIS YIELD FROM THIS REFINED ASSAY SYSTEM?

The differentially expressed proteins identified from SWATH and iTRAQ were pooled and subjected to further analysis to get a comprehensive picture. Comparison of K562/S vs S+IM yielded 416 differentiators while that of K562/S+IM vs R yielded 730 differentiators as against 169 and 56 differentiators obtained from meta-analysis respectively. This increase in number of differentiators was thought to improve the outcome of functional association network to delineate Bcr-Abl downstream components and those involved in resistance. Assessment of overlap among differentiators identified by this study and meta-analysis revealed presence of 26 common differentiators in S vs S+IM comparison group (Fig. 6.2-A) and 16 common differentiators in S+IM vs R comparison group (Fig. 6.2-B).

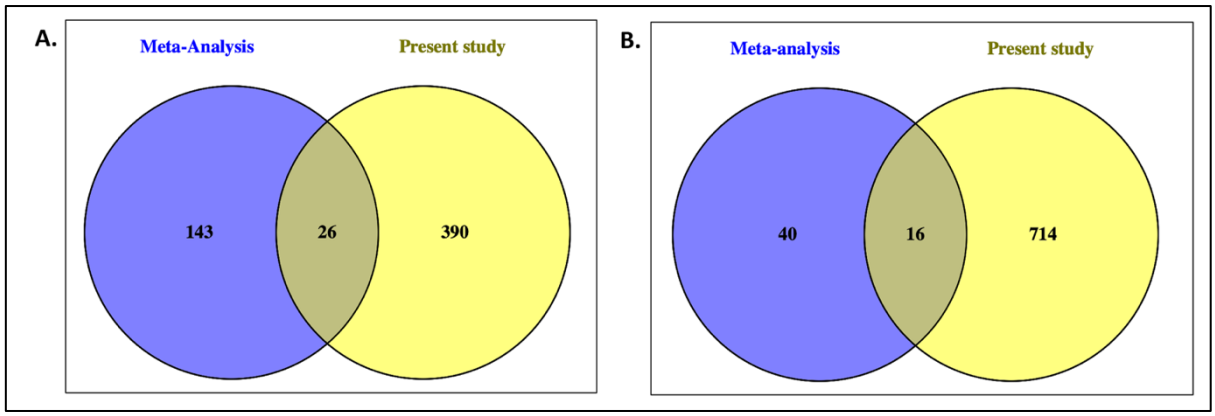


Fig. 6.2. Overlap of differentiators identified from studies in meta-analysis and present study: *(A) Differentiators from S vs S+IM comparison group. (B) Differentiators from S+IM vs R comparison group.*

Functional categorization of differentiators from both the comparison groups (Table 5.6) revealed a complete overlap among the biological processes modulated upon inhibition of Bcr-Abl activity in sensitive cells and in resistance, but with more number of proteins found to be altered in each biological process in resistant cells, implying the possibility of Bcr-Abl downstream processes being modulated further in resistance. Sub-categorization of signaling molecules, apart from identifying proteins that are altered only in resistant cells (Fig. 5.21-C), also indicated the presence of some common differentiators between both the comparison groups (Fig. 5.21-B) confirming that Bcr-Abl downstream components are indeed modulated in resistant cells. In a situation where IM resistance is mediated by restoration of Bcr-Abl oncogenic signaling, identifying and targeting the altered Bcr-Abl pathway components would help in alleviating IM resistance. Hence to further assess which Bcr-Abl downstream components were altered in resistance, a STRING analysis was carried out using differentiators identified from K562/S+IM vs R comparison group, with Abl as one of the inputs (Fig. 5.22-B) and a functional hierarchy of Abl was established (Fig. 5.23-D). A similar functional hierarchy with Abl was established from STRING analysis of proteins modulated upon IM treatment (K562/S vs S+IM), which served as a reference for Bcr-Abl downstream pathway components (Fig.5.22-A, 5.23-C).

From the above interactions, one key observation was that in both the comparison groups, the functional association of Abl with multiple proteins converged onto 14-3-3 family proteins indicating their potential as hub molecule. Apart from this, Grb2, STAT-5 and MAPK-14 were common to both the comparison groups and hence were found to be the components of Bcr-Abl pathway modulated in resistance. Grb2 being an adapter protein failed to qualify as a molecule that could be explored for its potential as therapeutic target for IM-resistant CML. STAT-5 being a well-established Bcr-Abl downstream target, its role in resistance and potential as therapeutic target has been extensively studied [155-159]. The stress induced Ser/Thr kinase, MAPK-14 (p38-α MAPK), though known for its role in resistance in various cancers is not explored much in CML. Thus 14-3-3 family proteins and p38 MAPK were identified as key differentiators from proteomic analysis.

DID THE DIFFERENTIATORS IDENTIFIED IN THIS STUDY ENHANCE OUR UNDERSTANDING OF IM RESISTANCE IN CML-BC OR IDENTIFY A THERAPEUTIC TARGET?

Validation of key differentiators identified in this study by western blotting revealed that 14-3-3 ε apart from being a component of Bcr-Abl downstream pathway is also modulated in resistance (Fig. 5.24-B). With respect to p38-MAPK, no significant change in levels was observed, probably owing to the specificity of antibody towards 3 different isoforms α, β and γ of p38-MAPK (Fig.5.25-A). However, phosphorylation of p38-MAPK was found to be reduced upon inhibition of Bcr-Abl activity in K562/S cells which significantly increased in K562/R cells (Fig. 5.25-B), suggesting its involvement in Bcr-Abl signaling and increased activity in resistance. Assessment of the status of 14-3-3 ε and p38-MAPK in KU812/R cells (Fig.5.26), which also harbors Bcr-Abl independent resistant mechanism, revealed a similar modulation. This consistency observed among heterogenous IM-resistant CML-BC cell lines (K562 of erythroleukemic and KU812 of basophilic lineage) further substantiates the association of 14-3-3 ε and p38-MAPK with IM-resistance. Hence, in this study 14-

3-3 ε and p38-MAPK were further investigated to understand their role in IM-resistance and assessed for their potential as therapeutic targets.

14-3-3 proteins are a family of 28-33 kDa small acidic scaffolding proteins that specifically bind to phosphorylated Ser/Thr residues of the target proteins, thereby interacting with more than 700 proteins involved in a variety of cellular function. While some functions are found to be redundant among its 7 isoforms, there are increasing reports suggesting the existence of isoform specific functions. Of these, 14-3-3 ε was found to play a vital role in regulation of growth factor receptor signaling due to its ability to specifically interact with signaling molecules that undergo extensive tyrosine phosphorylation [160]. With reference to the role of 14-3-3 proteins in CML, limited reports are available. Mancini *et.al.*, demonstrated that in presence of p210 Bcr-Abl, 14-3-3 σ sequesters c-abl preventing its nuclear localization and c-abl induced apoptosis [161]. However, no report associating 14-3-3 ε with Bcr-Abl signaling or CML is available. In this study, to check if reduction of 14-3-3 ε level leads to development of IM-resistant phenotype, a 14-3-3 ε knockout was generated in K562/S cells using CRISPR-cas9 system which resulted in a moderate yet significant increase in IC_{50} of IM (Fig.5.28), confirming its contribution towards IM-resistance.

p38-MAPK is a member of stress activated protein kinase family, associated with multitude of biological functions including cell cycle regulation, apoptosis, actin cytoskeleton reorganization and cytokine production [114]. It exhibits a dual role in tumorigenesis wherein it can suppress [162] or promote tumor growth and drug resistance in a context dependent manner [163, 164]. Even in case of CML, contradictory reports on the role of p38-MAPK were obtained. Parmar *et.al.,* using KT-1 cell line have shown that IM treatment inhibits the growth of cells expressing Bcr-Abl, in a p38-MAPK dependent manner by phosphorylating and activating the pathway. Inhibition of p38-MAPK activity reversed the anti-proliferative effect induced by imatinib, thereby making the cells unresponsive [165]. Kohmura *et.al.* on the other hand, though reported a similar activation of p38 -MAPK upon treating K562 cell line with IM, showed that inhibition of p38-MAPK activity affected only differentiation of K562 cells without altering the anti-proliferative effect induced by IM in these

cells [166]. In this study however, in contrast with the above reports, reduction in phosphorylation of p38-MAPK was observed upon IM treatment in K562/S cells, which significantly increased in K562/R cells (Fig. 5.25-B). To understand if this increased p38 activity contributes to IM resistance, p38-MAPK activity was inhibited in K562/R cells using inhibitor SB203580. The inhibition sensitized K562/R cells to IM by reducing their IC_{50} by 7-fold (Fig.5.30), demonstrating the importance of p38-MAPK signaling in conferring IM resistance as well as the potential of its inhibition in overcoming IM-resistance in CML-BC. A similar observation was made by Giafis *et.al.,* wherein p38-MAPK signaling was found to confer resistance to As_2O_3 treatment in CML-derived KT-1 cell line [167].

Inhibition of p38-MAPK has been identified as a potential therapeutic strategy to overcome resistance in multiple cancers like glioblastoma, metastatic breast cancer, myelodysplastic syndrome and ovarian cancer as well as in treating various heart diseases, based on which many inhibitors have entered phase I or phase II clinical trial as a single agent or in combination with other drugs [164]. While many inhibitors have failed due to their adverse side effects, losmapimod is one p38-MAPK inhibitor that resulted in good tolerability among patients with myocardial infarction in phase II study and has entered phase III clinical trial [168, 169], upholding the prospect of clinical use of p38-MAPK inhibitor.

In summary, based on identification of key differentiators from proteomic analysis and their functional validation, the study identified an association between downregulation of 14-3-3 ε, a key component of Bcr-Abl pathway and imatinib resistance in CML-BC, thereby providing a lead to further delineate the upstream modulators of 14-3-3 ε which would help in gaining deeper understanding on how Bcr-Abl pathway is modulated in resistance. This study also endorses that p38-MAPK phosphorylation is an event which confers resistance by modulating Bcr-Abl pathway component independent of Bcr-Abl activity, the mechanism B proposed in Fig 6.1 Further, upon inhibition of Ser/Thr kinase activity of p38-MAPK using specific inhibitor, the sensitivity of K562/R

cells to imatinib was restored. Thus, this study reports for the first time that p38-MAPK is a potential therapeutic target for IM-resistant CML-BC (Fig 6.3).

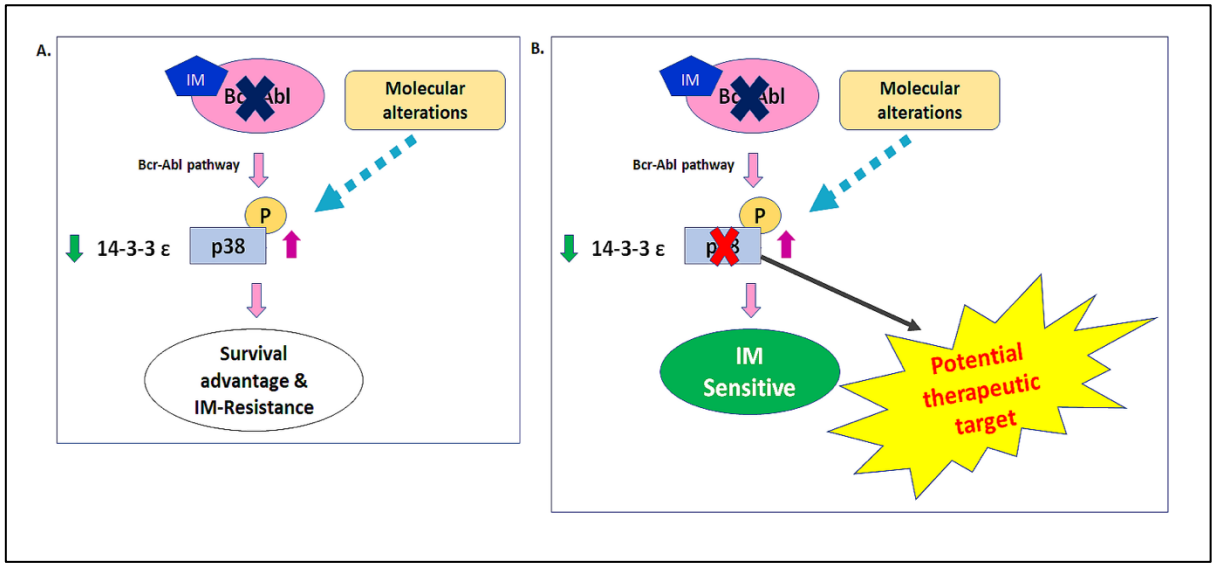


Fig.6.3. Summary of key findings in this study: *A. Modulation of downstream components of Bcr-Abl pathway i.e., downregulation of 14-3-3 ε and increased phosphorylation of p38-MAPK, in a Bcr-Abl independent manner, contributes to development of IM-resistant phenotype. B. Inhibition of Ser/Thr kinase activity of p38 sensitizes the cells to IM, thereby signifying its potential as therapeutic target to overcome to IM-resistance.*

7. CONCLUSION

Chronic myeloid leukemia, though considered as an epitome of targeted therapy, is not impervious to the problem of drug resistance. About 40% patients in advanced stage of CML fail to respond to any of the existing TKIs and this brings in a dire need to identify alternate therapeutic targets for treating CML-BC. Using proteomic analysis of CML-BC cell lines this study for the first time identified that downregulation of 14-3-3 ε, a key component of Bcr-Abl pathway, contributes to development of IM-resistant phenotype in CML-BC cell lines. This in turn suggests that identifying its upstream modulators would help in gaining a better insight on the modulation of other Bcr-Abl pathway components as well as role of 14-3-3 ε in IM-resistance.

In a finding that has translational potential, this study demonstrated for the first time that p38-MAPK activity, which was found to be upregulated in IM-resistant cells, can be targeted to overcome resistance in CML. Further being a component of Bcr-Abl downstream signalling, inhibition of its activity would also alleviate Bcr-Abl dependent resistance by inhibiting Bcr-Abl downstream signalling in CML-CP and BC wherein IM resistance is due to kinase domain mutations (Fig. 7.1).

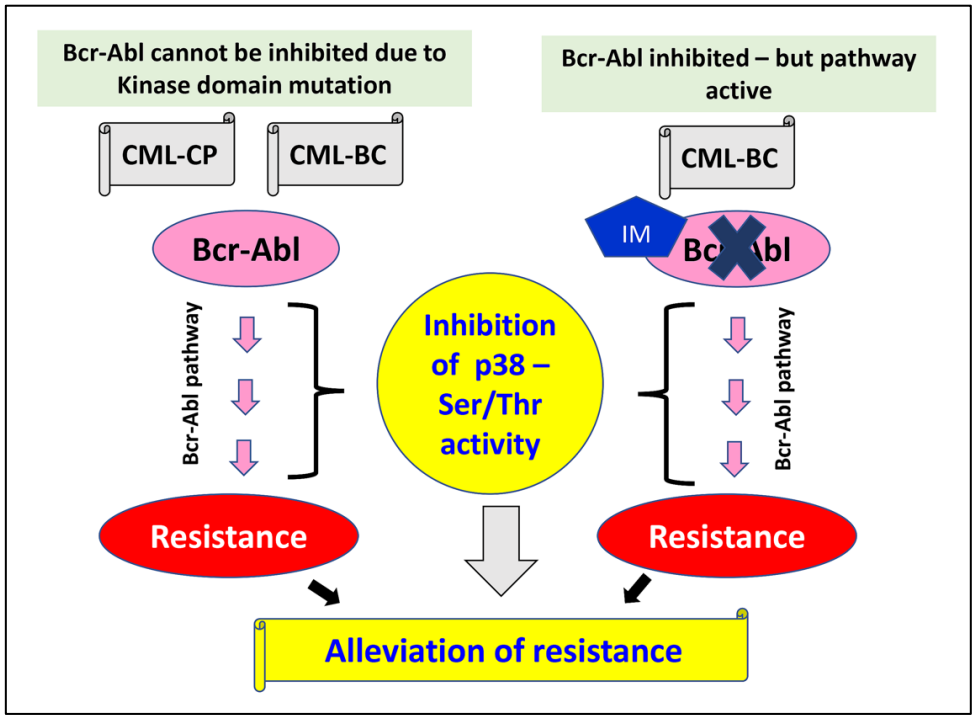


Fig. 7.1. Contribution of this study to alleviate IM resistance: *p38-MAPK a component of Bcr-Abl pathway, is identified as the potential therapeutic target for alleviating IM-resistance CML by inhibiting Bcr-Abl downstream signaling pathway.*

8.APPENDIX - 1

Table A1: Data extraction sheet for K562 +/- IM comparison group – List of differentiators

Study 1	**Study 2**	**Study 3**	**Study 3 (contd.)**	**Study 4**
Uniport acc.no.	**Uniport acc.no.**	**Uniport acc.no.**	**Uniport acc.no.**	**Uniport acc.no.**
Q99766	P23588	P01106	Q92878	P69905
P30043	P31943	Q54AF1	P15259	P69891
P18669	Q9UQ80	P55957	P00938	P02100
P51970	P11021	Q14790	P14618	P69892
Q9NQT5	P78371	Q07817	P00338	P68871
Q8TCB7	P63241	O43521	P07738	P69892
P07951	P06748	P98170	P05091	P22830
O14645	P18206	P53350		P36551
P17022	P43487	P57775		P08397
Q99729	P82970	Q92560		Q01581
Q15765	P52565	Q96Q27		P00352
Q03403	Q9UQ80	Q13191		P62750
P42773	P11021	Q13616		P62424
O15519	P38646	P61086		P62081
Q00169	P04083	Q9NWF9		P61353
Q9UQB9	P55072	P05198		P18124
Q00987	Q9UQ80	P13667		Q9NYK5
P52565		P07237		P09651
Q16816		P18850		P62304
Q9NZH6		P18848		P60228
P01241		P17861		Q92522
P62258		O75460		P84243
P35247		Q53QY0		P16403
P29016		P30291		Q99613
P21730		Q9UJX3		P05198
Q96NS1		P30281		P62807
P49223		Q00534		P68431
Q8IUB2		P11802		P0C0S8
P00742		P45974		P62805
Q92688		Q5LJB1		P04818
P47944		Q5VZ98		O00116
O43715		Q7Z6M2		Q8WVY7
Q93083		Q8IX29		P31040
P19883		P62736		P43304
P28072		P20929		P22392
P48739		O75116		P31153
Q9BY27		Q13509		Q9Y4E8
		P08670		P07858
		P11177		P61457
				Q13423

Table A1 (contd.): Data extraction sheet for K562 +/- IM comparison group – List of differentiators

Study 4 (contd.)	**Study 5**
Uniport acc.no.	**Uniport acc.no.**
P09874	O00571
Q9Y376	P49006
P52292	Q5T1J5
O43264	Q9Y6H1
Q9UG63	Q13067
P47755	Q13069
Q8WW33	Q13070
Q9Y5A9	
P27816	
P35658	
P20290	
Q969Z0	
Q9UN86	
P31689	
P08195	
P13639	
Q9BZK3	
P14735	
P17844	
Q8NE71	
Q92890	
Q13765	
Q12849	
P62310	
Q99536	
Q8IX12	
P45880	
SEC61B	

Table A2: Data extraction sheet for K562 S vs R comparison group – List of differentiators

Study 1	**Study 3**
Uniport acc.no.	**Uniport acc.no.**
P07226	P11142
Q9H939	P20700
P06493	P13796
P35247	O60506
Q96NS1	P08670
P49223	Q71U36
P20337	O43175
P11233	P12268
O14645	P14618
P17022	Q9BWF3
O95755	Q99729
O00287	Q15181
P04637	P04406
Q9H165	Q8WUM4
P61758	P40926
Q99729	P23193
P42773	Q12904
Q9NZH6	P78417
P01241	P54819
Q00169	Q15056
Q93038	P30040
Q96PN8	O95336
Q16816	P62826
P50152	
P62262	
P01913	
P30041	
P00736	
P02760	
Q8IUB2	
Q03403	
Q9H9S4	
Q93083	
P00747	

Table A3: Differentiators identified by SWATH analysis from K562 S vs S+IM comparison group

Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)
HBZ	6.58	UNC45A	2.37	MCTS1	1.78	BLVRB	1.50
KRT6C	5.99	CSNK2A2	2.37	COPG1	1.76	NDUFV2	1.50
MRPL10	4.35	JUP	2.37	UROD	1.73	MTHFD1	1.50
TMEM9	3.97	CSTB	2.36	HMBS	1.73	RAE1	1.49
HBG2	3.96	ATP6V0C	2.33	ATP5J	1.72	CSTF2	1.49
HBG1	3.76	DYNLRB1	2.31	PPIE	1.72	NUTF2	1.49
HEMGN	3.70	PXN	2.27	TRIM33	1.70	PGM3	1.48
HRSP12	3.69	RAB35	2.25	MPC2	1.70	POLR2B	1.47
SUCLG1	3.66	SLC25A24	2.25	NARS	1.70	HNRNPUL2	1.47
SLAIN2	3.57	DNPH1	2.25	BPNT1	1.70	CAPN1	1.46
GNS	3.44	GLB1	2.23	HSDL2	1.69	HDAC1	1.46
SNX2	3.37	PPIH	2.23	MRPS31	1.68	MYL6	1.46
METAP1	3.27	RECQL	2.23	THRAP3	1.65	SMARCA4	1.45
KRT16	3.24	PLD3	2.22	HNRNPM	1.65	CTSB	1.45
NDUFS5	3.23	RPS6KA3	2.22	CD59	1.64	MDN1	1.45
EHD1	3.16	TMSB10	2.20	PDHB	1.63	SLC25A1	1.45
UQCRQ	3.13	EPN1	2.15	PRPF40A	1.63	HP1BP3	1.44
SACM1L	3.04	AKAP1	2.14	SELENBP1	1.63	RHOG	1.44
AFG3L2	3.04	BAX	2.09	APIP	1.63	TMED10	1.44
DPM1	3.01	AP2M1	2.05	TRA2A	1.63	FUS	1.43
CREG1	2.99	PPP2R4	2.05	TCEA1	1.62	CARS	1.43
CAST	2.95	PCYOX1	2.04	VAT1	1.61	SARNP	1.42
LYAR	2.95	NPM3	2.04	TAGLN2	1.60	RTCB	1.42
EML3	2.84	GGCT	2.04	STMN1	1.60	CTSL	1.41
BRE	2.83	PYM1	2.01	POLD2	1.59	PCMT1	1.41
PPP2R2A	2.75	CIAO1	2.00	NONO	1.59	SEC63	1.41
STAT5B	2.74	PCYT2	1.99	ARF5	1.59	PSMC4	1.40
PAK2	2.72	DCD	1.96	CTSD	1.58	MIF	1.40
TYMS	2.69	PRKAR2B	1.95	PGD	1.58	MCM5	1.40
SLC2A1	2.66	EWSR1	1.94	RBM12	1.58	PGK1	1.40
ADD2	2.65	SEPHS1	1.94	DYNC1H1	1.57	GRB2	1.40
PYCR2	2.61	TAF15	1.90	QDPR	1.57	HMGB2	1.40
PIN1	2.60	PTBP3	1.88	COPE	1.56	CCAR2	1.39
COPS5	2.58	KDM1A	1.86	OAT	1.54	AKR1C2	1.39
CPNE3	2.57	HADHA	1.86	HDGF	1.51	RPA2	1.39
TRMT1L	2.56	HMGB1	1.84	UBE2V2	1.51	PPIL1	1.39
CTSH	2.56	SNRPA	1.82	THOP1	1.51	HNRNPH3	1.39
PRRC2A	2.54	NUP210	1.82	SPTA1	1.51	AARS	1.39
MRPL46	2.49	CSTF3	1.80	CAT	1.51	MAP4	1.38
MTX1	2.44	SEPT8	1.80	CORO1C	1.51	LUC7L2	1.38
ACOT13	2.46	HUWE1	1.79	GATAD2A	1.51	WARS	1.37

Table A3 (Contd.): Differentiators identified by SWATH analysis from K562 S vs S+IM comparison group

Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)
MTA2	1.37	SRSF10	1.29	SRI	1.23	TTLL12	1.19
RBM10	1.37	NANS	1.29	PICALM	1.23	XRCC6	1.19
UCHL3	1.37	PGRMC2	1.29	PSIP1	1.23	GSTO1	1.19
FUBP1	1.37	CRKL	1.28	TKT	1.23	CDC37	1.19
UPF1	1.36	VARS	1.28	HNRNPU	1.23	ELAVL1	1.18
LRRC47	1.36	HAT1	1.28	PHGDH	1.23	ATIC	1.18
ECI1	1.36	ATP6V1A	1.28	RBMX	1.23	SFPQ	1.18
SF3B1	1.36	HNRNPD	1.28	FARSB	1.23	ALDOA	1.18
CNIH4	1.36	VTN	1.27	XPO1	1.22	ALYREF	1.18
DEK	1.36	CKB	1.27	PCBP1	1.22	TUBB4B	1.18
ACADVL	1.35	PRPSAP2	1.27	SRP14	1.22	STIP1	1.18
CIAPIN1	1.35	GSTK1	1.27	TPI1	1.22	G3BP2	1.18
ANP32B	1.35	RAB1B	1.27	TARDBP	1.22	ALDH18A1	1.18
MYL4	1.35	SRSF3	1.27	ATP5J2	1.22	AK2	1.18
UBQLN1	1.35	HDLBP	1.27	PSMD12	1.22	EDF1	1.17
PLIN3	1.35	PEBP1	1.26	LBR	1.22	TMED9	1.17
IMMT	1.35	SNRNP40	1.26	SF3B2	1.22	MSN	1.17
MDH1	1.34	AKR7A2	1.26	UBE2I	1.21	ABCF1	1.17
HSPB1	1.34	TALDO1	1.26	ETFA	1.21	ACLY	1.17
SMARCA5	1.34	ENSA	1.26	TMPO	1.21	MAPK14	1.17
MGST3	1.34	EIF4E	1.25	CPSF6	1.21	DDX6	1.17
DRG1	1.34	FLNA	1.25	VASP	1.21	ST13	1.17
RNPEP	1.34	RNH1	1.25	GANAB	1.21	SGTA	1.17
SMU1	1.33	PSMD3	1.25	CAPZB	1.21	ETFB	1.17
SLC25A13	1.33	PABPN1	1.25	FH	1.21	MARS	1.17
HNRNPH2	1.33	PSMA6	1.25	TACO1	1.21	EPRS	1.16
PUF60	1.33	SRRM2	1.25	FEN1	1.21	HNRNPR	1.16
IGF2BP3	1.33	AP2A1	1.25	FKBP5	1.21	CLTC	1.16
ILF3	1.32	ADH5	1.25	MTPN	1.20	TUBB	1.16
KIF5B	1.32	ACADM	1.24	DDX17	1.20	PFN1	1.16
ISYNA1	1.31	PGAM1	1.24	HNRNPC	1.20	ACO2	1.16
GSR	1.31	PSMD5	1.24	MCM2	1.20	FDPS	1.15
ANP32E	1.31	PSMD6	1.24	PDCD6	1.20	ERP29	1.15
TXNDC17	1.31	TCEB1	1.24	ACAT1	1.20	RALY	1.15
ACTL6A	1.31	HNRNPH1	1.24	ANP32A	1.20	OSTC	1.15
SOD1	1.31	GLO1	1.24	HEXB	1.19	PHB2	1.15
ACSM3	1.30	THOC3	1.23	PCNA	1.19	MAT2B	1.15
GSTP1	1.30	PAFAH1B2	1.23	CAPZA1	1.19	PRDX3	1.15
PPT1	1.30	RCC2	1.23	PSMA7	1.19	SRP9	1.15
PARK7	1.30	ATP5C1	1.23	PTBP1	1.19	NDUFS8	1.14
GNPDA1	1.30	U2AF2	1.23	MPP1	1.19	CNDP2	1.14

Table A3 (Contd.): Differentiators identified by SWATH analysis from K562 S vs S+IM comparison group

Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)
MSH2	1.29	VKORC1L1	0.68	NOC2L	0.85
SAE1	1.14	SLC3A2	0.65	DDX21	0.85
ILF2	1.14	ARHGEF2	0.65	TMEM33	0.84
IARS	1.23	SON	0.64	ADSS	0.84
U2AF1	1.19	YWHAB	0.62	RPS3	0.83
SHMT2	1.14	EIF4G2	0.61	AHSA1	0.81
EIF6	1.14	HEATR1	0.60	SLC1A5	0.80
VCP	1.14	GRWD1	0.60	RBM25	0.79
ATP5O	1.13	RRM2	0.59	NTMT1	0.78
PARP1	1.13	WDR74	0.59	SLC25A6	0.77
STOML2	1.13	UCK2	0.53	BRIX1	0.77
RPL23	1.12	CHP1	0.51	EIF3K	0.76
PSMA5	1.12	THOC1	0.48	ALDH1A2	0.76
PDCD5	1.11	NUP35	0.42	GTPBP4	0.76
CYCS	1.11	DCTD	0.37	SAMSN1	0.73
OLA1	1.10	TFB2M	0.36	TPT1	0.72
GOT2	1.08	UBE2S	0.30	RRS1	0.72
EIF4G1	0.90	PNPT1	0.70	LTV1	0.70
MAT2A	0.89	ZPR1	0.68		
MRPS27	0.86	EIF3L	0.68		

Table A4: Differentiators identified by iTRAQ from K562 S vs S+IM comparison group

Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)	Gene Name	Fold change (S+IM/S)
HBG2	6.08	DDX21	0.41	EEF1B2	0.28
KRT10	1.76	NUDC	0.41	RPL7	0.28
NDE1	1.61	CCT2	0.40	DDX5	0.26
RBBP4	1.22	PA2G4	0.39	EEF1G	0.23
ORC2	1.00	RPL3	0.34	RPS3A	0.23
VIM	0.95	RPL4	0.34	RPS4X	0.21
EEF2	0.95	CCT8	0.33	RPLP2	0.21
EEF1A1	0.94	TARS	0.32	NFXL1	0.08
NHP2	0.64	RPL7A	0.32		
ABCF2	0.62	RPS2	0.30		
RPL22	0.48	HSPA8	0.30		
EIF3A	0.42	FASN	0.28		

Table A5: Differentiators identified by SWATH analysis from K562 S+IM vs R comparison group.

Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)
RPL35	20.06	MEPCE	3.77	TBCE	2.71	CKB	2.40
GTSF1	10.49	REEP6	3.71	HSPB1	2.70	MAK16	2.36
RDH10	8.93	DHX30	3.61	HSPA4L	2.69	IDH1	2.34
RNF40	7.99	MAP2K2	3.57	GALE	2.69	SURF4	2.31
TELO2	7.32	TLN1	3.51	RPS15	2.69	PPWD1	2.29
ANP32B	6.89	BTF3L4	3.43	STAM	2.68	RPL32	2.29
RSU1	6.47	PFKP	3.41	NMT1	2.68	UBE2V2	2.29
HMGA1	6.18	NAA10	3.41	PLP2	2.67	ANP32A	2.28
TMEM43	6.13	FN3K	3.41	VCL	2.66	GID8	2.28
HYPK	6.13	UNC13D	3.39	COPG1	2.64	DNTTIP2	2.27
ENY2	5.98	LUC7L	3.34	EPHX2	2.64	PPP1R7	2.27
POLE3	5.96	RBM17	3.33	CIRBP	2.63	UNC45A	2.27
HCLS1	5.68	CDK2AP1	3.31	CIAO1	2.62	IMPA1	2.25
PPP1R14A	5.43	CUTA	3.29	SMC1A	2.62	VCP	2.25
C12orf57	5.38	PPP1R14B	3.28	ARHGDIB	2.61	MYL6	2.25
HBE1	5.27	FLNB	3.20	ATP5L	2.60	TCOF1	2.25
TRIP13	4.96	ACTN1	3.17	CTSB	2.59	ARL3	2.24
PTRF	4.84	DCTN2	3.13	S100A11	2.58	NUBP2	2.23
SERPINB1	4.79	MACROD1	3.10	COPB1	2.54	QKI	2.23
AK1	4.66	CAPZA2	3.08	PSMB7	2.54	PRKRA	2.23
MAGEB2	4.62	NUP155	3.07	TUBB6	2.53	LONP1	2.22
CALB1	4.56	PDLIM1	3.04	GCN1	2.53	GSE1	2.22
MTFR1L	4.51	ALDH2	3.03	RAD50	2.53	CCDC58	2.21
UBXN1	4.48	POR	3.01	MCTS1	2.52	DYNLRB1	2.21
CD44	4.47	IQGAP1	2.98	RAD23A	2.52	MAPK1	2.20
H2AFX	4.45	RALA	2.97	MAPK14	2.50	MBD3	2.19
C14orf142	4.42	RPS7	2.96	ADSS	2.49	SH3GL1	2.19
NT5C	4.38	UBE2S	2.94	PKM	2.49	ZC3H4	2.19
COTL1	4.27	HSPBP1	2.94	ISOC1	2.48	ECHS1	2.19
DSTN	4.27	AKR1A1	2.93	UTP14A	2.48	PPP5C	2.17
MRFAP1	4.08	SNX5	2.90	NUP205	2.47	MYBBP1A	2.17
CWC15	4.06	TWF2	2.88	PSAP	2.47	TMA16	2.17
ALDOC	4.03	RAB10	2.85	UQCRH	2.44	PA2G4	2.16
SERPINB9	3.99	KPNA4	2.85	HNRNPH3	2.44	SNRPGP15	2.15
SEC16A	3.98	SNRNP200	2.84	H2AFY	2.44	CTH	2.15
NUP62	3.96	SACM1L	2.83	HK1	2.43	NUCB2	2.14
LZIC	3.87	MRPS5	2.81	EXOSC7	2.43	SYAP1	2.13
AAGAB	3.85	PLIN2	2.78	PWP1	2.42	SRSF7	2.13
SLC1A5	3.83	ANXA2	2.77	MDC1	2.41	ABI1	2.12
SNRPD1	3.78	LGALS1	2.77	POLR1C	2.40	PLEC	2.12
SAR1A	2.72	ATG3	2.72	SERPINB6	2.40	PSMC1	2.12

Table A5 (contd.): Differentiators identified by SWATH analysis from K562 S+IM vs R comparison group.

Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)
FAM49B	2.11	MAGEC1	1.90	EBNA1BP2	1.71	SNAP23	1.62
YWHAG	2.11	PNP	1.90	HNRNPL	1.70	RBMX	1.62
CMBL	2.10	MAP4	1.89	EIF5A	1.70	OGDH	1.61
SOD1	2.09	SH3GLB1	1.89	PLIN3	1.70	SF1	1.61
HMGN2	2.08	CHMP4A	1.88	STAT5A	1.70	RAP1B	1.61
FERMT3	2.08	PRDX1	1.87	EIF5	1.70	FIP1L1	1.61
TPP1	2.06	PNN	1.87	SPC24	1.69	SRRM1	1.61
FAH	2.06	CFL1	1.87	EMD	1.69	CLIC1	1.61
CDV3	2.06	PCK2	1.87	SNAP29	1.68	UBE2L3	1.60
SMC3	2.06	NUP133	1.86	CRYZ	1.68	APMAP	1.60
NUMA1	2.05	CTPS1	1.86	BPNT1	1.68	RNPS1	1.59
UCHL5	2.04	MED15	1.85	LASP1	1.68	GOT1	1.59
IDH3B	2.04	RPL8	1.85	HNRNPA0	1.68	NIFK	1.59
GNL3	2.04	SRSF11	1.84	VAT1	1.67	HNRNPC	1.59
ANP32E	2.03	CCDC47	1.83	NPM3	1.67	COMT	1.59
PABPN1	2.02	LMNA	1.83	CBS	1.67	RPS17	1.59
SRSF5	2.00	PRPF4	1.82	RPL9	1.67	RPL7	1.59
SRM	1.99	PAFAH1B2	1.82	PPA2	1.66	EPRS	1.59
DNAJC9	1.99	KRT8	1.82	RPS28	1.66	RPL4	1.59
FARSA	1.99	DDX21	1.81	COPS4	1.66	RPS12	1.58
TRIP6	1.99	IPO4	1.81	TPT1	1.66	CCT2	1.58
UBR4	1.99	RPS13	1.81	RPL14	1.66	PRKDC	1.58
TSR1	1.98	NOL6	1.81	UBAP2L	1.65	TNPO1	1.58
ISYNA1	1.98	CYFIP1	1.80	VASP	1.65	PRPSAP2	1.58
NBAS	1.97	BANF1	1.79	FRG1	1.65	TAGLN2	1.57
KRT19	1.97	HIST1H4A	1.78	HINT1	1.65	BLMH	1.57
LSM8	1.96	PRKAR2A	1.78	RBM4	1.64	RPL18	1.56
BCL7B	1.95	NCL	1.78	MTA2	1.64	TRMT112	1.56
MGST3	1.94	DIAPH1	1.77	RPL23A	1.64	ACIN1	1.56
CDC42	1.94	AHNAK	1.76	STMN1	1.64	CAP1	1.56
AHSA1	1.94	CDK1	1.76	TROVE2	1.64	MRPL13	1.56
TXNRD2	1.92	YWHAE	1.76	PSMD4	1.63	CHERP	1.55
HPRT1	1.92	PACSIN2	1.76	PPM1F	1.63	NQO1	1.55
CNDP2	1.92	PGLS	1.76	CSTF3	1.63	ARHGEF2	1.55
ANXA5	1.92	RPL10	1.75	OTUB1	1.63	ATXN2L	1.55
BCCIP	1.91	PRDX2	1.74	PDCD5	1.63	EFHD2	1.55
LAS1L	1.91	UBTF	1.72	PRDX3	1.63	CHCHD2	1.55
FDXR	1.91	DNM2	1.72	XPO5	1.62	MARS	1.55
BAG2	1.91	ADH5	1.72	CTSL	1.62	HN1	1.55
AGK	1.90	FKBP4	1.71	GRB2	1.62	CISD1	1.55
FLOT1	1.90	HNRNPF	1.71	NUP50	1.62	MAPRE1	1.54

Table A5 (contd.): Differentiators identified by SWATH analysis from K562 S+IM vs R comparison group.

Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)
SLTM	1.54	IPO7	1.44	DNAJA1	1.36	TPI1	1.28
RPS25	1.54	RPL10A	1.44	RPS26	1.36	ARHGDIA	1.28
RPS24	1.53	DHX15	1.43	KPNA2	1.36	SMARCC1	1.27
SOD2	1.53	TCEA1	1.43	RNF2	1.35	ETFA	1.27
PCMT1	1.53	FEN1	1.43	LARS	1.35	VDAC2	1.27
VIM	1.52	ALDH1A2	1.43	PRMT1	1.35	LUC7L2	1.27
PSPC1	1.52	IMPDH2	1.43	PDCD6IP	1.35	LCP1	1.26
HEXIM1	1.52	EIF3K	1.43	STIP1	1.34	MATR3	1.26
VBP1	1.52	PIN4	1.42	NOP2	1.34	DDX17	1.26
DDX3X	1.52	ATP5C1	1.42	TARS	1.34	NME2	1.25
PDAP1	1.51	ACLY	1.42	RPS19	1.33	PGK1	1.25
METAP2	1.51	PML	1.42	ETFB	1.33	COPA	1.25
GNAI3	1.51	LRRC47	1.42	RRM1	1.33	ZYX	1.25
NAMPT	1.50	RPS8	1.41	U2AF2	1.33	RTCA	1.25
HNRNPDL	1.50	G6PD	1.41	ATP6V1A	1.33	TOMM70	1.25
SEPHS1	1.50	LMAN1	1.41	RPS11	1.33	U2AF1	1.25
SMU1	1.50	PPA1	1.41	TRA2A	1.32	PRDX6	1.24
FLNC	1.50	RPL27A	1.41	SAE1	1.32	PFDN6	1.24
DDOST	1.49	HNRNPAB	1.41	NAT10	1.32	HNRNPU	1.24
SRRT	1.49	NAP1L1	1.41	KHSRP	1.32	ST13	1.24
RAB14	1.49	THOC3	1.40	SRP72	1.32	MCM7	1.23
KPNB1	1.49	GTF2I	1.40	GLO1	1.32	PSMC2	1.23
PSMC5	1.48	BYSL	1.40	IARS	1.32	TPD52L2	1.23
ESD	1.48	TIMM44	1.40	C14orf166	1.31	UBA6	1.23
RPL3	1.48	PSMD11	1.40	PAK2	1.31	HADHA	1.23
PLRG1	1.48	TIMM13	1.40	MYH9	1.31	HADHB	1.22
FSCN1	1.47	CDC37	1.40	STAU1	1.31	RPS3A	1.22
SYNCRIP	1.47	RPLP1	1.40	EIF2S3	1.30	ENSA	1.21
AIMP1	1.47	SEC13	1.39	LSM3	1.30	PEBP1	1.20
SRSF9	1.47	UBE2M	1.39	AIFM1	1.30	ATIC	1.19
RPL24	1.46	RAB11B	1.39	SRSF3	1.30	FARSB	1.19
VDAC3	1.45	SEC31A	1.38	CAPN2	1.30	IGF2BP1	1.18
RPS4X	1.45	PSMC4	1.38	TUBA1C	1.29	SSRP1	1.18
HNRNPR	1.45	ALYREF	1.38	ALDOA	1.29	VAPA	1.18
RPL12	1.45	PFN1	1.38	MSN	1.29	PPIB	1.17
GPKOW	1.45	PRDX5	1.37	NUP214	1.28	PHB	1.17
LMNB2	1.45	NPEPPS	1.37	UBA2	1.28	TUBB	1.17
NUDT5	1.44	SARNP	1.37	ETF1	1.28	VDAC1	1.16
EEF1B2	1.44	HSPE1	1.37	DAZAP1	1.28	EEF1D	1.16
DDX5	1.44	AK2	1.36	RPL7A	1.28	HNRNPM	1.15
SUMO3	1.44	UAP1	1.36	OLA1	1.28	SRP9	1.15

Table A5 (contd.): Differentiators identified by SWATH analysis from K562 S+IM vs R comparison group.

Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)
ELAVL1	1.15	NAPA	0.70	DAP3	0.58	SCAMP3	0.46
PABPC1	1.15	PRMT5	0.70	IFITM3	0.58	GTF3C2	0.45
AARS	1.14	AP2A1	0.69	RTN4	0.57	ARHGAP4	0.45
RPL30	1.14	OCIAD1	0.69	DNAJC11	0.57	PCM1	0.45
MTCH2	1.12	MTPN	0.68	NDUFS8	0.57	PRKCB	0.45
CCT3	1.10	ALDH18A1	0.68	CBX1	0.56	GAPVD1	0.45
ACTL6A	0.87	LBR	0.68	USP10	0.56	HBA1	0.45
PSMB5	0.87	ZC3HAV1	0.68	TSFM	0.56	ATP5F1	0.45
EIF4E	0.87	ATP1A1	0.67	PEF1	0.55	LRPPRC	0.44
PDIA3	0.85	PDIA4	0.66	RFC5	0.55	PRRC2C	0.44
PHB2	0.84	DLD	0.66	SDHA	0.55	PSPH	0.44
GOT2	0.84	PYCR1	0.66	PDIA6	0.55	PGAM5	0.44
EIF3C	0.81	PCYOX1	0.66	NDUFS3	0.55	ICT1	0.43
IGF2BP3	0.81	MSH6	0.66	OSTC	0.54	WTAP	0.43
PFKL	0.80	TBL3	0.66	PICALM	0.54	ADD2	0.42
ACTR2	0.80	RAE1	0.65	SEPT8	0.54	SUGT1	0.42
CCT7	0.80	UBE2K	0.65	GSPT1	0.53	ERH	0.42
MTHFD1	0.79	EIF2A	0.65	NDUFB10	0.52	MBOAT7	0.42
COLGALT1	0.79	EWSR1	0.64	SELENBP1	0.52	RPF2	0.42
FH	0.78	CAB39	0.64	NFU1	0.52	GLOD4	0.41
HSD17B10	0.77	PRDX4	0.64	NDUFA5	0.52	NDUFS1	0.41
CFDP1	0.77	PGRMC2	0.63	CD59	0.51	QPRT	0.40
HNRNPH1	0.77	SSR4	0.63	HSDL2	0.51	FKBP8	0.40
ARF1	0.76	DIABLO	0.63	GLRX5	0.51	CSDE1	0.40
LAP3	0.76	PSMA6	0.63	USP14	0.51	IMP4	0.40
DECR1	0.75	HBZ	0.63	HSP90AA1	0.50	COX5B	0.40
GSTK1	0.75	PSMA7	0.62	ERP29	0.50	NDUFS5	0.39
PCBP1	0.75	LMAN2	0.62	USP39	0.50	FAM98B	0.39
MCM2	0.75	ABCF1	0.61	ASNS	0.49	MRPL14	0.39
SLC25A3	0.74	FDPS	0.61	C8orf33	0.49	KRT5	0.39
EXOSC9	0.74	HMBS	0.61	COX6B1	0.49	PSIP1	0.38
HAT1	0.74	MSH2	0.61	TXN	0.49	PPP4R2	0.38
MAT2A	0.74	TK1	0.61	NDUFV1	0.48	NDUFA7	0.37
MCM6	0.73	APOE	0.60	LIG1	0.48	YBX3	0.37
STRAP	0.73	CALR	0.60	MTHFD2	0.48	TSEN15	0.37
IGF2R	0.73	AKR7A2	0.60	MRPL41	0.48	TXNL1	0.37
LARP1	0.73	TFRC	0.59	MRPL49	0.48	COPB2	0.36
PPT1	0.72	ECH1	0.58	CENPV	0.47	BCAT2	0.36
HSP90B1	0.70	ATP5D	0.58	ALDH4A1	0.47	GTF2F1	0.36
NUP35	0.70	ACOT7	0.58	APOBEC3C	0.46	SARS2	0.36
VARS	0.70	RHOXF2B	0.58	NUP93	0.46	TECR	0.35

Table A5 (contd.): Differentiators identified by SWATH analysis from K562 S+IM vs R comparison group.

Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)
CAT	0.35	PTPRC	0.22	CWF19L1	0.30	TUBA1B	0.06
CPSF7	0.35	PZP	0.22	PLD3	0.30	POLD2	0.27
COX7A2	0.34	NAA50	0.22	AIMP2	0.30	MRPS35	0.27
ADD1	0.34	RCOR1	0.22	PXN	0.30	TCEB2	0.26
RPIA	0.34	MAGI1	0.21	UFL1	0.29	ISCU	0.26
ADSL	0.32	NDUFV2	0.20	ITGB1	0.29	COQ5	0.26
UQCRC2	0.32	HRSP12	0.20	EMC1	0.29	GCSH	0.26
BMP2K	0.32	UQCRB	0.20	LTF	0.29	PAGE5	0.25
A2M	0.31	HMGN1	0.19	MRPL23	0.28	ITIH2	0.25
FAM192A	0.31	MT-CO2	0.18	KRT6C	0.27	CPNE3	0.24
AP1M1	0.31	UQCRC1	0.17	UQCRQ	0.27	GRPEL1	0.24
MRPL19	0.31	THOC2	0.15	DCXR	0.27	CPSF3	0.24
PNPT1	0.31	HEBP2	0.13	DPP7	0.27	CYB5R3	0.24
ATP6V0C	0.30	RAB35	0.12	SMAP	0.22	NOSIP	0.24

Table A6: Differentiators identified by iTRAQ from K562 S+IM vs R comparison group

Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)	Gene Name	Fold change (R/S+IM)
HBE	39.67	IPYR	3.17	VIME	0.99
TGM2	10.31	FEN1	3.15	RM19	0.96
VDAC2	8.85	TLN1	3.00	ADT2	0.90
HSPB1	7.47	PRKDC	2.79	USP9X	0.74
KPYM	6.45	SCAM2	2.44	MDC1	0.62
VINC	5.98	LGUL	2.33	EZRI	0.47
ACSM3	5.24	ECHB	2.14	MARCS	0.04
ANXA2	4.86	PPM1F	2.09		
KCRB	4.75	ACADM	1.97		
HMGA1	4.64	K2C8	1.63		
PRDX6	4.46	COIL	1.55		
PDC6I	3.76	ETFB	1.46		
CISY	3.49	3HIDH	1.39		
NOSTN	3.33	FRDA	1.35		
THIL	3.27	SYAC	1.11		
FKBP4	3.18	SRSF4	1.10		
FLNC	3.18	CSN1	1.03		

9.APPENDIX – 2

Dendrograms for cluster analysis of datasets A-F

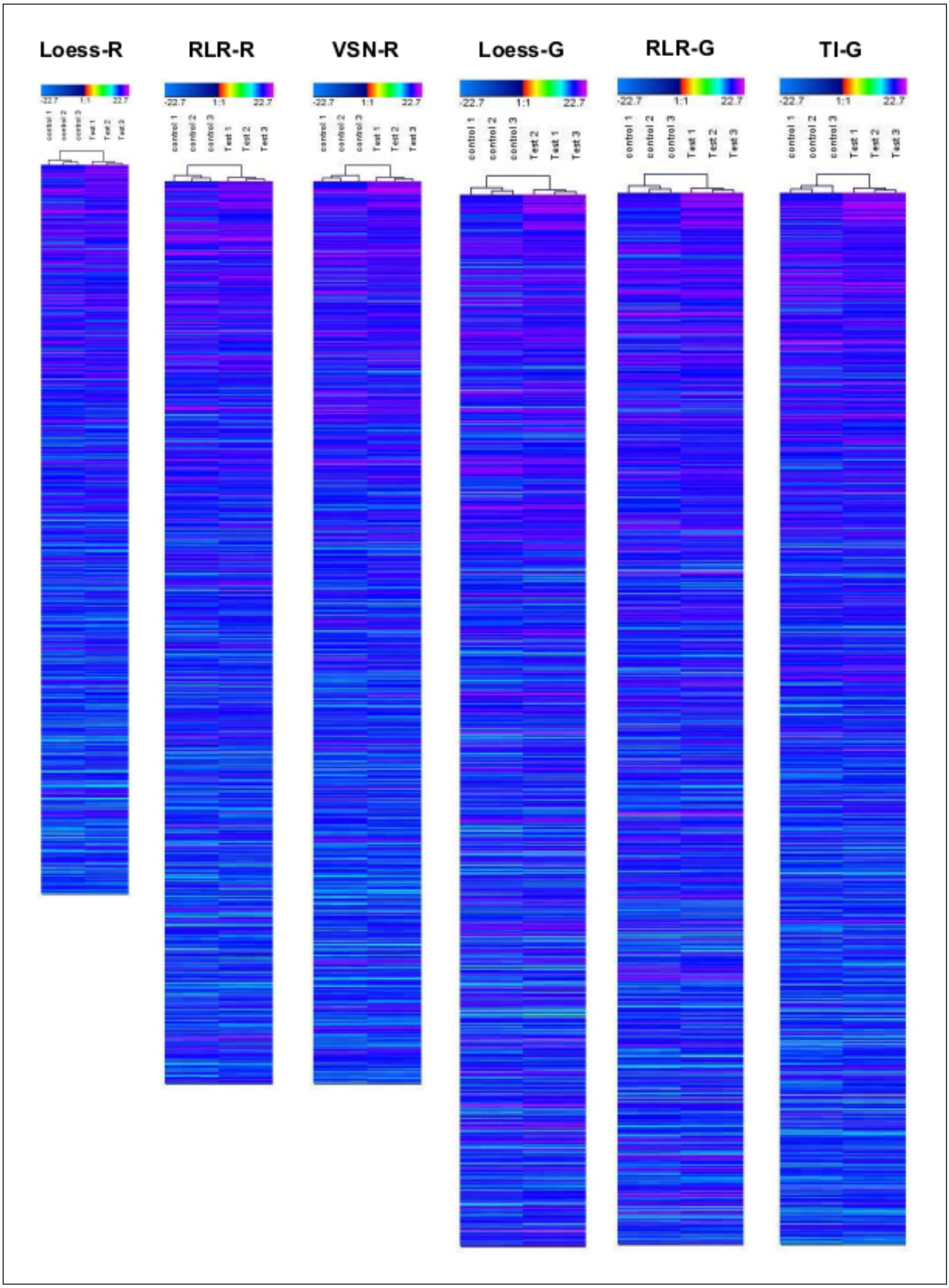


Fig. A1 *– Dendrograms representing hierarchical clustering of control and test groups of dataset A based on differentiators identified using p-value.*

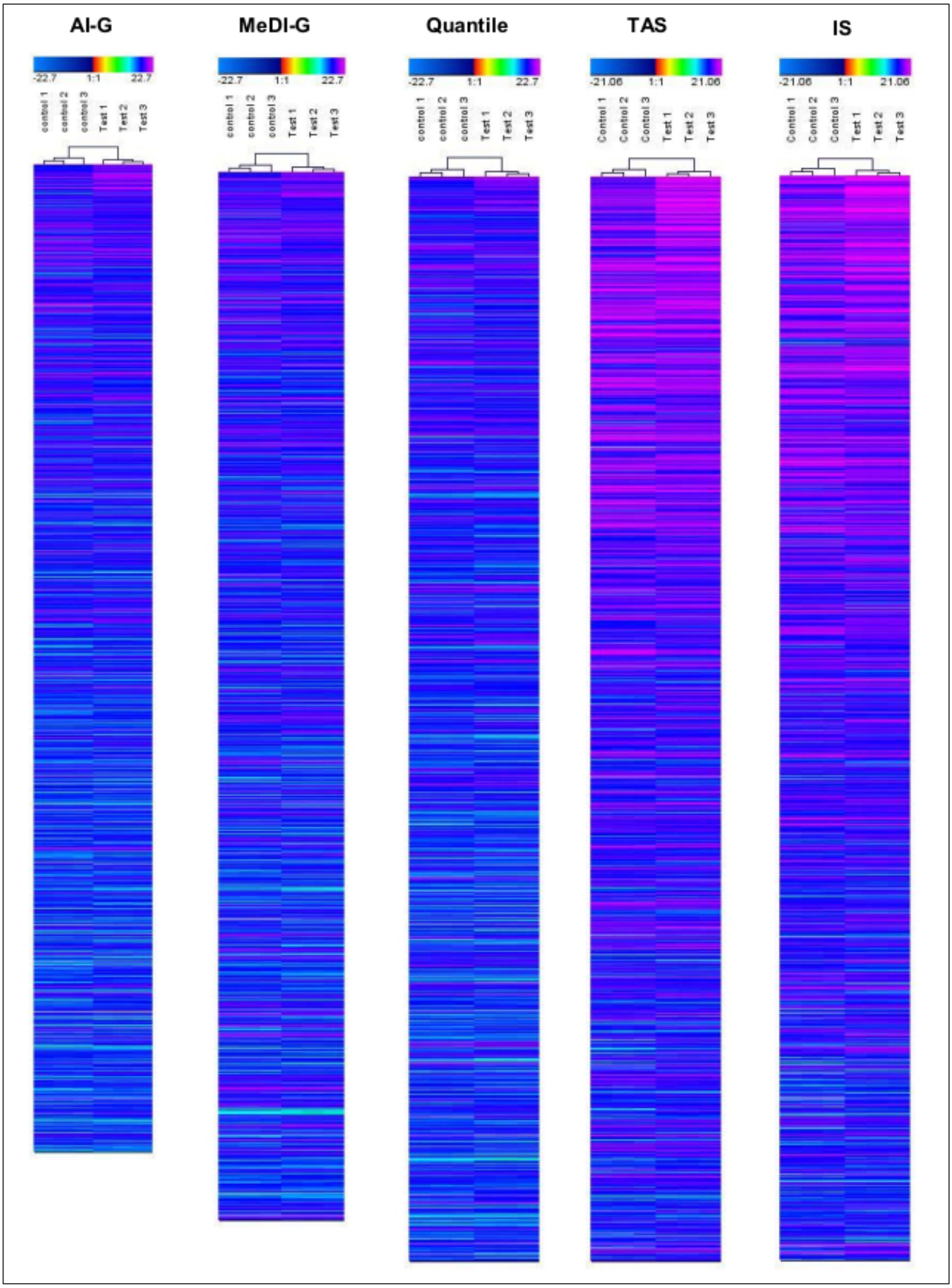


Figure A1 (contd.) - Dendrograms representing hierarchical clustering of control and test groups of dataset A based on differentiators identified using p-value.

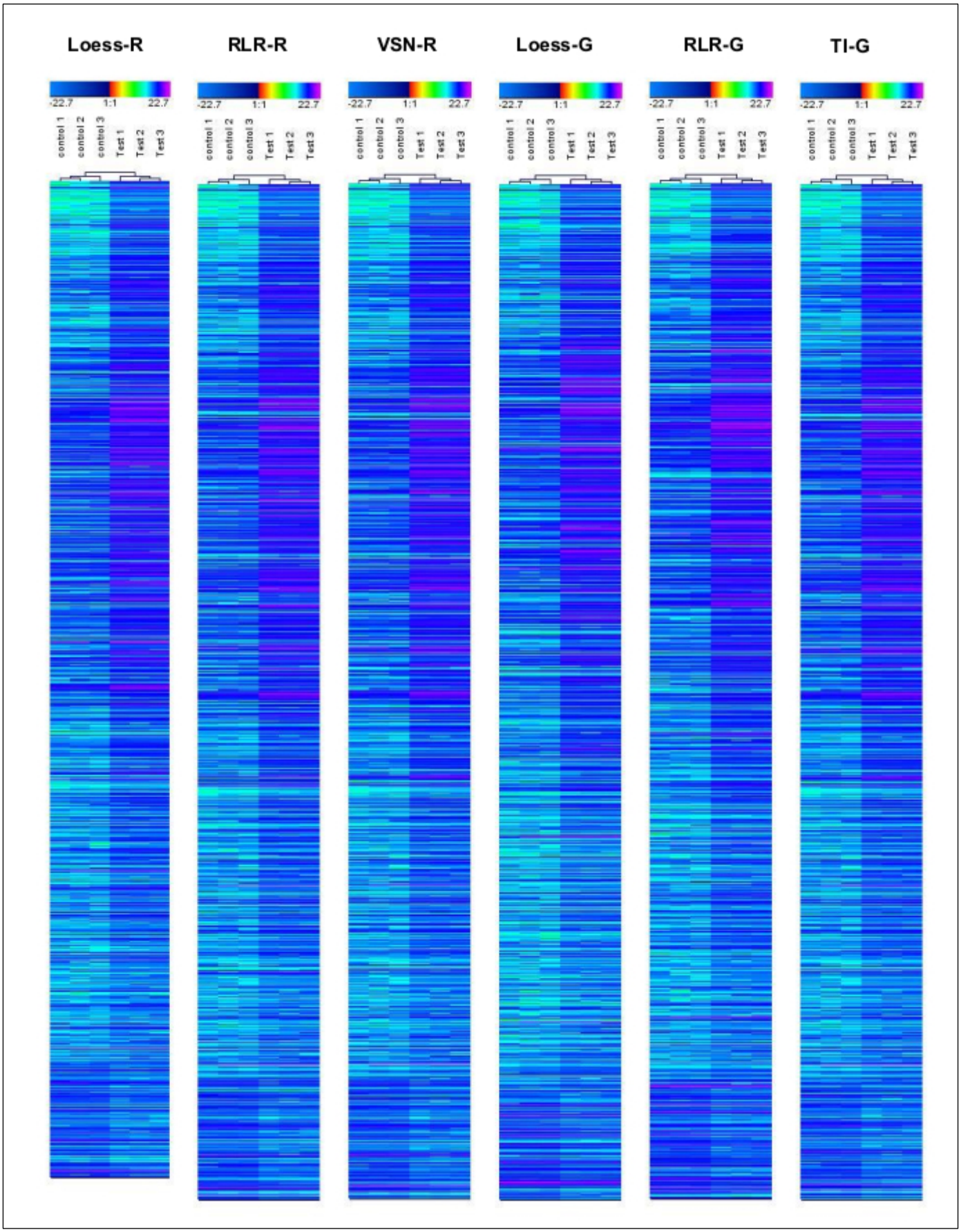


Figure A2 – Dendrograms representing hierarchical clustering of control and test groups of dataset A based on differentiators identified using fold change.

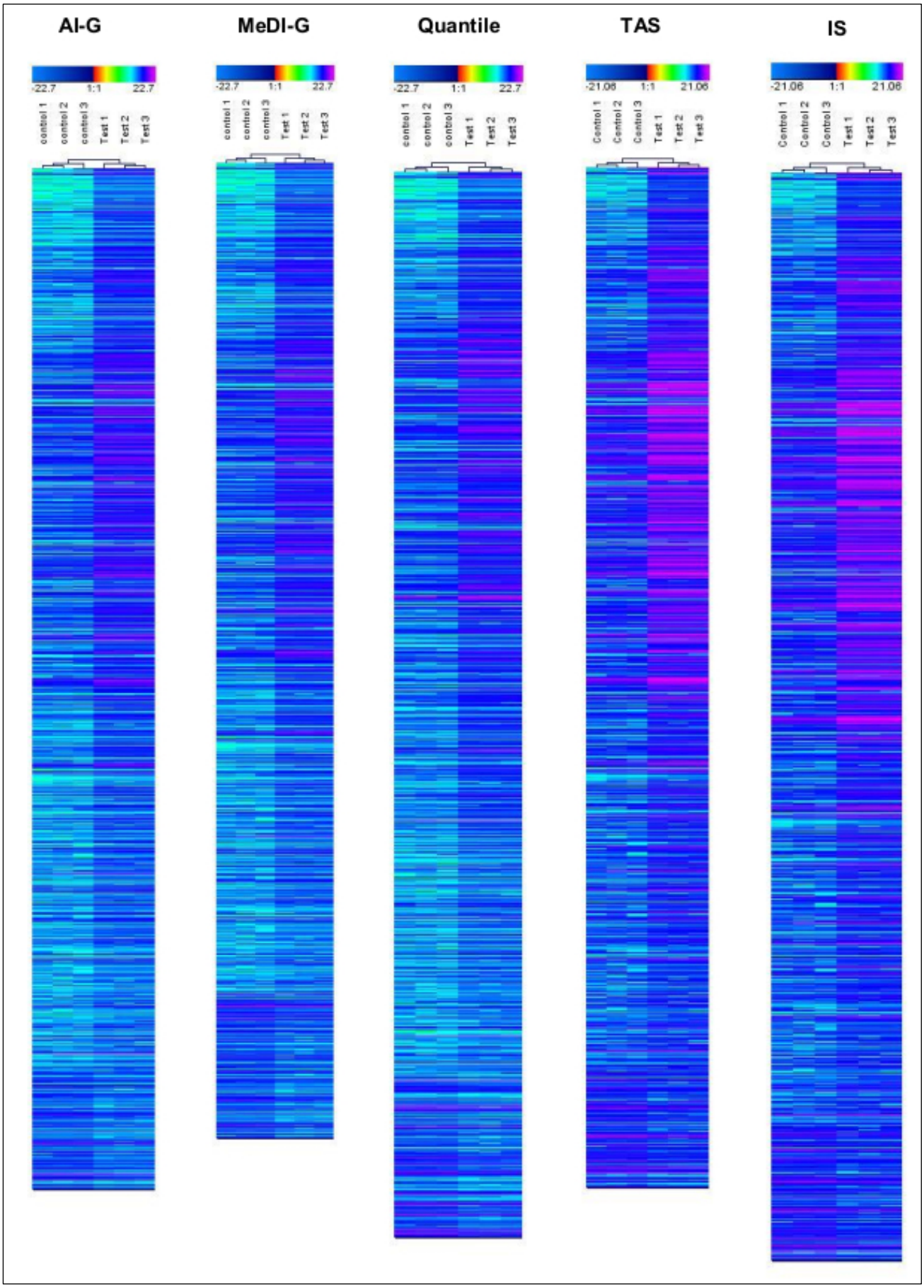


Figure A2 (contd.) – Dendrograms representing hierarchical clustering of control and test groups of dataset A based on differentiators identified using fold change.

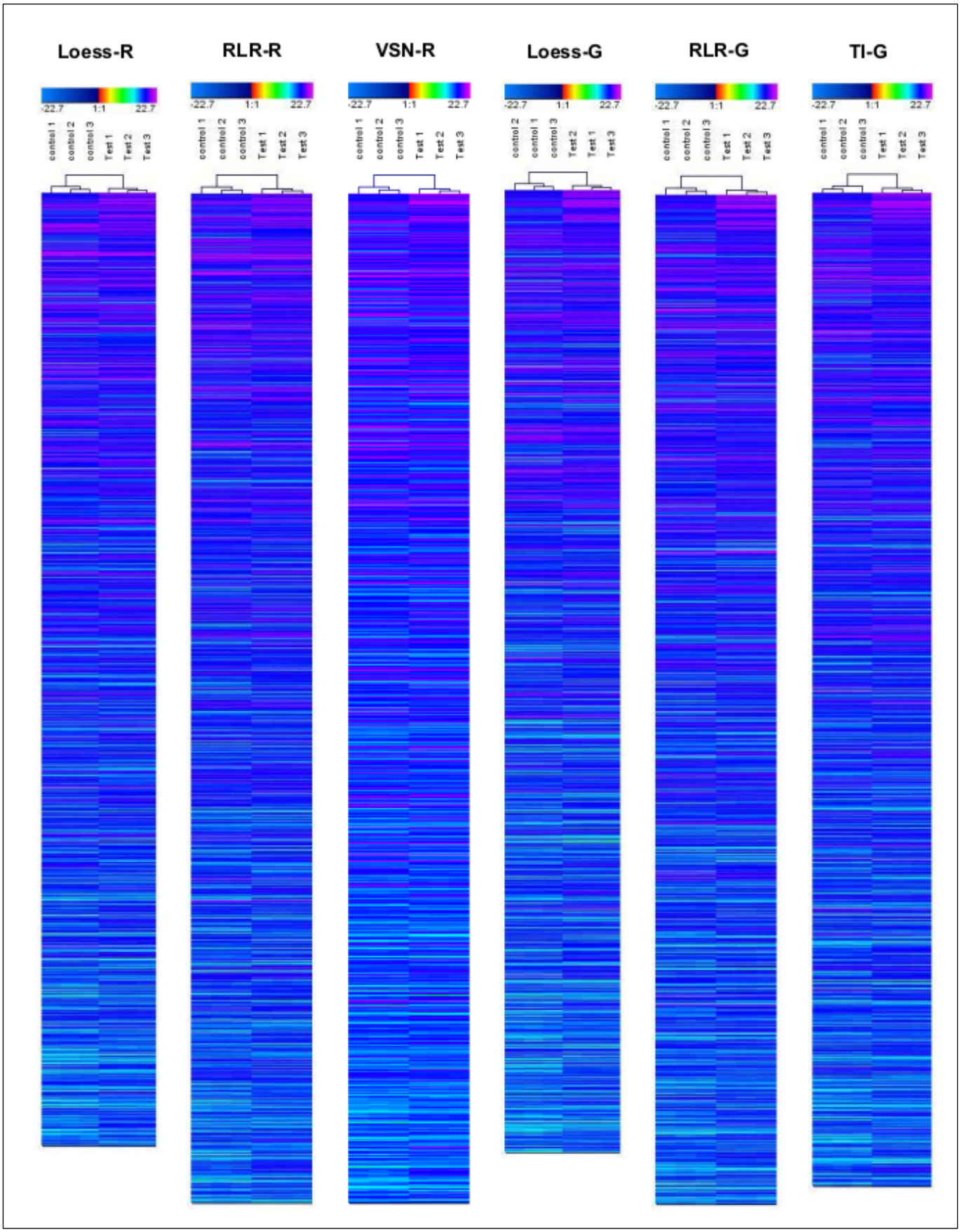


Figure A3 – Dendrograms representing hierarchical clustering of control and test groups of dataset A based on differentiators identified using both p-value & fold change.

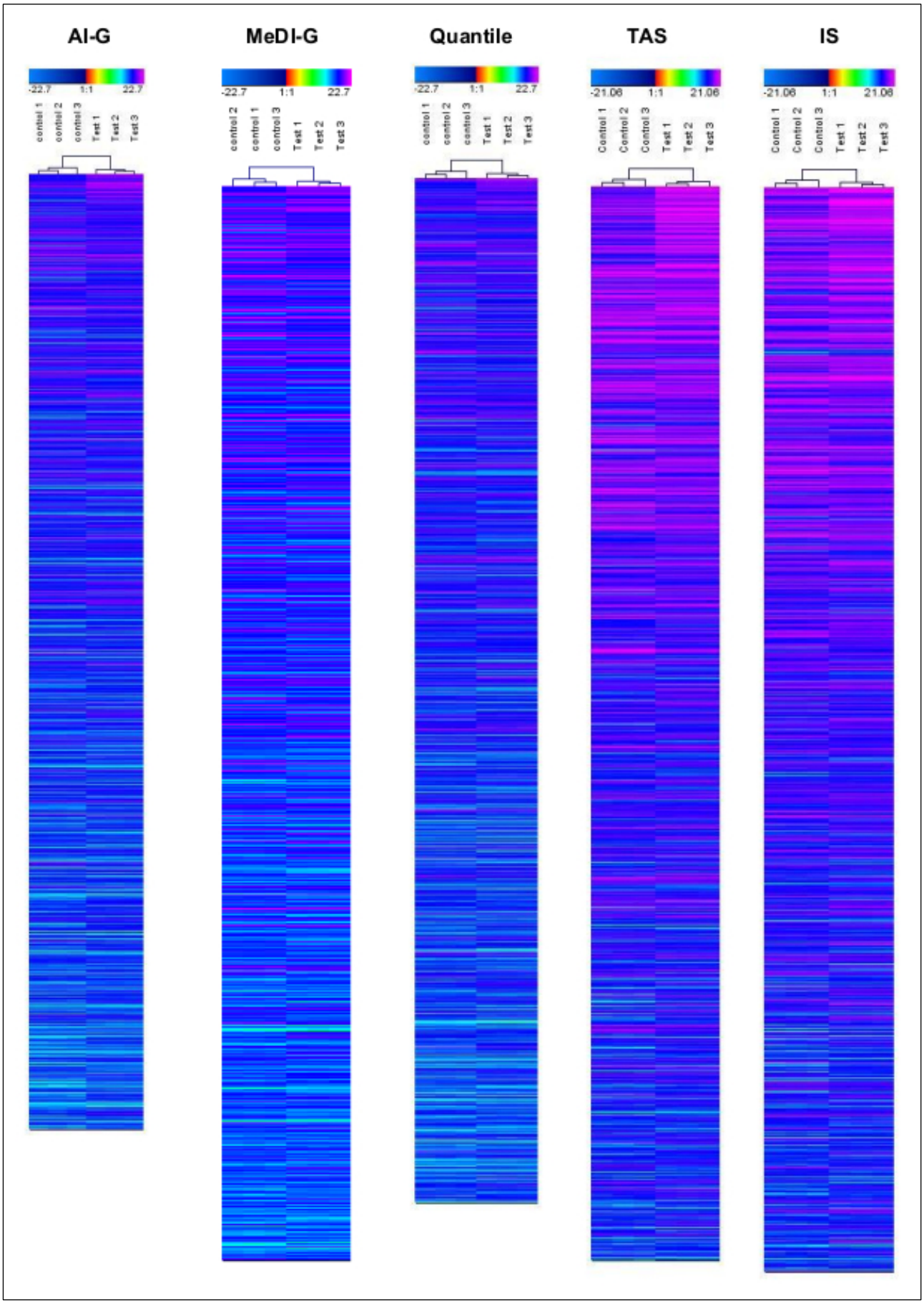


Figure A3 (contd.) - Dendrograms representing hierarchical clustering of control and test groups of dataset A based on differentiators identified using both p-value & fold change.

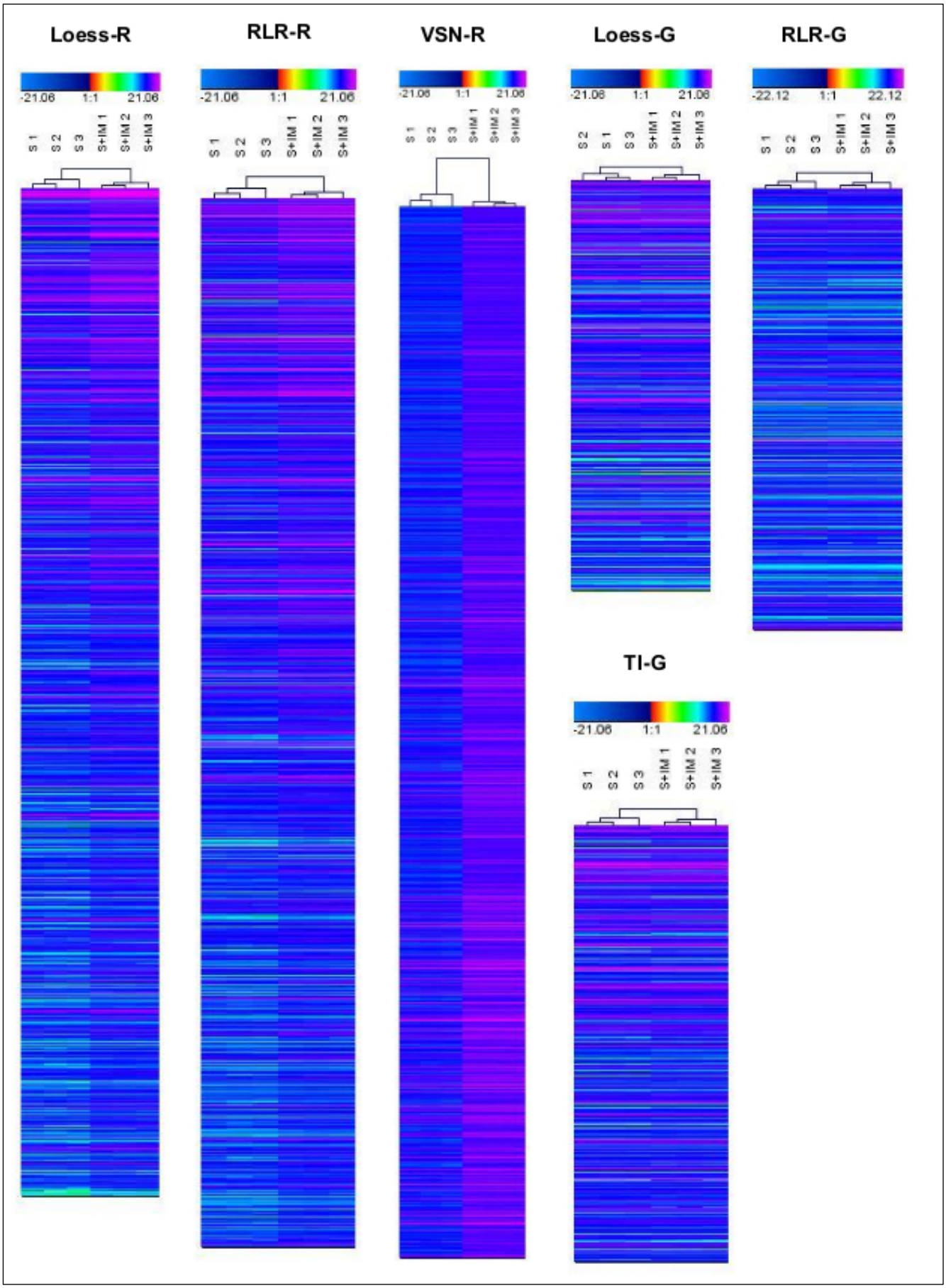


Figure A4 – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset B based on differentiators identified using p-value.

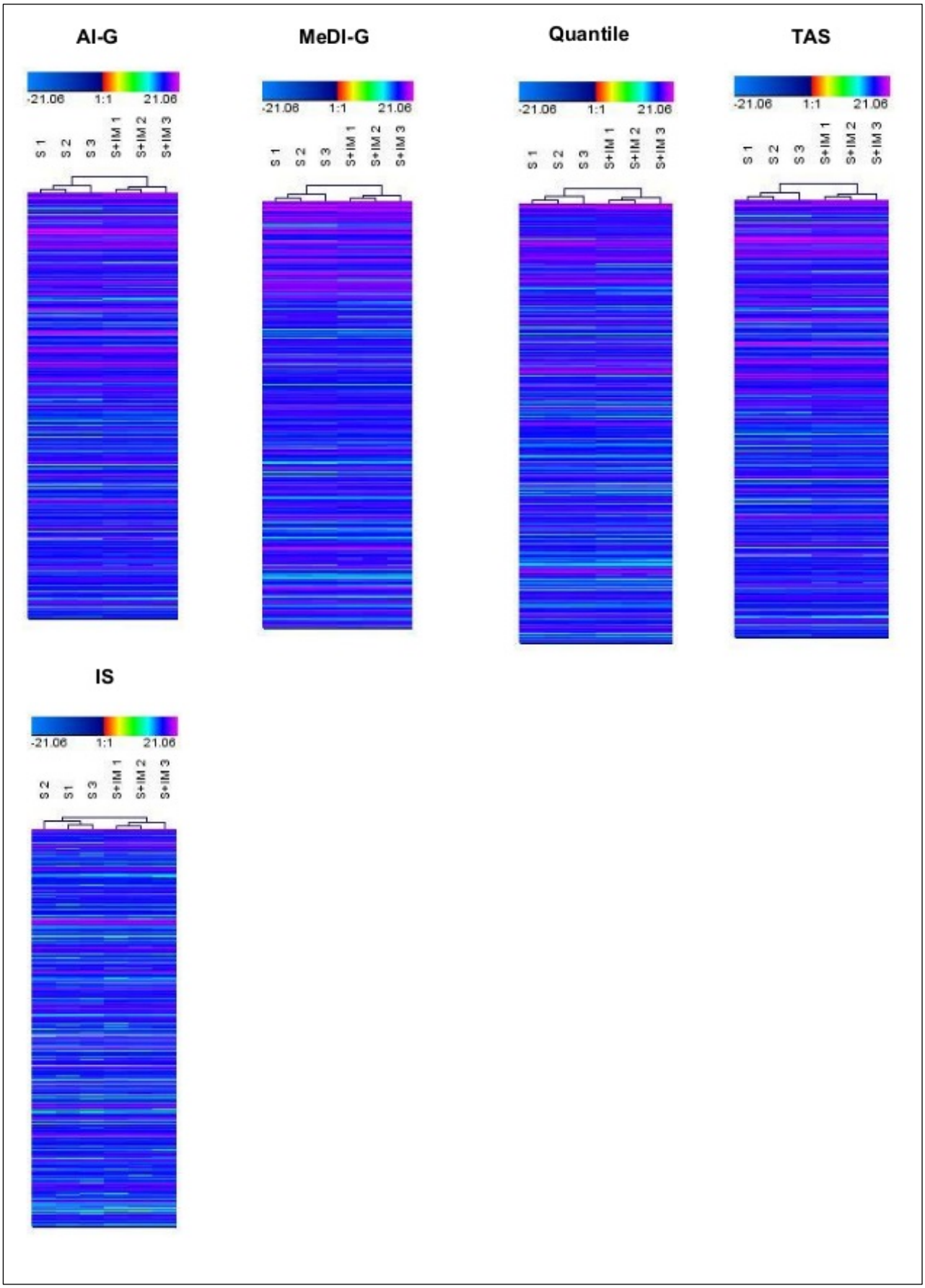


Figure A4 (Contd.) – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset B based on differentiators identified using p-value.

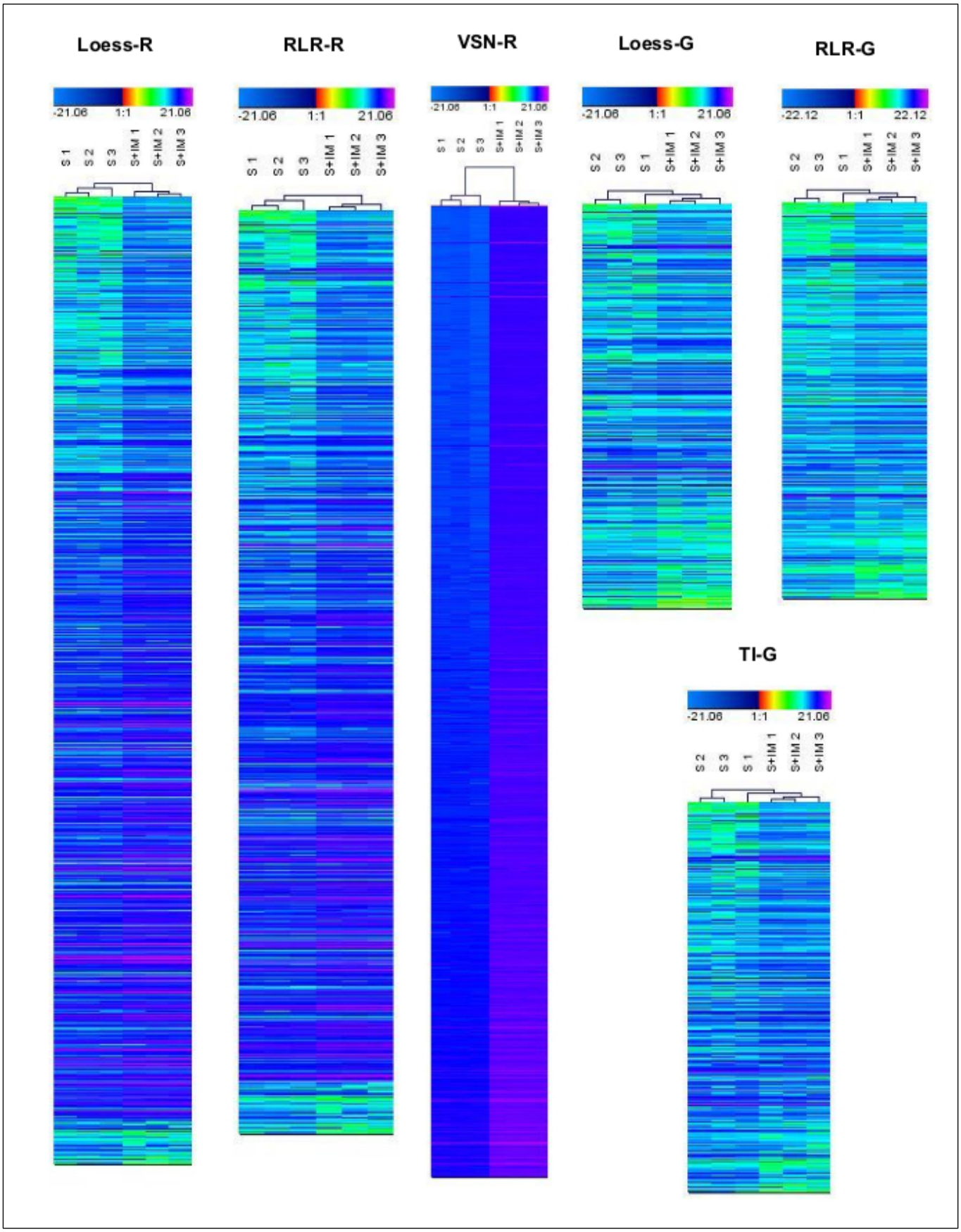


Figure A5 – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset B based on differentiators identified using fold change.

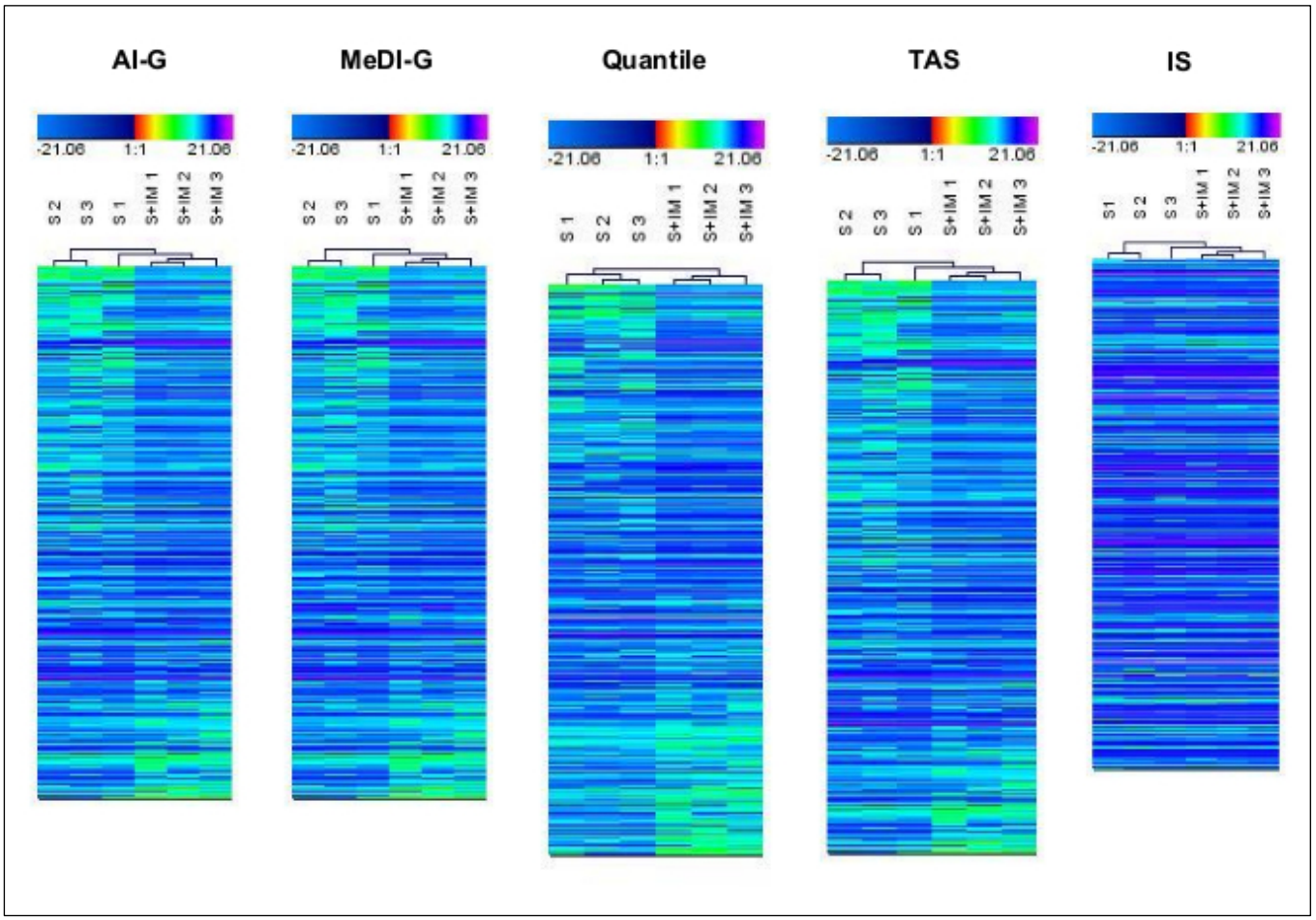


Figure A5 (contd..) – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset B based on differentiators identified using fold change.

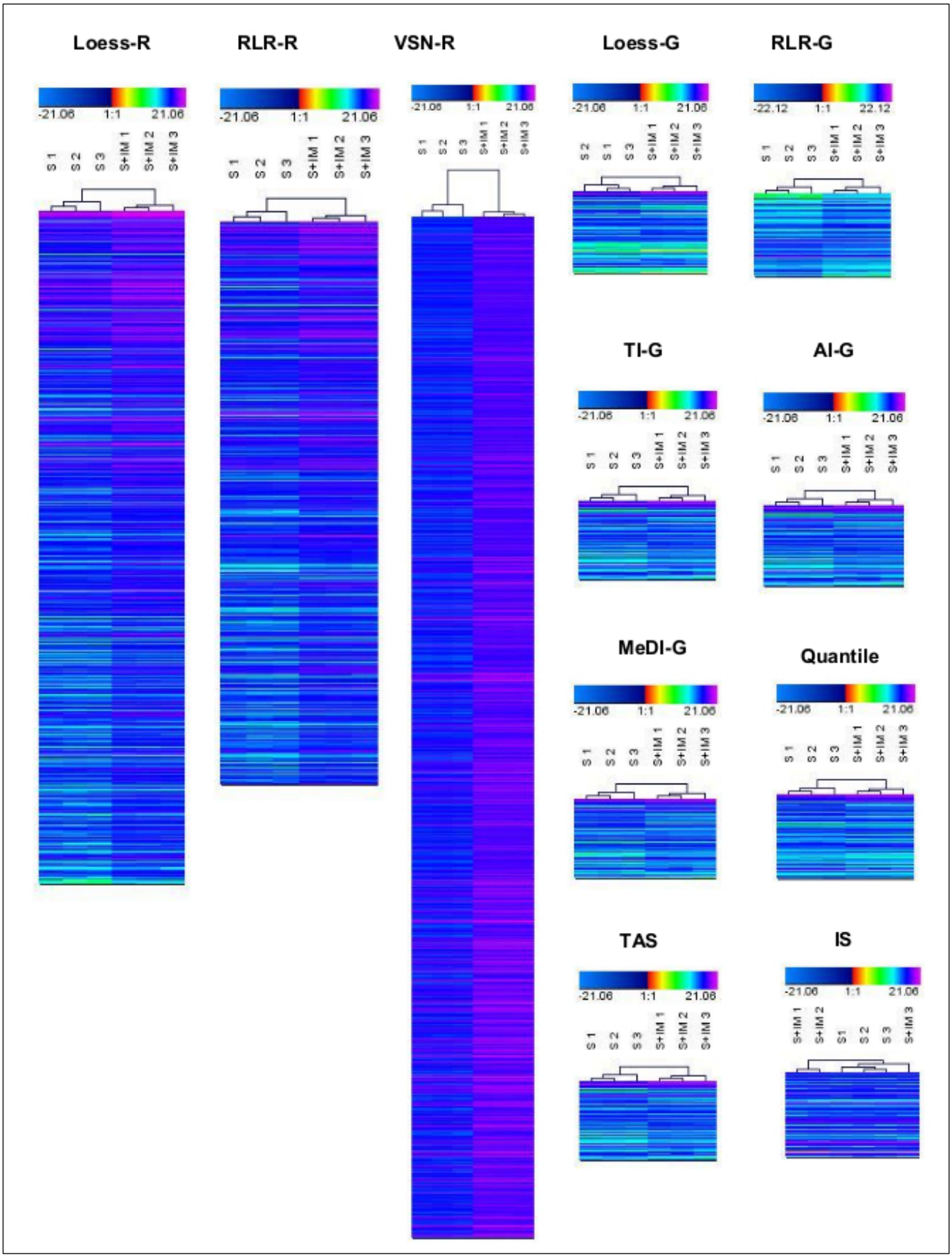


Figure A6 – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset B based on differentiators identified using both p-value and fold change.

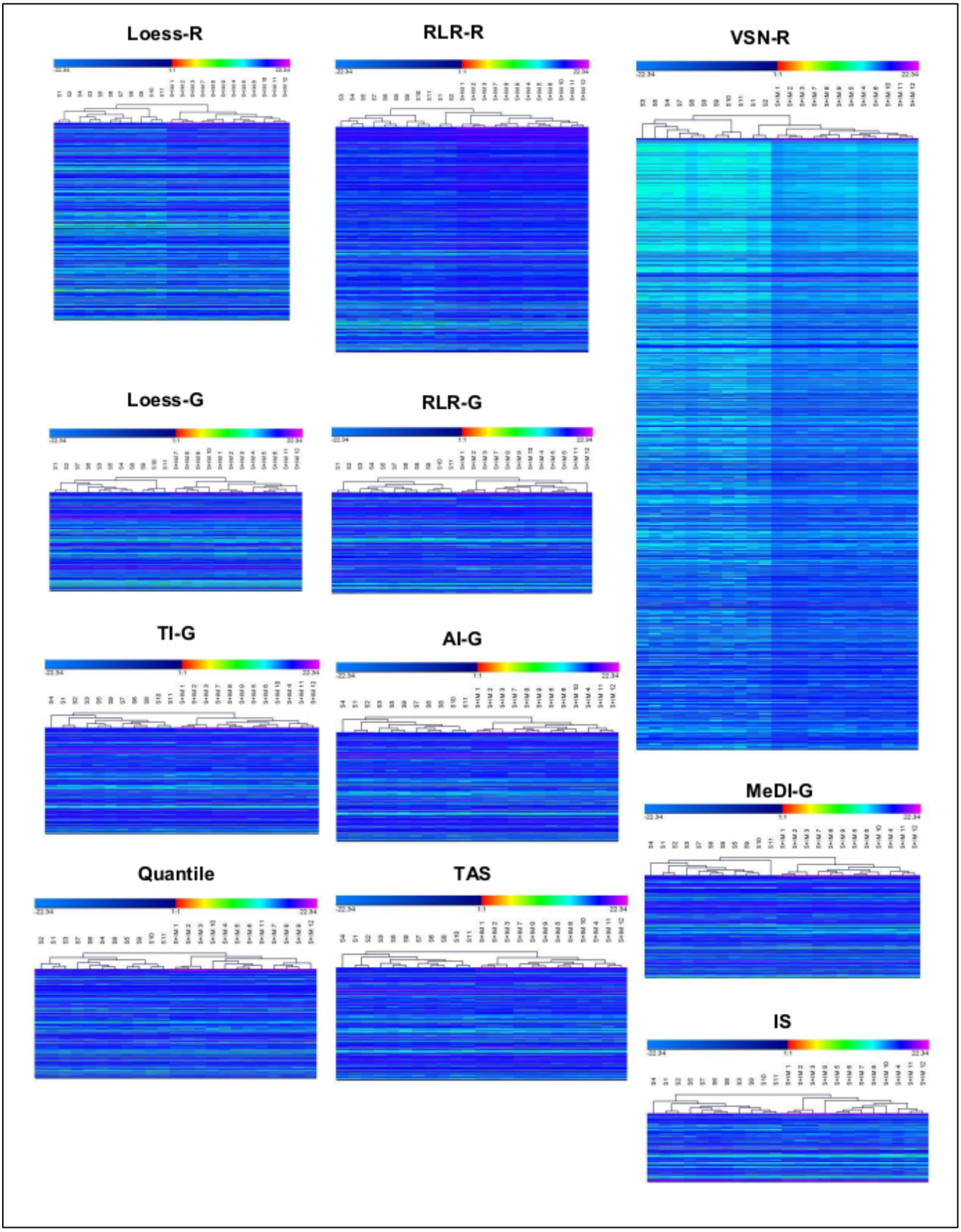


Figure A7 – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset C based on differentiators identified using p-value.

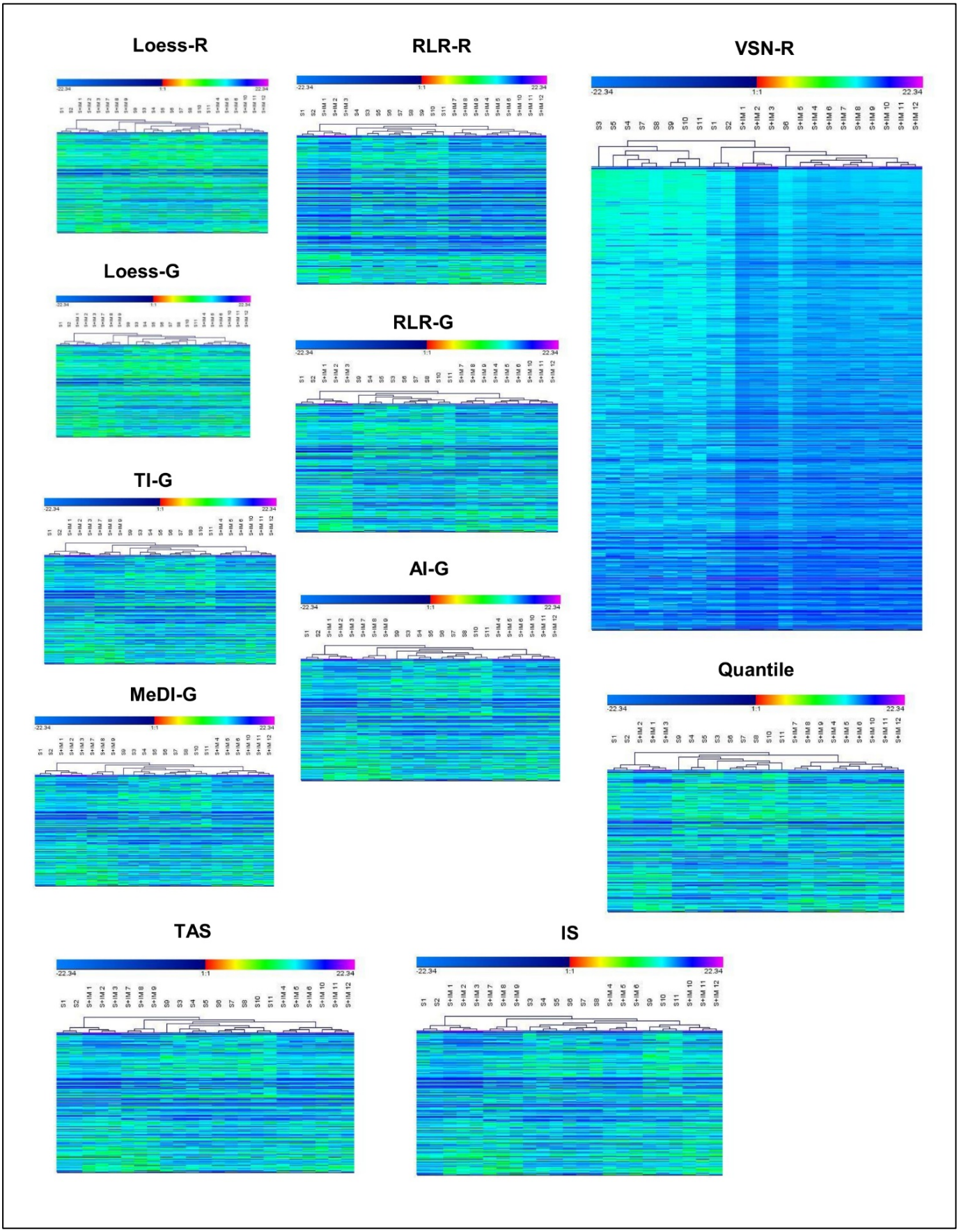


Figure A8 – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset C based on differentiators identified using fold change.

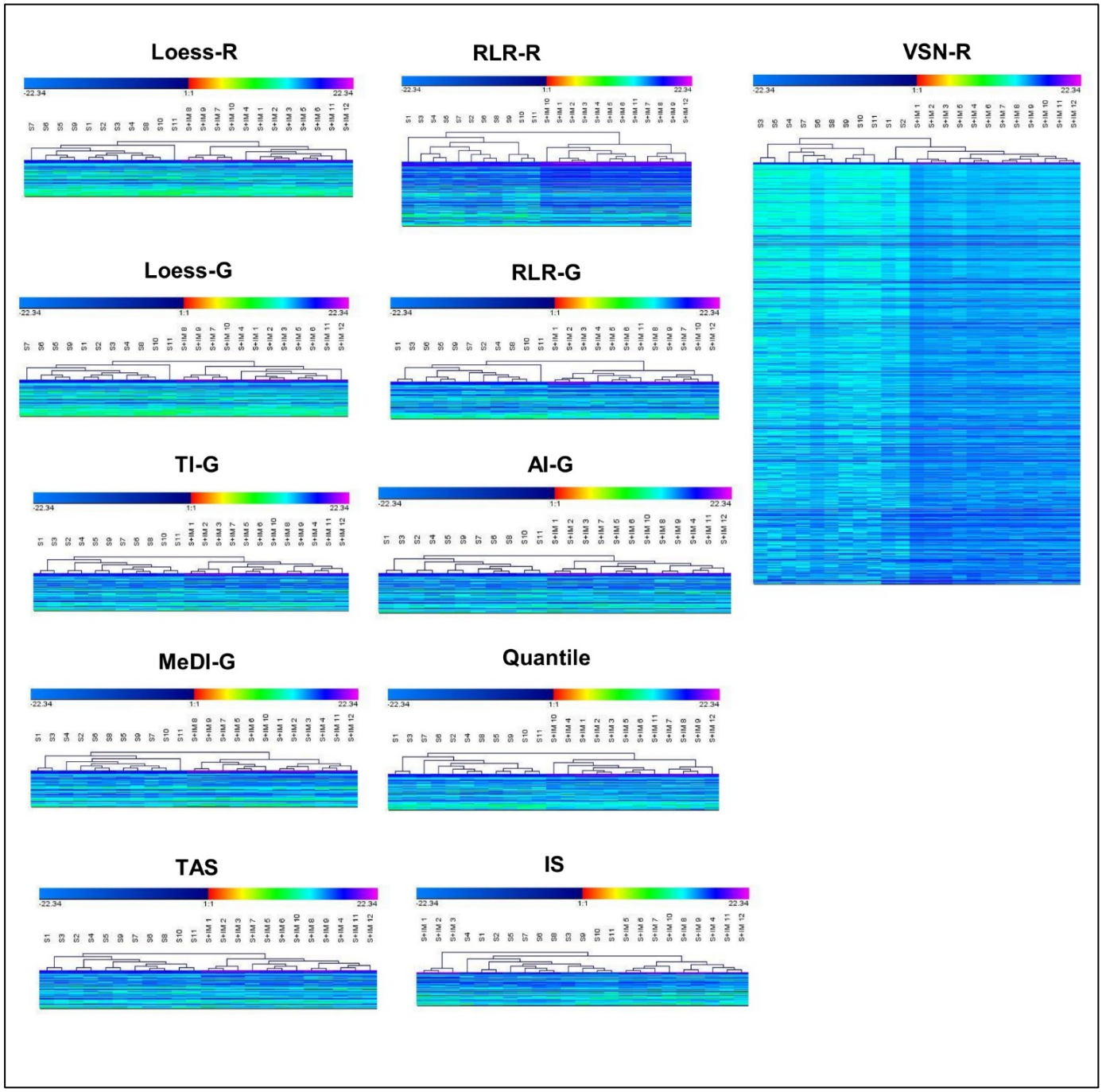


Figure A9 – Dendrograms representing hierarchical clustering of control (S) and test (S+IM) groups of dataset C based on differentiators identified using both p-value and fold change.

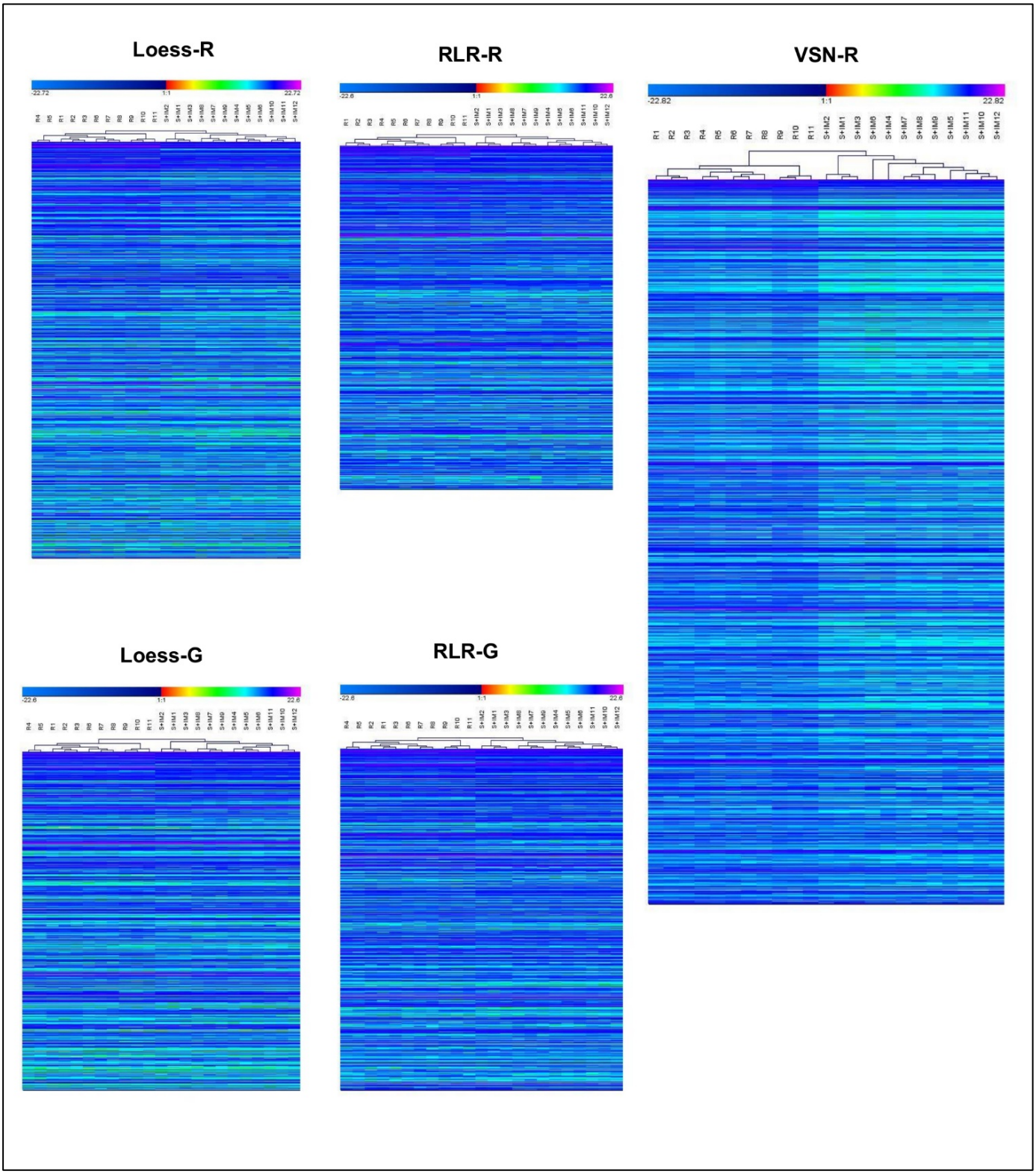


Figure A10 – Dendrograms representing hierarchical clustering of control (S+IM) and test (R) groups of dataset D based on differentiators identified using p-value.

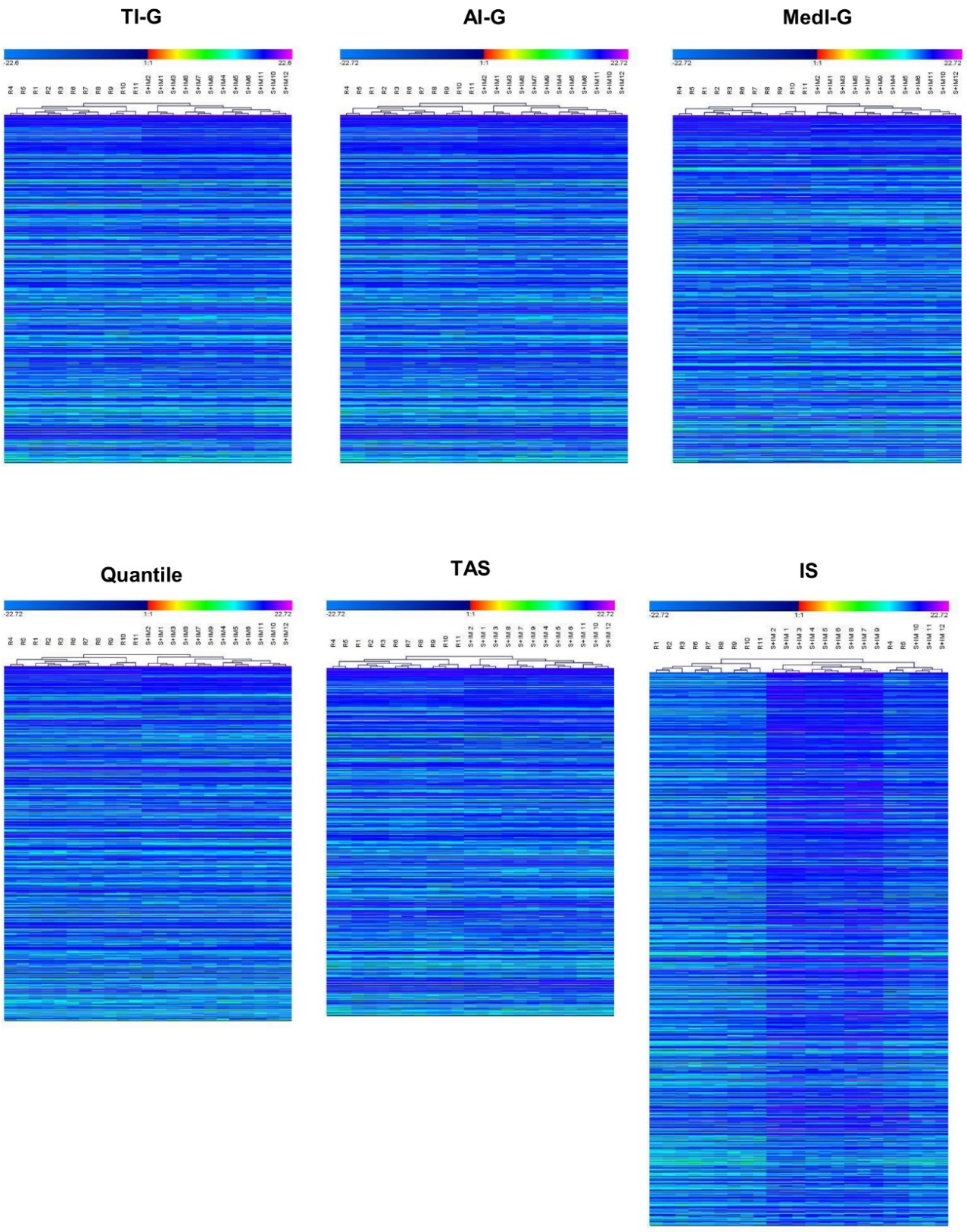


Figure A10 (contd…) – Dendrograms representing hierarchical clustering of control (S+IM) and test (R) groups of dataset D based on differentiators identified using p-value.

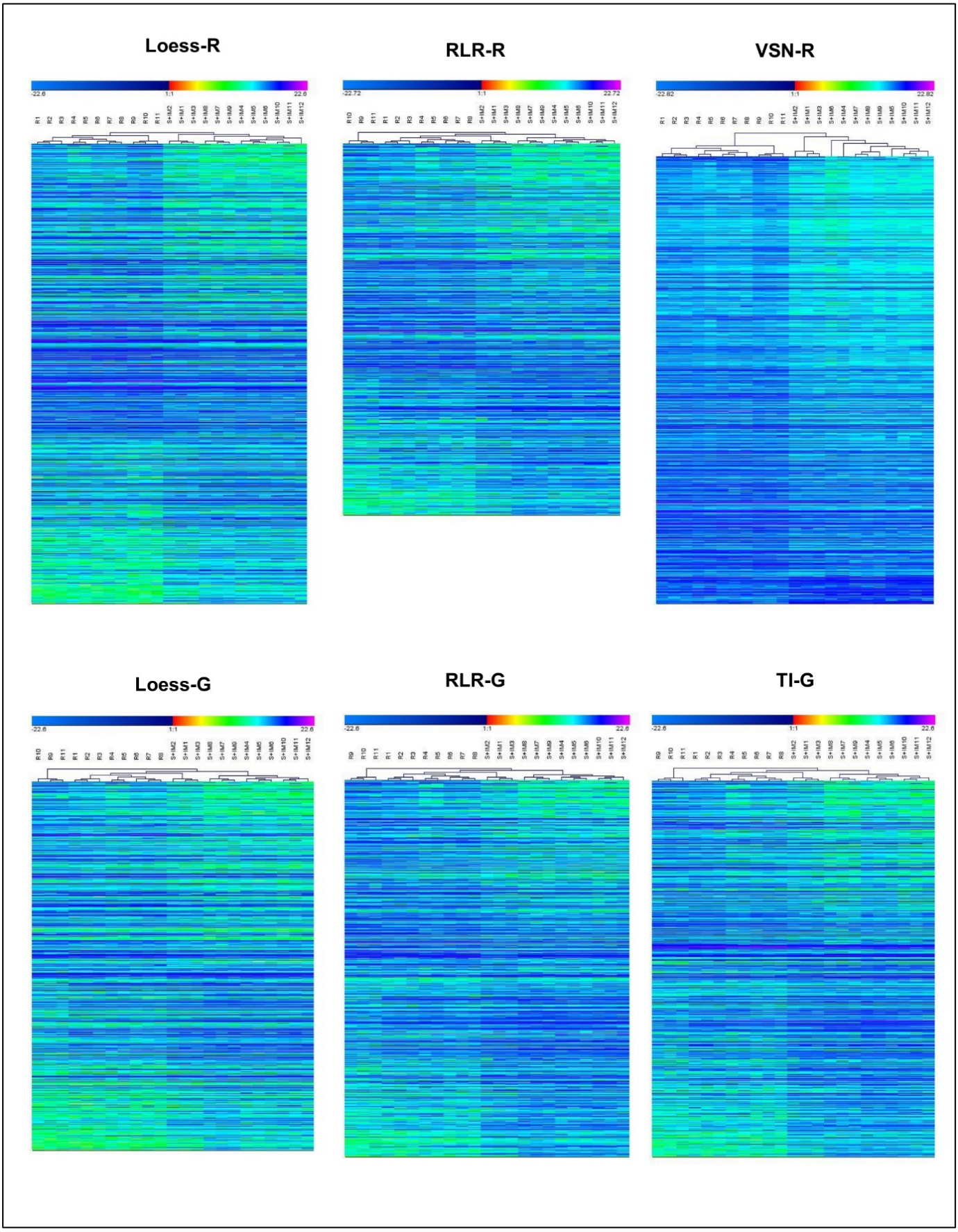


Figure A11 – Dendrograms representing hierarchical clustering of control (S+IM) and test (R) groups of dataset D based on differentiators identified using fold change.

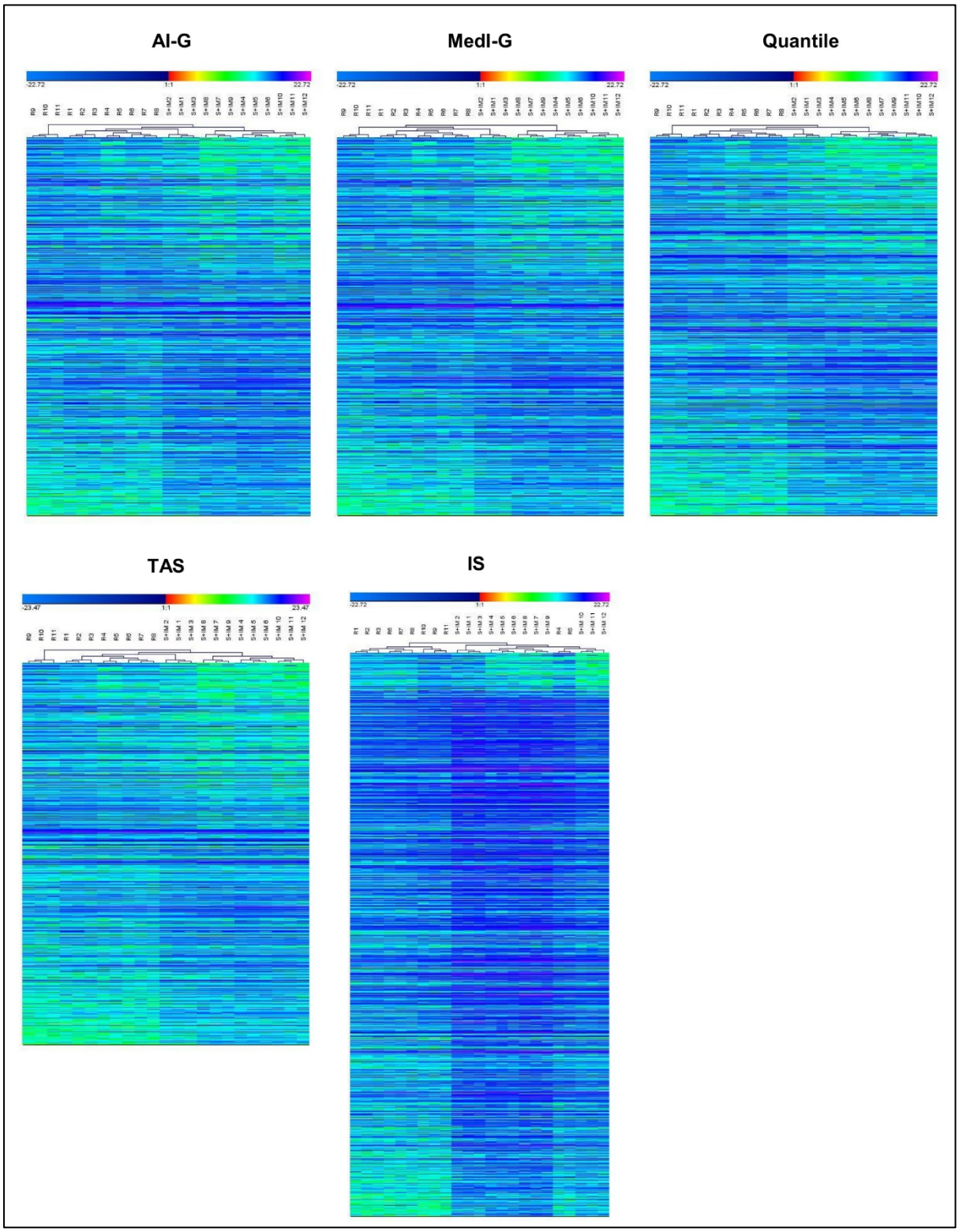


Figure A11 (contd..) – Dendrograms representing hierarchical clustering of control (S+IM) and test (R) groups of dataset D based on differentiators identified using fold change.

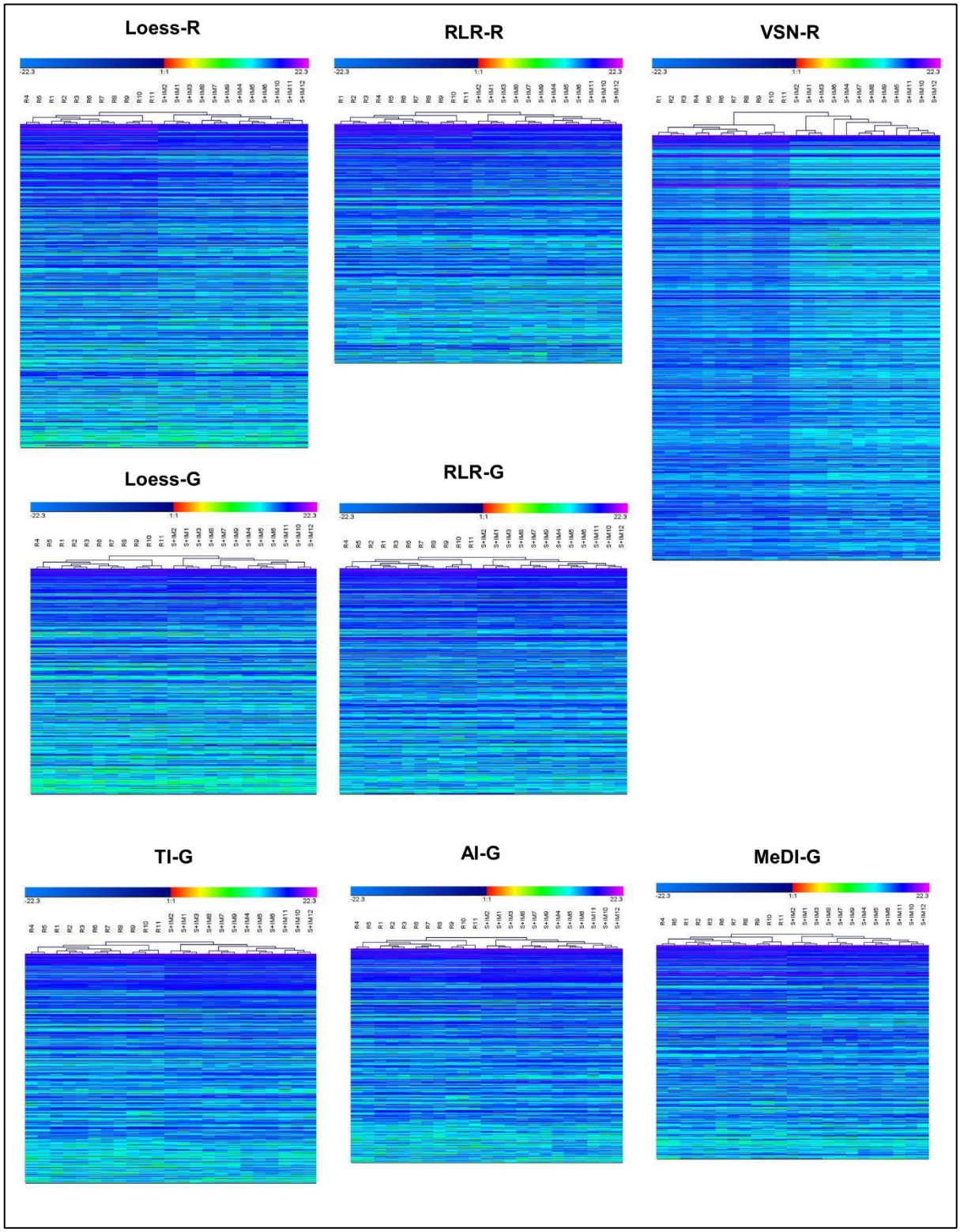


Figure A12 – Dendrograms representing hierarchical clustering of control (S+IM) and test (R) groups of dataset D based on differentiators identified using both p-value & fold change.

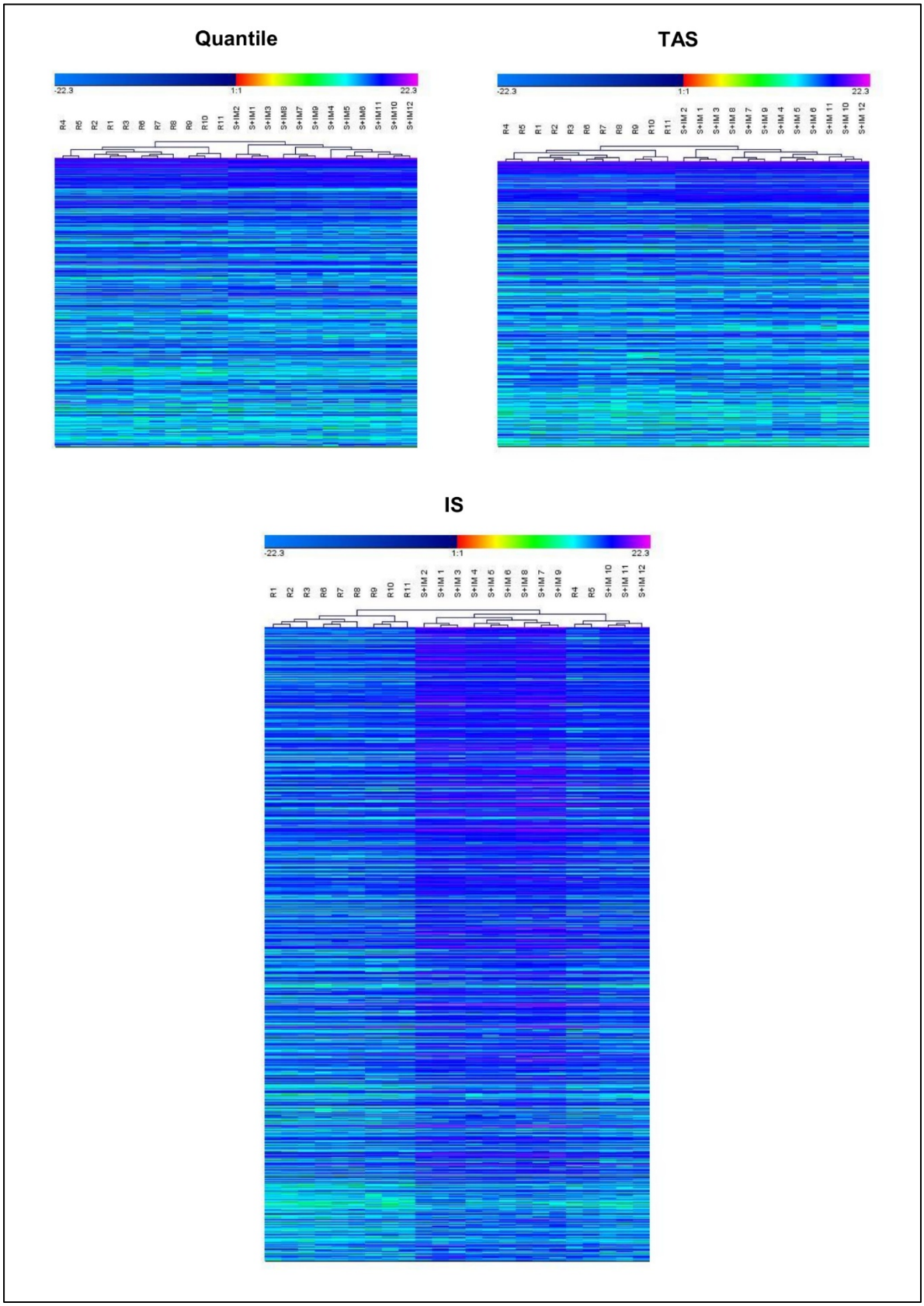


Figure A12 (contd..)– Dendrograms representing hierarchical clustering of control (S+IM) and test (R) groups of dataset D based on differentiators identified using both p-value & fold change.

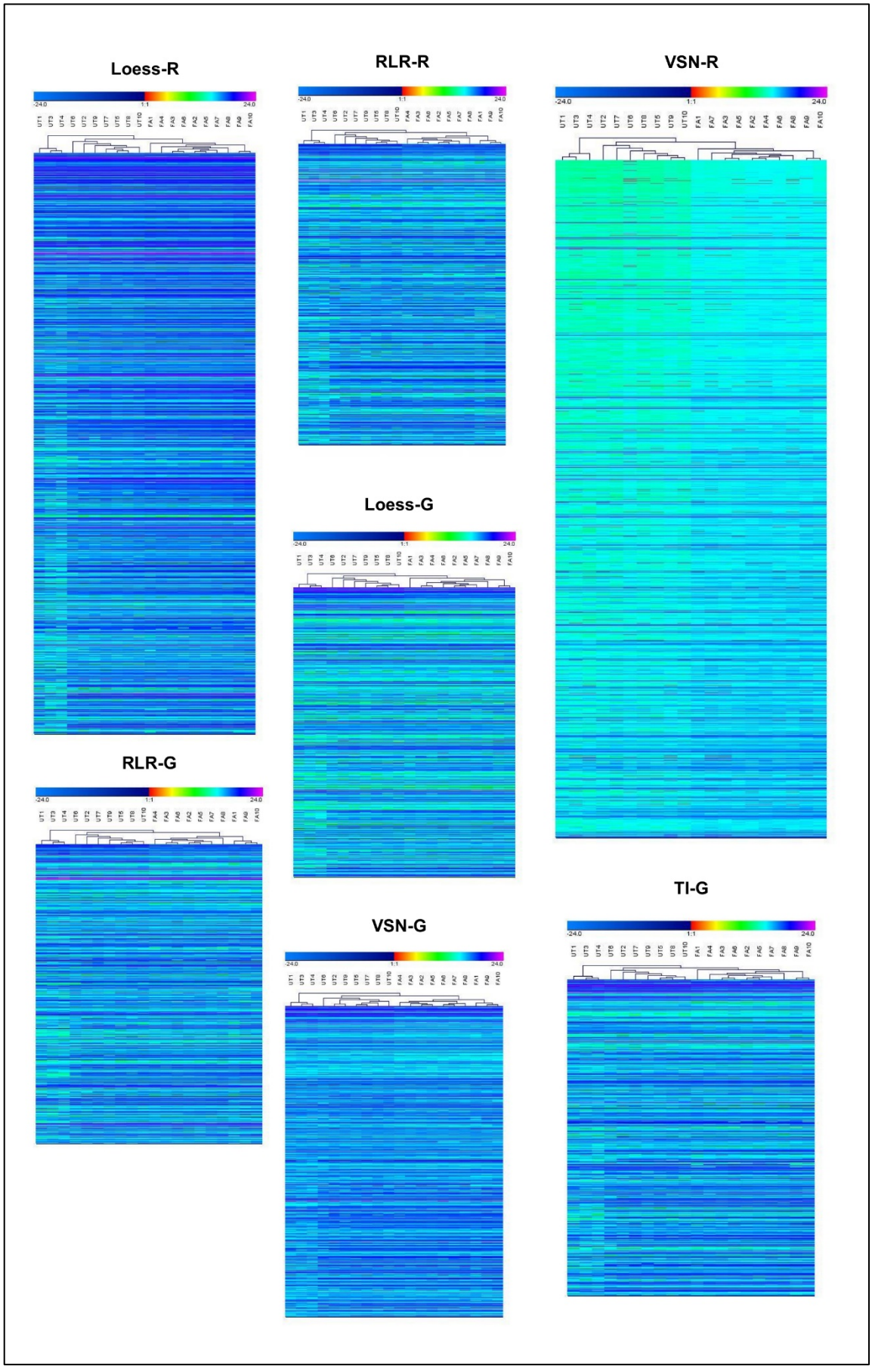


Figure A13 – Dendrograms representing hierarchical clustering of control (UT) and test (FA) groups of dataset E based on differentiators identified using p-value.

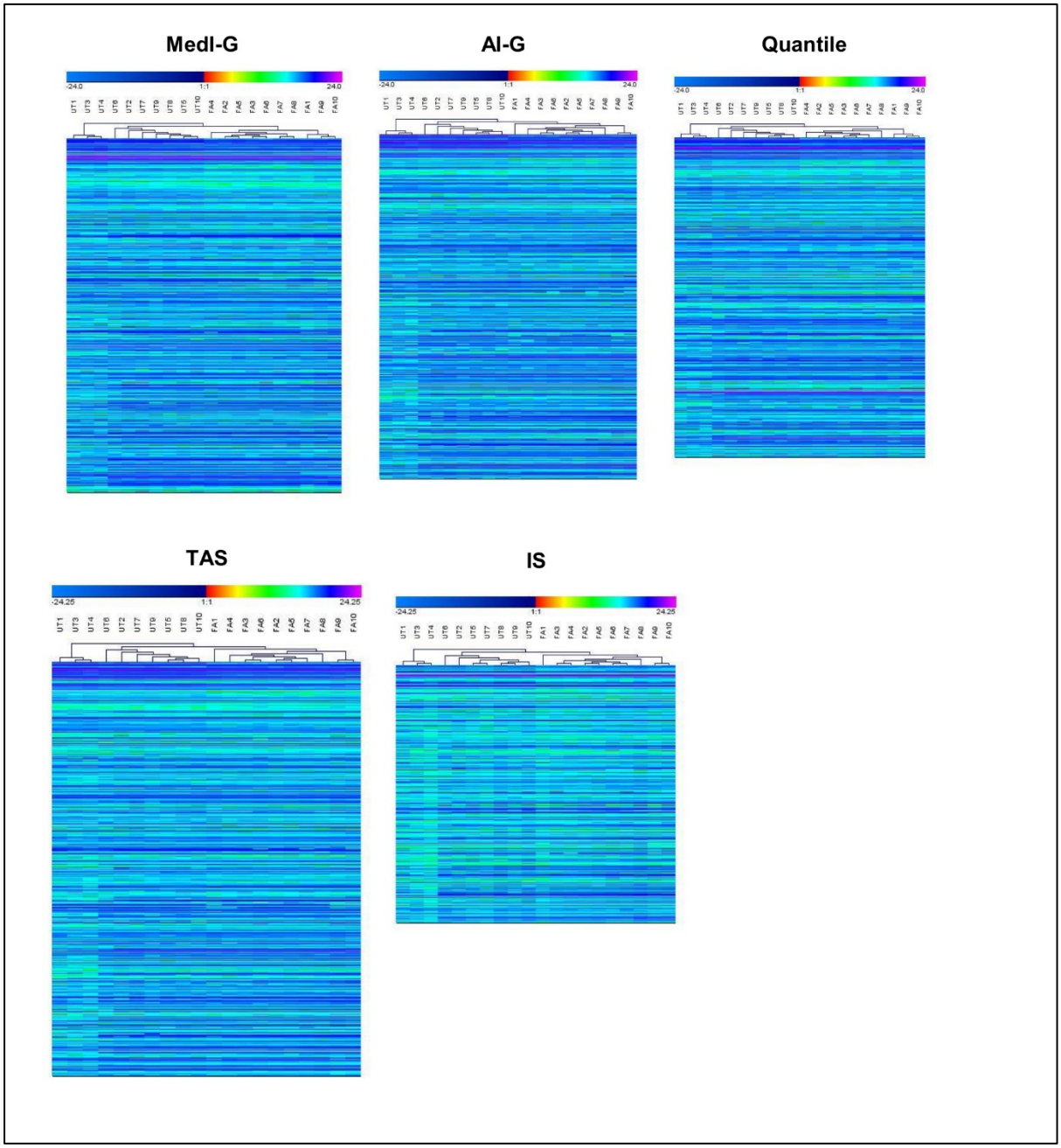


Figure A13(contd..) – Dendrograms representing hierarchical clustering of control (UT) and test (FA) groups of dataset E based on differentiators identified using p-value.

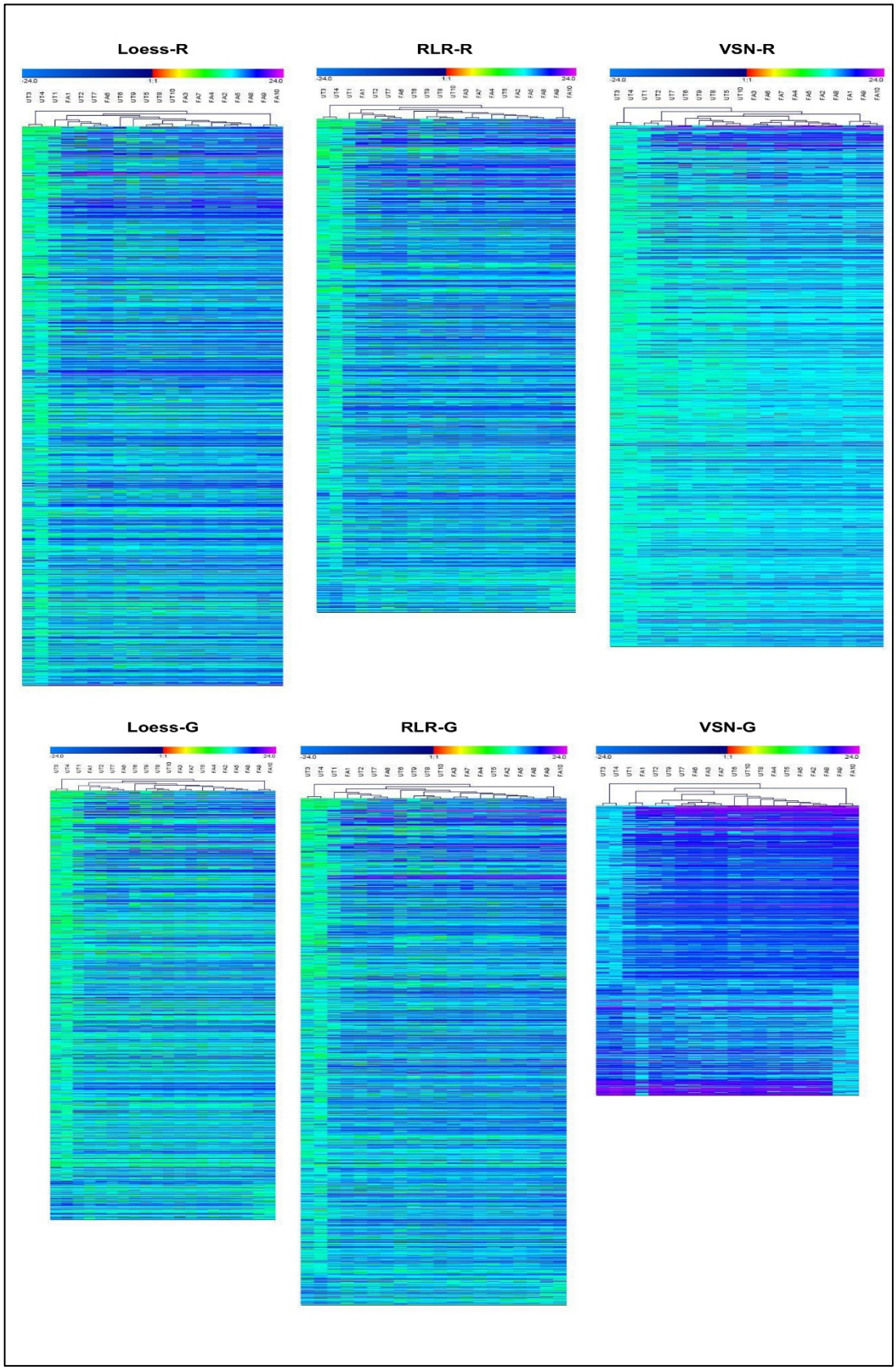


Figure A14 – Dendrograms representing hierarchical clustering of control (UT) and test (FA) groups of dataset E based on differentiators identified using fold change.

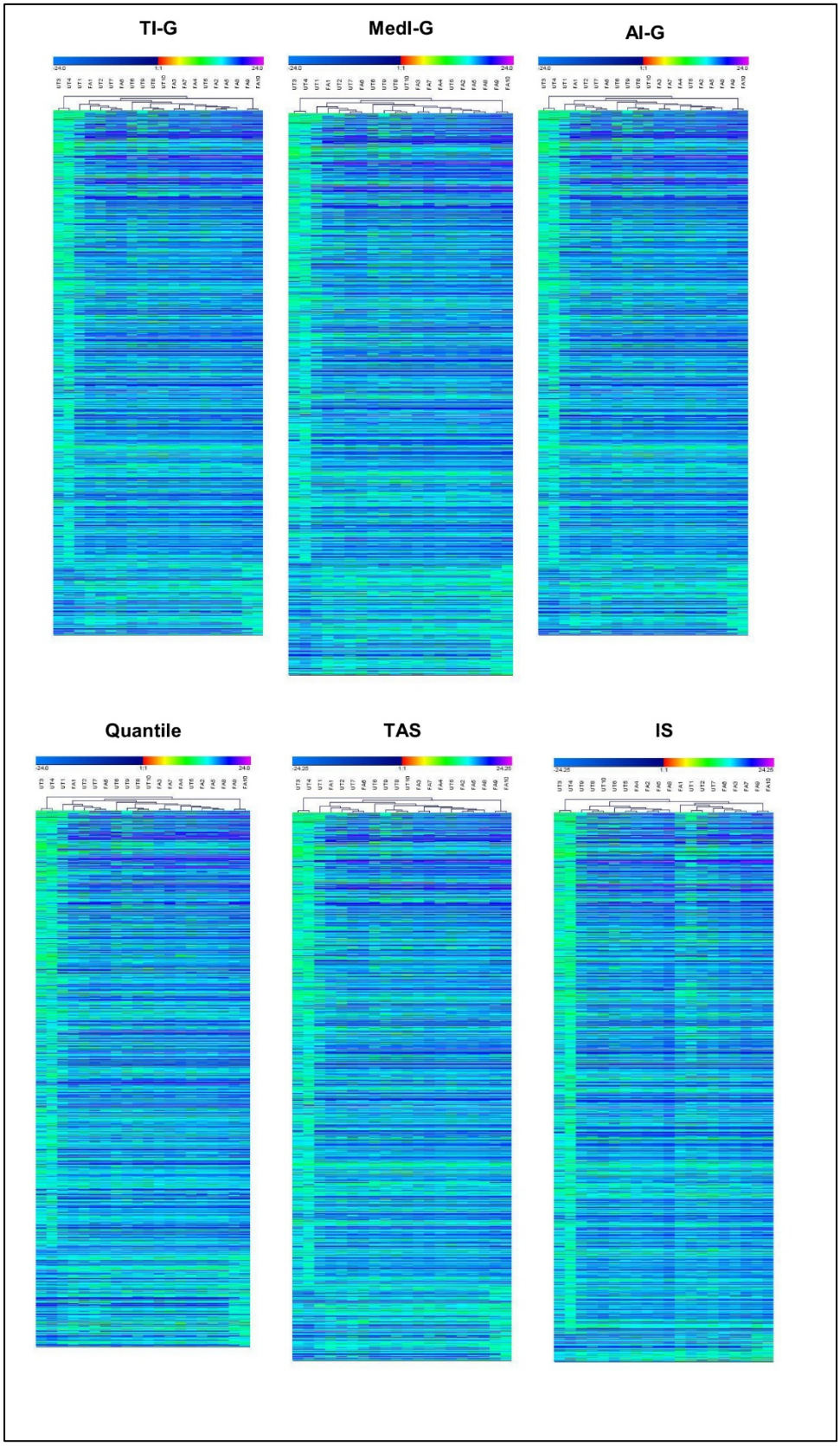


Figure A14 (contd.) – Dendrograms representing hierarchical clustering of control (UT) and test (FA) groups of dataset E based on differentiators identified using fold change.

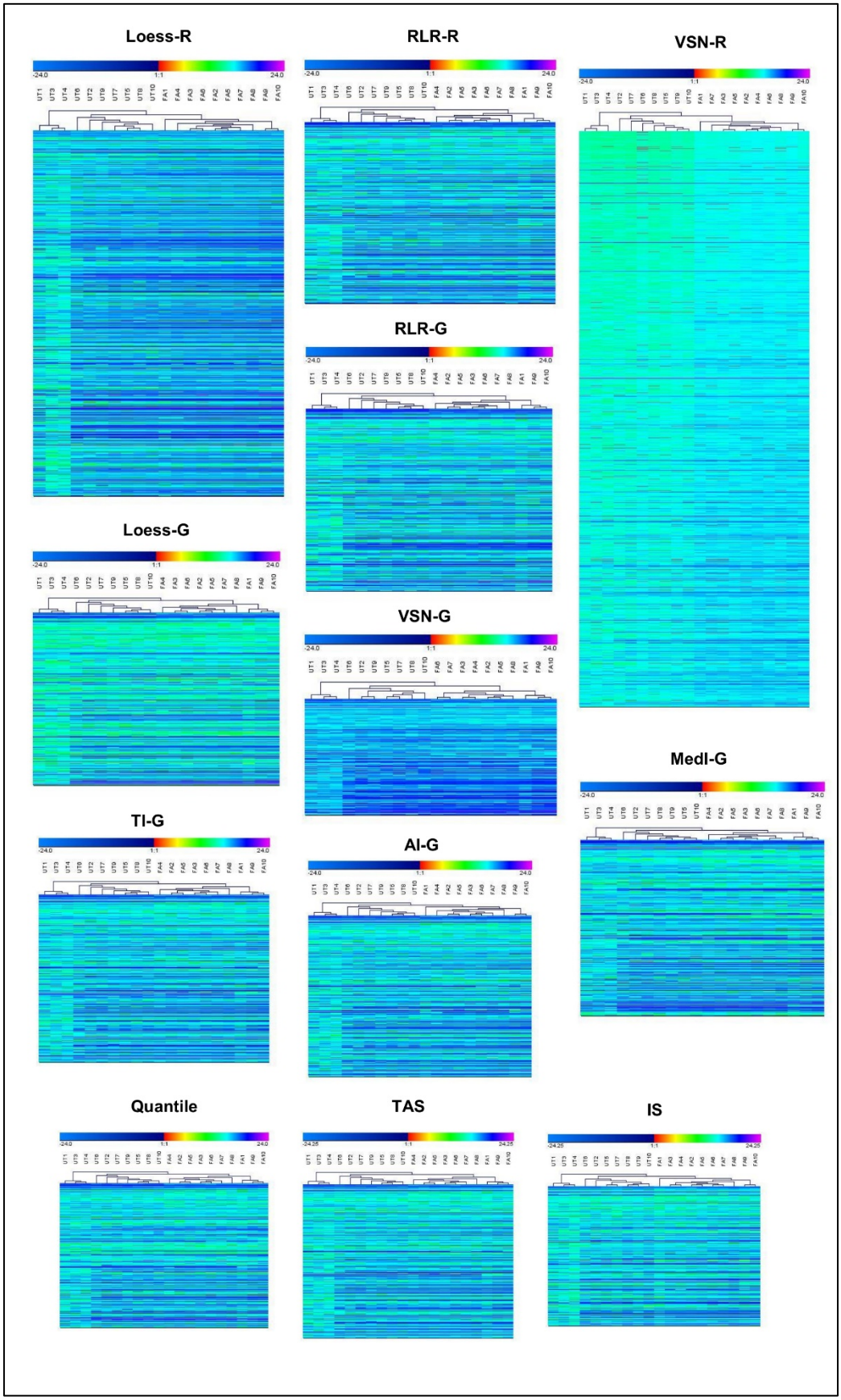


Figure A15 – Dendrograms representing hierarchical clustering of control (UT) and test (FA) groups of dataset E based on differentiators identified using both p-value & fold change.

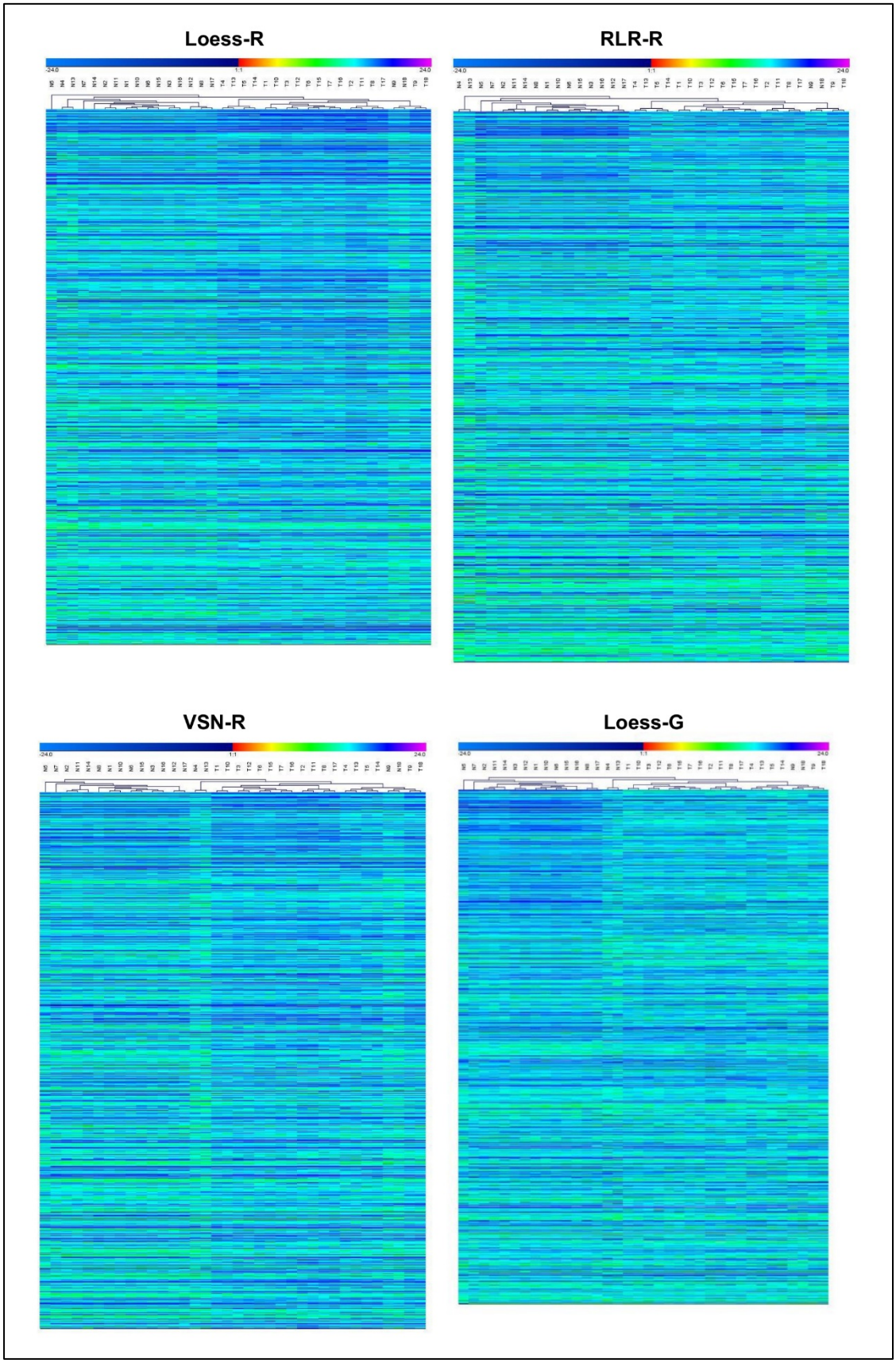


Figure A16 – Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using p-value.

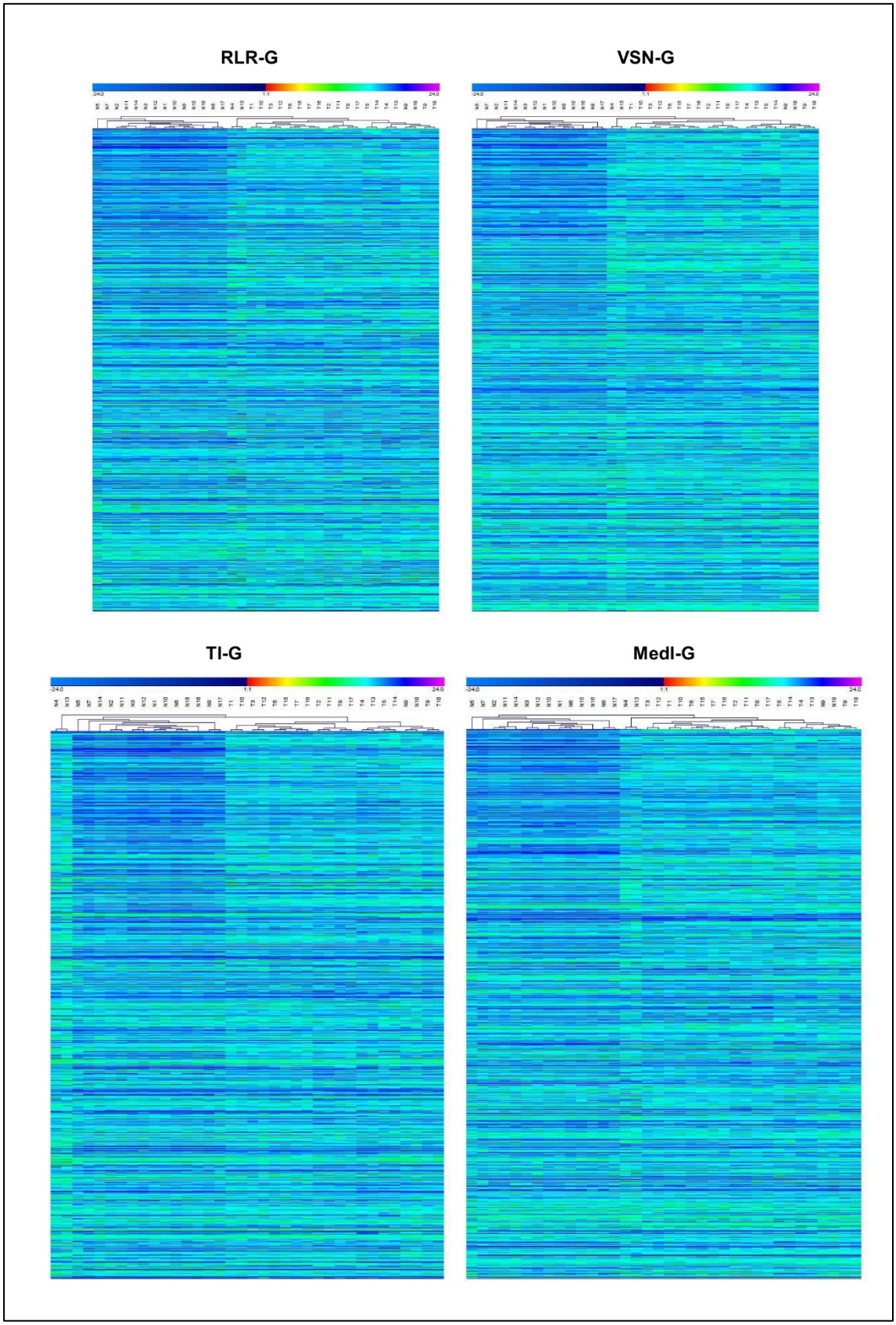


Figure A16 (contd.) – Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using p-value.

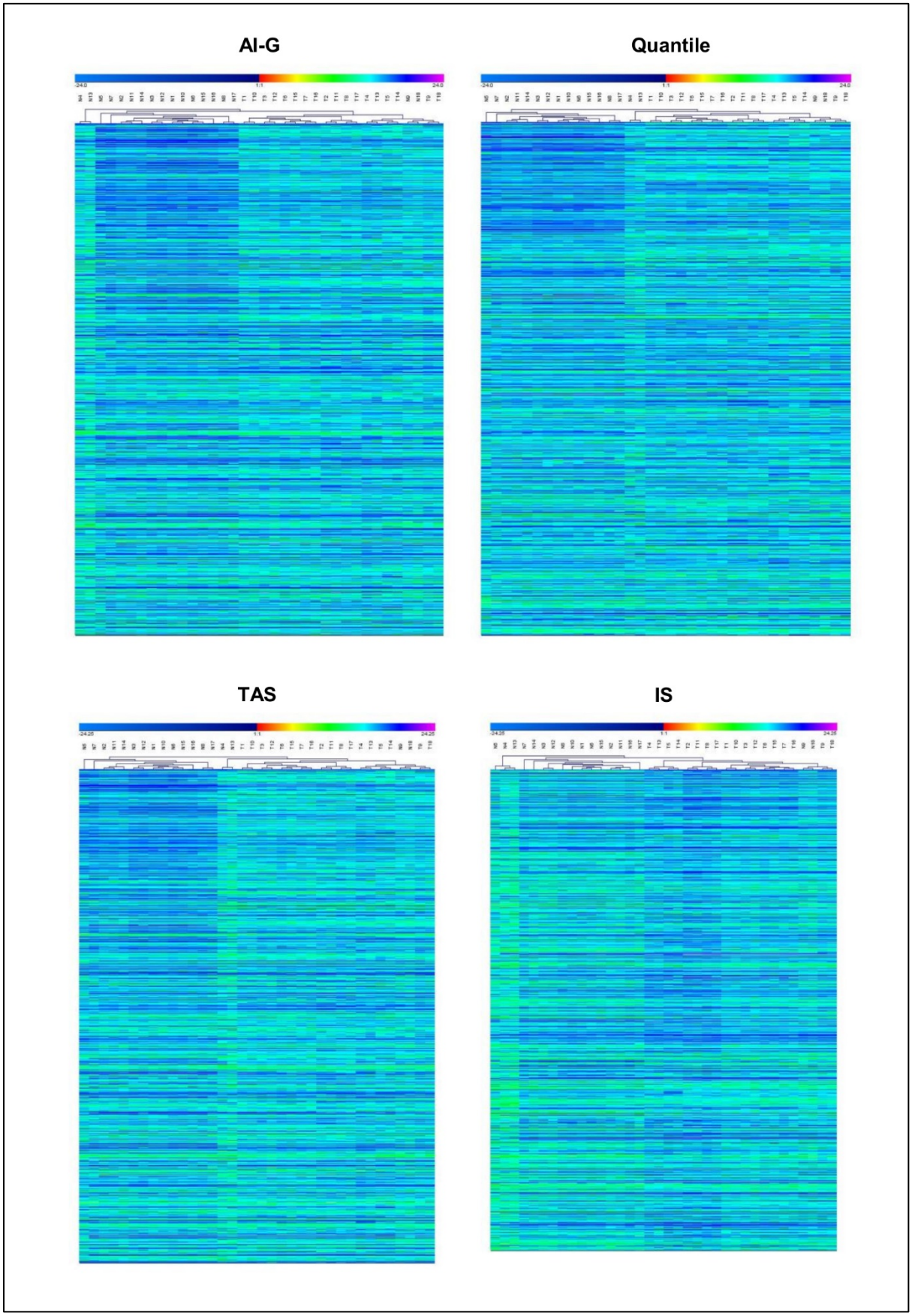


Figure A16 (contd.)– Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using p-value.

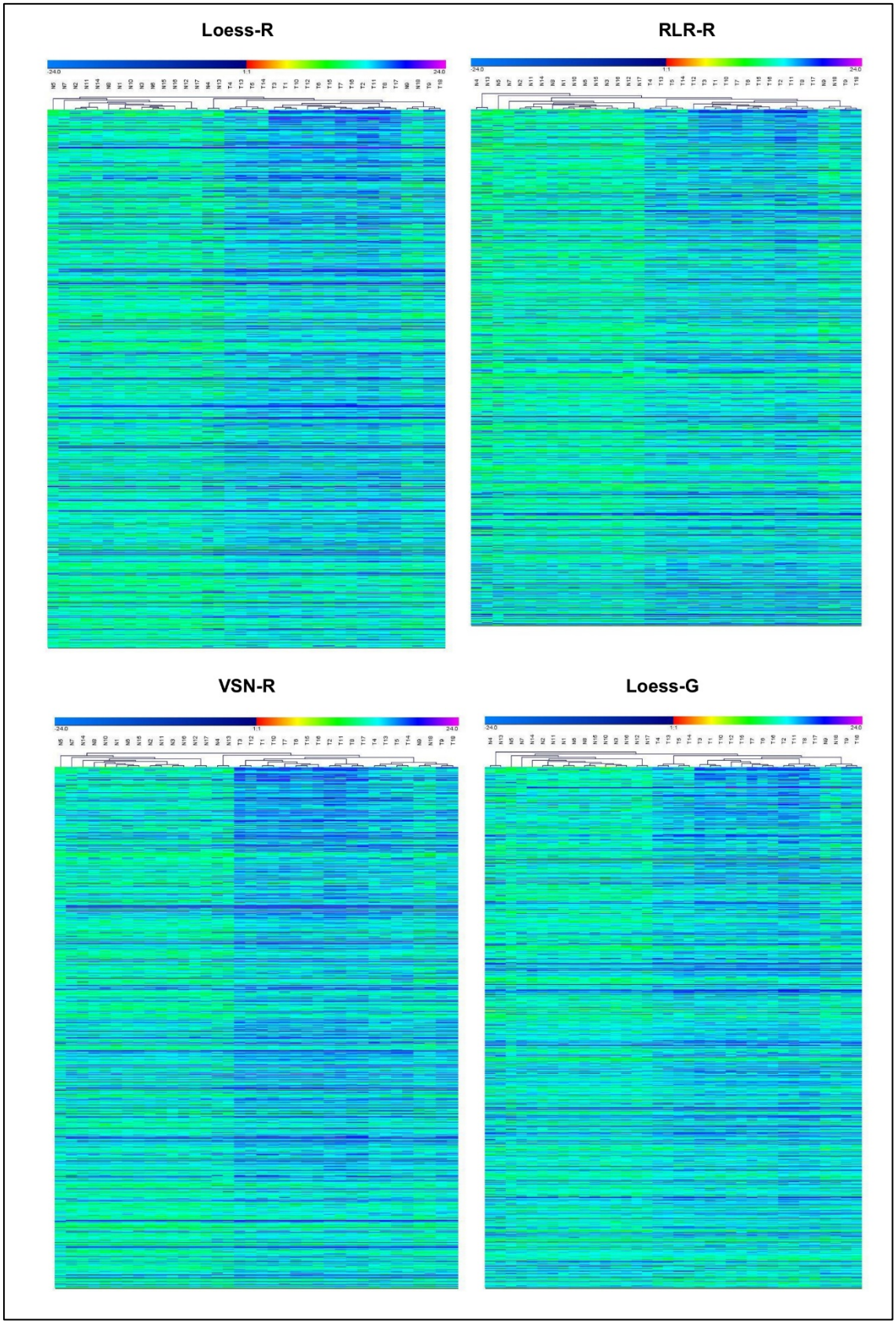


Figure A17 – Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using fold change.

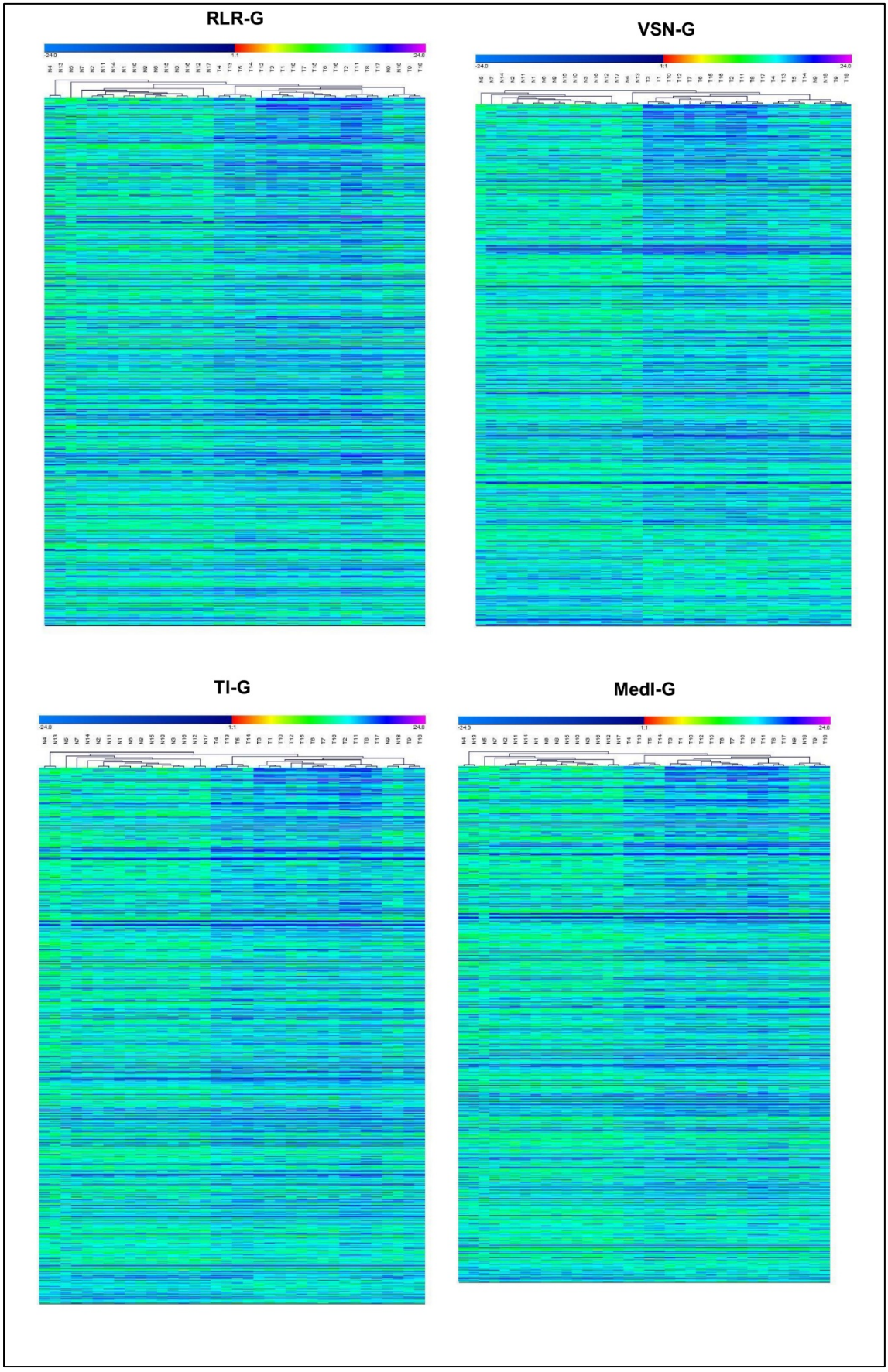


Figure A17 (contd.)– Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using fold change.

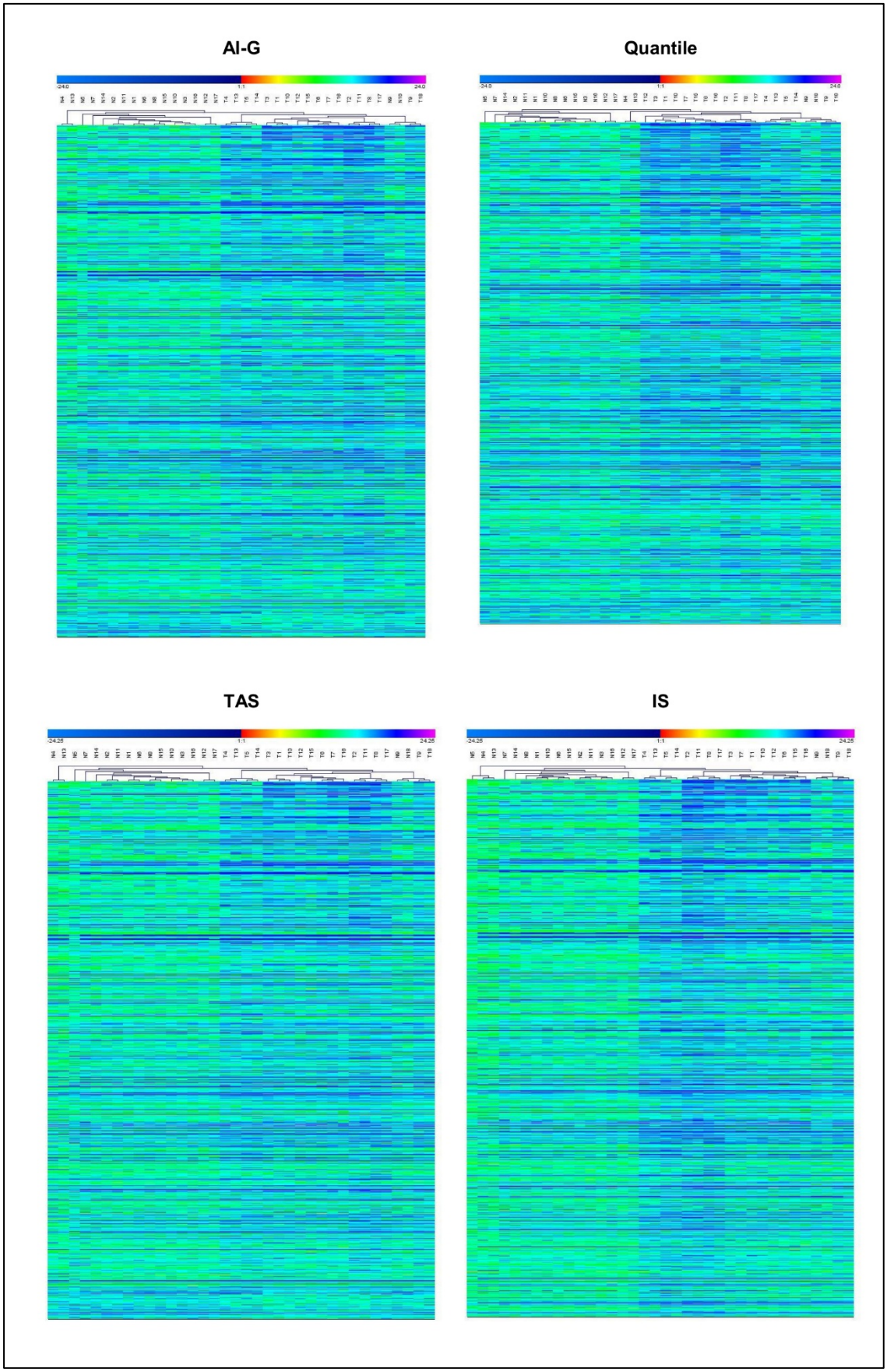


Figure A17 (contd.)– Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using fold change.

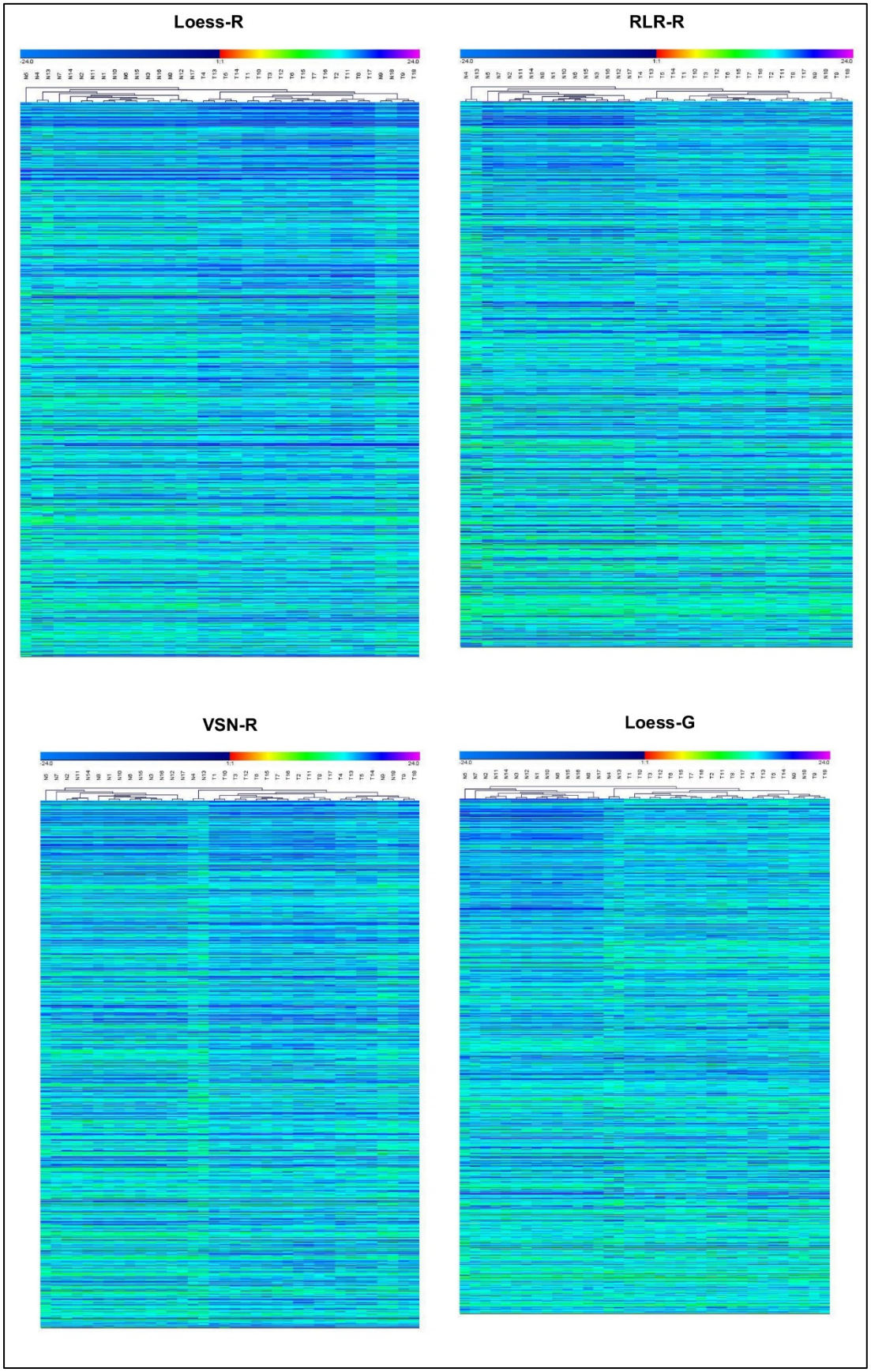


Figure A18 – Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using both p-value and fold change.

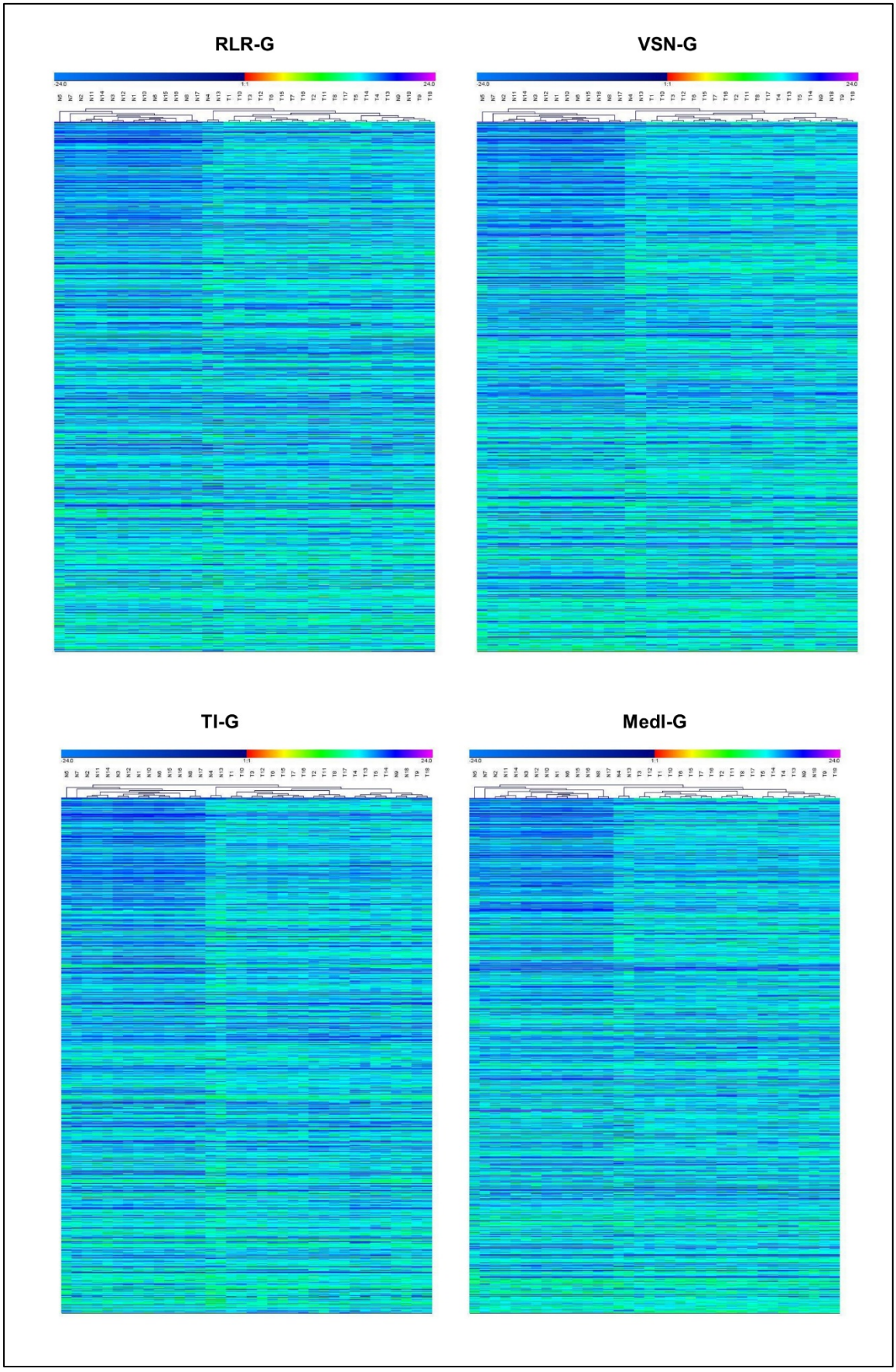


Figure A18 (contd.) – Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using both p-value and fold change.

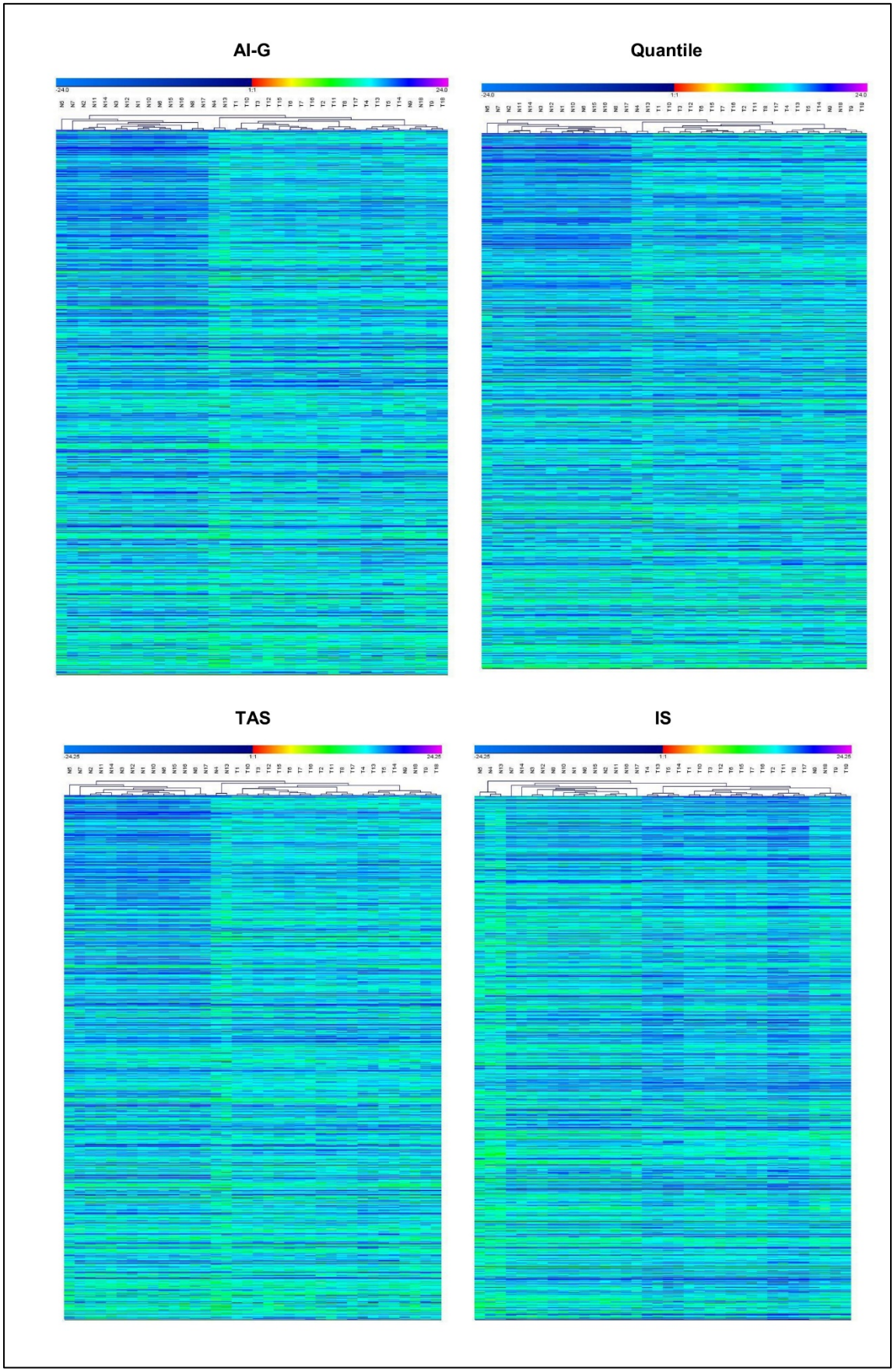


Figure A18 (contd.)– Dendrograms representing hierarchical clustering of control (N) and test (T) groups of dataset F based on differentiators identified using both p-value and fold change.

10. REFERENCES

1. Pasternak G, Hochhaus A, Schultheis B, Hehlmann R: **Chronic myelogenous leukemia: molecular and cellular aspects**. *Journal of cancer research and clinical oncology* 1998, **124**(12):643-660.
2. Ren R: **Mechanisms of BCR–ABL in the pathogenesis of chronic myelogenous leukaemia**. *Nature Reviews Cancer* 2005, **5**(3):172-183.
3. Kurzrock R, Kantarjian HM, Druker BJ, Talpaz M: **Philadelphia chromosome-positive leukemias: from basic mechanisms to molecular therapeutics**. *Annals of internal medicine* 2003, **138**(10):819-830.
4. Eaves C, Cashman J, Eaves A: **Defective regulation of leukemic hematopoiesis in chronic myeloid leukemia**. *Leukemia research* 1998, **22**(12):1085-1096.
5. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, Bloomfield CD, Cazzola M, Vardiman JW: **The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia**. *Blood* 2016, **127**(20):2391-2405.
6. Kurzrock R, Gutterman JU, Talpaz M: **The molecular genetics of Philadelphia chromosome-positive leukemias**. *The New England journal of medicine* 1988, **319**(15):990-998.
7. Ross TS, Mgbemena VE: **Re-evaluating the role of BCR/ABL in chronic myelogenous leukemia**. *Molecular & cellular oncology* 2014, **1**(3):e963450.
8. Pane F, Intrieri M, Quintarelli C, Izzo B, Muccioli GC, Salvatore F: **BCR/ABL genes and leukemic phenotype: from molecular mechanisms to clinical correlations**. *Oncogene* 2002, **21**(56):8652-8667.
9. Thompson PA, Kantarjian HM, Cortes JE: **Diagnosis and Treatment of Chronic Myeloid Leukemia in 2015**. *Mayo Clinic proceedings* 2015, **90**(10):1440-1454.
10. Calabretta B, Perrotti D: **The biology of CML blast crisis**. *Blood* 2004, **103**(11):4010-4022.
11. An X, Tiwari AK, Sun Y, Ding PR, Ashby CR, Jr., Chen ZS: **BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: a review**. *Leukemia research* 2010, **34**(10):1255-1268.
12. **BCR-ABL fusion gene** [https://www.cancer.gov/publications/dictionaries/cancer-terms/def/bcr-abl-fusion-gene]
13. Jabbour E, Kantarjian H: **Chronic myeloid leukemia: 2018 update on diagnosis, therapy and monitoring**. 2018, **93**(3):442-459.
14. Goldman JM, Melo JV: **Chronic myeloid leukemia--advances in biology and new approaches to treatment**. *The New England journal of medicine* 2003, **349**(15):1451-1464.
15. Soverini S, Mancini M, Bavaro L, Cavo M, Martinelli G: **Chronic myeloid leukemia: the paradigm of targeting oncogenic tyrosine kinase signaling and counteracting resistance for successful cancer therapy**. *Mol Cancer* 2018, **17**(1):49.
16. Ferrajoli A, Fizzotti M, Liberati AM, Grignani F: **Chronic myelogenous leukemia: an update on the biological findings and therapeutic approaches**. *Critical reviews in oncology/hematology* 1996, **22**(3):151-174.
17. Hantschel O: **Structure, regulation, signaling, and targeting of abl kinases in cancer**. *Genes Cancer* 2012, **3**(5-6):436-446.
18. Deininger MW, Goldman JM, Melo JV: **The molecular biology of chronic myeloid leukemia**. *Blood* 2000, **96**(10):3343-3356.

19. Pérez-Caro M, Sánchez-Garcia I: **BCR-ABL and Human Cancer**. In: *Apoptosis, Cell Signaling, and Human Diseases: Molecular Mechanisms.* Edited by Srivastava R. Totowa, NJ: Humana Press; 2007: 3-34.
20. Rumpold H, Webersinke G: **Molecular pathogenesis of Philadelphia-positive chronic myeloid leukemia - is it all BCR-ABL?** *Curr Cancer Drug Targets* 2011, **11**(1):3-19.
21. Laurent E, Talpaz M, Kantarjian H, Kurzrock R: **The BCR gene and philadelphia chromosome-positive leukemogenesis**. *Cancer research* 2001, **61**(6):2343-2355.
22. Warmuth M, Danhauser-Riedl S, Hallek M: **Molecular pathogenesis of chronic myeloid leukemia: implications for new therapeutic strategies**. *Annals of hematology* 1999, **78**(2):49-64.
23. Melo JV: **BCR-ABL gene variants**. *Bailliere's clinical haematology* 1997, **10**(2):203-222.
24. Butturini A, Arlinghaus RB, Gale RP: **BCR/ABL and leukemia**. *Leukemia research* 1996, **20**(6):523-529.
25. Balabanov S, Evans CA, Abraham SA, Pellicano F, Copland M, Walker MJ, Whetton AD, Holyoake TL: **Quantitative proteomics analysis of BMS-214662 effects on CD34 positive cells from chronic myeloid leukaemia patients**. *Proteomics* 2013, **13**(1):153-168.
26. O'Hare T, Deininger MWN, Eide CA, Clackson T, Druker BJ: **Targeting the BCR-ABL Signaling Pathway in Therapy-Resistant Philadelphia Chromosome-Positive Leukemia**. *Clinical Cancer Research* 2011, **17**(2):212-221.
27. Abram CL, Lowell CA: **The ins and outs of leukocyte integrin signaling**. *Annual review of immunology* 2009, **27**:339-362.
28. Patel AB, O'Hare T, Deininger MW: **Mechanisms of Resistance to ABL Kinase Inhibition in Chronic Myeloid Leukemia and the Development of Next Generation ABL Kinase Inhibitors**. *Hematology/oncology clinics of North America* 2017, **31**(4):589-612.
29. Kujak C, Kolesar JM: **Treatment of chronic myelogenous leukemia**. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists* 2016, **73**(3):113-120.
30. Santos FPS, Kantarjian H, Quintás-Cardama A, Cortes J: **Evolution of therapies for chronic myelogenous leukemia**. *Cancer J* 2011, **17**(6):465-476.
31. Wong S, Witte ON: **The BCR-ABL story: bench to bedside and back**. *Annu Rev Immunol* 2004, **22**:247-306.
32. Bhatia R, McCarthy JB, Verfaillie CM: **Interferon-alpha restores normal beta 1 integrin-mediated inhibition of hematopoietic progenitor proliferation by the marrow microenvironment in chronic myelogenous leukemia**. *Blood* 1996, **87**(9):3883-3891.
33. Roychowdhury S, Talpaz M: **Managing resistance in chronic myeloid leukemia**. *Blood reviews* 2011, **25**(6):279-290.
34. Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J: **Structural mechanism for STI-571 inhibition of abelson tyrosine kinase**. *Science* 2000, **289**(5486):1938-1942.
35. Nagar B: **c-Abl Tyrosine Kinase and Inhibition by the Cancer Drug Imatinib (Gleevec/STI-571)**. *The Journal of Nutrition* 2007, **137**(6):1518S-1523S.
36. Rossari F, Minutolo F, Orciuolo E: **Past, present, and future of Bcr-Abl inhibitors: from chemical development to clinical efficacy**. *Journal of hematology & oncology* 2018, **11**(1):84.

37. Reddy EP, Aggarwal AK: **The ins and outs of bcr-abl inhibition**. *Genes Cancer* 2012, **3**(5-6):447-454.
38. Nagar B, Bornmann WG, Pellicena P, Schindler T, Veach DR, Miller WT, Clarkson B, Kuriyan J: **Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571)**. *Cancer research* 2002, **62**(15):4236-4243.
39. An X, Tiwari AK, Sun Y, Ding P-R, Ashby Jr CR, Chen Z-S: **BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: a review**. *Leukemia research* 2010, **34**(10):1255-1268.
40. Loscocco F, Visani G, Galimberti S, Curti A, Isidori A: **BCR-ABL Independent Mechanisms of Resistance in Chronic Myeloid Leukemia**. *Frontiers in Oncology* 2019, **9**(939).
41. Hehlmann R: **How I treat CML blast crisis**. *Blood* 2012, **120**(4):737-747.
42. Holyoake TL, Helgason GV: **Do we need more drugs for chronic myeloid leukemia?** *Immunological reviews* 2015, **263**(1):106-123.
43. Ma L, Shan Y, Bai R, Xue L, Eide CA, Ou J, Zhu LJ, Hutchinson L, Cerny J, Khoury HJ *et al*: **A therapeutically targetable mechanism of BCR-ABL-independent imatinib resistance in chronic myeloid leukemia**. *Science translational medicine* 2014, **6**(252):252ra121.
44. Donato NJ, Wu JY, Stapley J, Gallick G, Lin H, Arlinghaus R, Talpaz M: **BCR-ABL independence and LYN kinase overexpression in chronic myelogenous leukemia cells selected for resistance to STI571**. *Blood* 2003, **101**(2):690-698.
45. O'Hare T, Eide CA, Deininger MW: **Persistent LYN signaling in imatinib-resistant, BCR-ABL-independent chronic myelogenous leukemia**. *Journal of the National Cancer Institute* 2008, **100**(13):908-909.
46. Jabbour E, Kantarjian HM, Jones D, Shan J, O'Brien S, Reddy N, Wierda WG, Faderl S, Garcia-Manero G, Verstovsek S *et al*: **Imatinib mesylate dose escalation is associated with durable responses in patients with chronic myeloid leukemia after cytogenetic failure on standard-dose imatinib therapy**. *Blood* 2009, **113**(10):2154-2160.
47. Eechoute K, Sparreboom A, Burger H, Franke RM, Schiavon G, Verweij J, Loos WJ, Wiemer EAC, Mathijssen RHJ: **Drug Transporters and Imatinib Treatment: Implications for Clinical Practice**. *Clinical Cancer Research* 2011, **17**(3):406-415.
48. Rosti G, Castagnetti F, Gugliotta G, Baccarani M: **Tyrosine kinase inhibitors in chronic myeloid leukaemia: which, when, for whom?** *Nature reviews Clinical oncology* 2017, **14**(3):141-154.
49. Zaharieva M, Amudov G, Konstantinov S, Guenova M: **Modern Therapy of Chronic Myeloid Leukemia**. *Edited by Margarita Guenova and Gueorgui Balatzenko* 2013:227.
50. Yang K, Fu LW: **Mechanisms of resistance to BCR-ABL TKIs and the therapeutic strategies: A review**. *Crit Rev Oncol Hematol* 2015, **93**(3):277-292.
51. Shamroe CL, Comeau JM: **Ponatinib: A new tyrosine kinase inhibitor for the treatment of chronic myeloid leukemia and Philadelphia chromosome-positive acute lymphoblastic leukemia**. *The Annals of pharmacotherapy* 2013, **47**(11):1540-1546.
52. Maia RC, Vasconcelos FC, Souza PS, Rumjanek VM: **Towards Comprehension of the ABCB1/P-Glycoprotein Role in Chronic Myeloid Leukemia**. *Molecules* 2018, **23**(1):119.
53. Schmidt S: **Short overview on the current treatment of chronic myeloid leukemia in chronic phase**. *Memo* 2016, **9**(4):157-162.

54. Cortes JE, Kim DW, Pinilla-Ibarz J, le Coutre P, Paquette R, Chuah C, Nicolini FE, Apperley JF, Khoury HJ, Talpaz M *et al*: **A phase 2 trial of ponatinib in Philadelphia chromosome-positive leukemias**. *The New England journal of medicine* 2013, **369**(19):1783-1796.
55. Shi Y, Rand AJ, Crow JH, Moore JO, Lagoo AS: **Blast phase in chronic myelogenous leukemia is skewed toward unusual blast types in patients treated with tyrosine kinase inhibitors: a comparative study of 67 cases**. *American journal of clinical pathology* 2015, **143**(1):105-119.
56. Zabriskie MS, Eide CA, Tantravahi SK, Vellore NA, Estrada J, Nicolini FE, Khoury HJ, Larson RA, Konopleva M, Cortes JE *et al*: **BCR-ABL1 compound mutations combining key kinase domain positions confer clinical resistance to ponatinib in Ph chromosome-positive leukemia**. *Cancer cell* 2014, **26**(3):428-442.
57. Mitchell R, Hopcroft LEM, Baquero P, Allan EK, Hewit K, James D, Hamilton G, Mukhopadhyay A, O'Prey J, Hair A *et al*: **Targeting BCR-ABL-Independent TKI Resistance in Chronic Myeloid Leukemia by mTOR and Autophagy Inhibition**. *Journal of the National Cancer Institute* 2018, **110**(5):467-478.
58. Singhal MK, Sengar M, Nair R: **Summary of the published Indian data on chronic myeloid leukemia**. *South Asian Journal of Cancer* 2016, **5**(3):162-165.
59. Ganesan P, Kumar L: **Chronic Myeloid Leukemia in India**. *Journal of global oncology* 2017, **3**(1):64-71.
60. Massimino M, Stella S, Tirro E, Romano C, Pennisi MS, Puma A, Manzella L, Zanghi A, Stagno F, Di Raimondo F *et al*: **Non ABL-directed inhibitors as alternative treatment strategies for chronic myeloid leukemia**. 2018, **17**(1):56.
61. Winer ES, DeAngelo DJ: **A Review of Omacetaxine: A Chronic Myeloid Leukemia Treatment Resurrected**. *Oncology and Therapy* 2018, **6**(1):9-20.
62. Khajapeer KV, Baskaran R: **Hsp90 Inhibitors for the Treatment of Chronic Myeloid Leukemia**. *Leuk Res Treatment* 2015, **2015**:757694-757694.
63. Van Norman GA: **Phase II Trials in Drug Development and Adaptive Trial Design**. *JACC Basic Transl Sci* 2019, **4**(3):428-437.
64. Govekar R: **Identification of Therapeutic Targets for Cancer: Proteomic Technologies and Strategies are the Key to Success**. *Biomed Res J* 2015, **2**(2):166-178.
65. Xie F, Liu T, Qian WJ, Petyuk VA, Smith RD: **Liquid chromatography-mass spectrometry-based quantitative proteomics**. *The Journal of biological chemistry* 2011, **286**(29):25443-25449.
66. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R: **Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis**. *Molecular & Cellular Proteomics* 2012, **11**(6):O111. 016717.
67. Ludwig C, Gillet L, Rosenberger G: **Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial**. 2018, **14**(8):e8126.
68. Fuller HR, Morris GE: **Quantitative Proteomics Using iTRAQ Labeling and Mass Spectrometry**. 2012.
69. Huang Q, Yang L, Luo J, Guo L, Wang Z, Yang X, Jin W, Fang Y, Ye J, Shan B *et al*: **SWATH enables precise label-free quantification on proteome scale**. *Proteomics* 2015, **15**(7):1215-1223.
70. Wasinger VC, Zeng M, Yau Y: **Current Status and Advances in Quantitative Proteomic Mass Spectrometry**. *International Journal of Proteomics* 2013, **2013**:12.

71. Chu H-W, Chang K-P, Hsu C-W, Chang IY-F, Liu H-P, Chen Y-T, Wu C-C: **Identification of Salivary Biomarkers for Oral Cancer Detection with Untargeted and Targeted Quantitative Proteomics Approaches**. *Molecular & Cellular Proteomics* 2019, **18**(9):1796.
72. Rehman I, Evans CA, Glen A, Cross SS, Eaton CL, Down J, Pesce G, Phillips JT, Yen OS, Thalmann GN *et al*: **iTRAQ Identification of Candidate Serum Biomarkers Associated with Metastatic Progression of Human Prostate Cancer**. *PloS one* 2012, **7**(2):e30885.
73. Leonardi A, Palmigiano A, Mazzola EA, Messina A, Milazzo EMS, Bortolotti M, Garozzo D: **Identification of human tear fluid biomarkers in vernal keratoconjunctivitis using iTRAQ quantitative proteomics**. *Allergy* 2014, **69**(2):254-260.
74. Subbannayya Y, Syed N, Barbhuiya MA, Raja R, Marimuthu A, Sahasrabuddhe N, Pinto SM, Manda SS, Renuse S, Manju HC *et al*: **Calcium calmodulin dependent kinase kinase 2 - a novel therapeutic target for gastric adenocarcinoma**. *Cancer Biology & Therapy* 2015, **16**(2):336-345.
75. Luo Y, Mok TS, Lin X, Zhang W, Cui Y, Guo J, Chen X, Zhang T, Wang T: **SWATH-based proteomics identified carbonic anhydrase 2 as a potential diagnosis biomarker for nasopharyngeal carcinoma**. *Scientific reports* 2017, **7**:41191.
76. Ortea I, Ruiz-Sánchez I, Cañete R, Caballero-Villarraso J, Cañete MD: **Identification of candidate serum biomarkers of childhood-onset growth hormone deficiency using SWATH-MS and feature selection**. *Journal of proteomics* 2018, **175**:105-113.
77. Feng X, Bai X, Ni J, Wasinger VC, Beretov J, Zhu Y, Graham P, Li Y: **CHTOP in Chemoresistant Epithelial Ovarian Cancer: A Novel and Potential Therapeutic Target**. *Frontiers in Oncology* 2019, **9**(557).
78. Vähätupa M, Nättinen J, Jylhä A, Aapola U, Kataja M, Kööbi P, Järvinen TAH, Uusitalo H, Uusitalo-Järvinen H: **SWATH-MS Proteomic Analysis of Oxygen-Induced Retinopathy Reveals Novel Potential Therapeutic Targets**. *Investigative Ophthalmology & Visual Science* 2018, **59**(8):3294-3306.
79. Fatima I, Sadaf S, Musharraf SG, Hashmi N, Akhtar MW: **CD5 molecule-like and transthyretin as putative biomarkers of chronic myeloid leukemia - an insight from the proteomic analysis of human plasma**. *Scientific reports* 2017, **7**(1):40943.
80. Alaiya AA, Aljurf M, Shinwari Z, Almohareb F, Malhan H, Alzahrani H, Owaidah T, Fox J, Alsharif F, Mohamed SY *et al*: **Protein signatures as potential surrogate biomarkers for stratification and prediction of treatment response in chronic myeloid leukemia patients**. *International journal of oncology* 2016, **49**(3):913-933.
81. Brehme M, Hantschel O, Colinge J, Kaupe I, Planyavsky M, Kocher T, Mechtler K, Bennett KL, Superti-Furga G: **Charting the molecular network of the drug target Bcr-Abl**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(18):7414-7419.
82. Preisinger C, Kolch W: **The Bcr-Abl kinase regulates the actin cytoskeleton via a GADS/Slp-76/Nck1 adaptor protein pathway**. *Cellular signalling* 2010, **22**(5):848-856.
83. Patel H, Nteliopoulos G, Nikolakopoulou Z, Jackson A, Gordon MY: **Antibody arrays identify protein-protein interactions in chronic myeloid leukaemia**. *British journal of haematology* 2011, **152**(5):611-614.
84. Kuzelova K, Pluskalova M, Grebenova D, Pavlaskova K, Halada P, Hrkal Z: **Changes in cell adhesivity and cytoskeleton-related proteins during imatinib-induced apoptosis of leukemic JURL-MK1 cells**. *Journal of cellular biochemistry* 2010, **111**(6):1413-1425.

85. Pizzatti L, Sa LA, de Souza JM, Bisch PM, Abdelhay E: **Altered protein profile in chronic myeloid leukemia chronic phase identified by a comparative proteomic study**. *Biochimica et biophysica acta* 2006, **1764**(5):929-942.
86. Griffiths SD, Burthem J, Unwin RD, Holyoake TL, Melo JV, Lucas GS, Whetton AD: **The use of isobaric tag peptide labeling (iTRAQ) and mass spectrometry to examine rare, primitive hematopoietic cells from patients with chronic myeloid leukemia**. *Molecular biotechnology* 2007, **36**(2):81-89.
87. Park J, Kim S, Oh JK, Kim JY, Yoon SS, Lee D, Kim Y: **Identification of differentially expressed proteins in imatinib mesylate-resistant chronic myelogenous cells**. *Journal of biochemistry and molecular biology* 2005, **38**(6):725-738.
88. Xiong L, Zhang J, Yuan B, Dong X, Jiang X, Wang Y: **Global proteome quantification for discovering imatinib-induced perturbation of multiple biological pathways in K562 human chronic myeloid leukemia cells**. *Journal of proteome research* 2010, **9**(11):6007-6015.
89. Arvaniti K, Papadioti A, Kinigopoulou M, Theodorou V, Skobridis K, Tsiotis G: **Proteome Changes Induced by Imatinib and Novel Imatinib Derivatives in K562 Human Chronic Myeloid Leukemia Cells**. *Proteomes* 2014, **2**(3):363-381.
90. Balabanov S, Gontarewicz A, Ziegler P, Hartmann U, Kammer W, Copland M, Brassat U, Priemer M, Hauber I, Wilhelm T *et al*: **Hypusination of eukaryotic initiation factor 5A (eIF5A): a novel therapeutic target in BCR-ABL-positive leukemias identified by a proteomics approach**. *Blood* 2007, **109**(4):1701-1711.
91. Rosenhahn J, Weise A, Michel S, Hennig K, Hartmann I, Schiefner J, Schubert K, Liehr T, von Eggeling F, Loncarevic IF: **Cytogenetic characterisation and proteomic profiling of the Imatinib-resistant cell line KCL22-R**. *International journal of oncology* 2007, **31**(1):121-128.
92. Zhang J, Jin Z, Du Q, Li R, Yao F, Huang B, Xu N, Xu L, Luo X, Liu X: **Analysis of altered proteins related to blast crisis in chronic myeloid leukemia by proteomic study**. *International journal of laboratory hematology* 2012, **34**(3):267-273.
93. Ferrari G, Pastorelli R, Buchi F, Spinelli E, Gozzini A, Bosi A, Santini V: **Comparative proteomic analysis of chronic myelogenous leukemia cells: inside the mechanism of imatinib resistance**. *Journal of proteome research* 2007, **6**(1):367-375.
94. Colavita I, Esposito N, Martinelli R, Catanzano F, Melo JV, Pane F, Ruoppolo M, Salvatore F: **Gaining insights into the Bcr-Abl activity-independent mechanisms of resistance to imatinib mesylate in KCL22 cells: a comparative proteomic approach**. *Biochimica et biophysica acta* 2010, **1804**(10):1974-1987.
95. Pocaly M, Lagarde V, Etienne G, Dupouy M, Lapaillerie D, Claverol S, Vilain S, Bonneu M, Turcq B, Mahon FX *et al*: **Proteomic analysis of an imatinib-resistant K562 cell line highlights opposing roles of heat shock cognate 70 and heat shock 70 proteins in resistance**. *Proteomics* 2008, **8**(12):2394-2406.
96. Pocaly M, Lagarde V, Etienne G, Ribeil JA, Claverol S, Bonneu M, Moreau-Gaudry F, Guyonnet-Duperat V, Hermine O, Melo JV *et al*: **Overexpression of the heat-shock protein 70 is associated to imatinib resistance in chronic myeloid leukemia**. *Leukemia* 2007, **21**(1):93-101.
97. Toman O, Kabickova T, Vit O, Fiser R, Polakova KM, Zach J, Linhartova J, Vyoral D, Petrak J: **Proteomic analysis of imatinib-resistant CML-T1 cells reveals calcium homeostasis as a potential therapeutic target**. *Oncology reports* 2016, **36**(3):1258-1268.

98. Quintas-Cardama A, Qiu YH, Post SM, Zhang Y, Creighton CJ, Cortes J, Kornblau SM: **Reverse phase protein array profiling reveals distinct proteomic signatures associated with chronic myeloid leukemia progression and with chronic phase in the CD34-positive compartment**. *Cancer* 2012, **118**(21):5283-5292.
99. Zhang QY, Mao JH, Liu P, Huang QH, Lu J, Xie YY, Weng L, Zhang Y, Chen Q, Chen SJ *et al*: **A systems biology understanding of the synergistic effects of arsenic sulfide and Imatinib in BCR/ABL-associated leukemia**. *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(9):3378-3383.
100. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R: **InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams**. *BMC bioinformatics* 2015, **16**(1):169-169.
101. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P *et al*: **STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets**. *Nucleic acids research* 2019, **47**(D1):D607-D613.
102. Alikian M, Gerrard G, Subramanian PG, Mudge K, Foskett P, Khorashad JS, Lim AC, Marin D, Milojkovic D, Reid A *et al*: **BCR-ABL1 kinase domain mutations: methodology and clinical evaluation**. *American journal of hematology* 2012, **87**(3):298-304.
103. Sana TR, Waddell K, Fischer SM: **A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome**. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* 2008, **871**(2):314-321.
104. Navarro P, Kuharev J, Gillet LC, Bernhardt OM, MacLean B, Rost HL, Tate SA, Tsou CC, Reiter L, Distler U *et al*: **A multicenter study benchmarks software tools for label-free proteome quantification**. *Nat Biotechnol* 2016, **34**(11):1130-1136.
105. Tan SLW, Chadha S, Liu Y, Gabasova E, Perera D, Ahmed K, Constantinou S, Renaudin X, Lee M, Aebersold R *et al*: **A Class of Environmental and Endogenous Toxins Induces BRCA2 Haploinsufficiency and Genome Instability**. *Cell* 2017, **169**(6):1105-1118.e1115.
106. Guo T, Kouvonen P, Koh CC, Gillet LC, Wolski WE, Rost HL: **Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps**. 2015, **21**(4):407-413.
107. Rosenberger G, Koh CC, Guo T, Rost HL, Kouvonen P, Collins BC, Heusel M, Liu Y, Caron E, Vichalkovski A *et al*: **A repository of assays to quantify 10,000 human proteins by SWATH-MS**. *Scientific data* 2014, **1**:140031.
108. Chawade A, Alexandersson E, Levander F: **Normalyzer: a tool for rapid evaluation of normalization methods for omics data sets**. *Journal of proteome research* 2014, **13**(6):3114-3120.
109. Valikangas T, Suomi T, Elo LL: **A systematic evaluation of normalization methods in quantitative label-free proteomics**. *Briefings in bioinformatics* 2018, **19**(1):1-11.
110. Karpievitch YV, Dabney AR, Smith RD: **Normalization and missing value imputation for label-free LC-MS analysis**. *BMC Bioinformatics* 2012, **13**(16):S5.
111. Callister SJ, Barry RC, Adkins JN, Johnson ET, Qian W-j, Webb-Robertson B-JM, Smith RD, Lipton MS: **Normalization Approaches for Removing Systematic Biases Associated with Mass Spectrometry and Label-Free Proteomics**. *Journal of proteome research* 2006, **5**(2):277-286.

112. Kultima K, Nilsson A, Scholz B, Rossbach UL, Fälth M, Andrén PE: **Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides**. *Molecular & cellular proteomics : MCP* 2009, **8**(10):2285-2295.
113. Young PR, McLaughlin MM, Kumar S, Kassis S, Doyle ML, McNulty D, Gallagher TF, Fisher S, McDonnell PC, Carr SA *et al*: **Pyridinyl imidazole inhibitors of p38 mitogen-activated protein kinase bind in the ATP site**. *The Journal of biological chemistry* 1997, **272**(18):12116-12121.
114. Cuenda A, Rousseau S: **p38 MAP-Kinases pathway regulation, function and role in human diseases**. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 2007, **1773**(8):1358-1375.
115. Funding AT, Johansen C, Kragballe K, Otkjær K, Jensen UB, Madsen MW, Fjording MS, Finnemann J, Skak-Nielsen T, Paludan SR *et al*: **Mitogen- and Stress-Activated Protein Kinase 1 Is Activated in Lesional Psoriatic Epidermis and Regulates the Expression of Pro-Inflammatory Cytokines**. *Journal of Investigative Dermatology* 2006, **126**(8):1784-1791.
116. Chihara D, Ito H, Matsuda T, Katanoda K, Shibata A, Saika K, Sobue T, Matsuo K: **Decreasing trend in mortality of chronic myelogenous leukemia patients after introduction of imatinib in Japan and the U.S**. *Oncologist* 2012, **17**(12):1547-1550.
117. Di Felice E, Roncaglia F, Venturelli F, Mangone L, Luminari S, Cirilli C, Carrozzi G, Giorgi Rossi P: **The impact of introducing tyrosine kinase inhibitors on chronic myeloid leukemia survival: a population-based study**. *BMC cancer* 2018, **18**(1):1069-1069.
118. Westfall DE, Zhang L, Song S, Lee S: **Concurrent megakaryocytic and erythroid chronic myelogenous leukemia blast crisis**. *Archives of pathology & laboratory medicine* 2008, **132**(6):1021-1025.
119. Yu D, Yu LL: **Optimal conditions of chemotherapeutic sensitivity in K562 cell line using tetrazolium dye assay**. *Zhongguo yao li xue bao = Acta pharmacologica Sinica* 1993, **14**(2):137-140.
120. Kubonishi I, Miyoshi I: **Establishment of a Ph1 chromosome-positive cell line from chronic myelogenous leukemia in blast crisis**. *International journal of cell cloning* 1983, **1**(2):105-117.
121. Kishi K: **A new leukemia cell line with Philadelphia chromosome characterized as basophil precursors**. *Leukemia research* 1985, **9**(3):381-390.
122. Virgili A, Nacheva EP: **Genomic amplification of BCR/ABL1 and a region downstream of ABL1 in chronic myeloid leukaemia: a FISH mapping study of CML patients and cell lines**. *Mol Cytogenet* 2010, **3**:15-15.
123. Soverini S, Branford S, Nicolini FE, Talpaz M, Deininger MWN, Martinelli G, Müller MC, Radich JP, Shah NP: **Implications of BCR-ABL1 kinase domain-mediated resistance in chronic myeloid leukemia**. *Leukemia research* 2014, **38**(1):10-20.
124. Rejali L, Poopak B, Hasanzad M, Sheikhsofla F, Varnoosfaderani AS, Safari N, Rabieipoor S: **Characterizing of Four Common BCR-ABL Kinase Domain Mutations (T315I, Y253H, M351T and E255K) in Iranian Chronic Myelogenous Leukemia Patients With Imatinib Resistance**. *Iran J Cancer Prev* 2015, **8**(3):e2334-e2334.
125. Bouchet S, Dulucq S, Pasquet JM, Lagarde V, Molimard M, Mahon FX: **From in vitro to in vivo: intracellular determination of imatinib and nilotinib may be related with clinical outcome**. *Leukemia* 2013, **27**(8):1757-1759.
126. Jiang X, Zhao Y, Smith C, Gasparetto M, Turhan A, Eaves A, Eaves C: **Chronic myeloid leukemia stem cells possess multiple unique features of resistance to BCR-ABL targeted therapies**. *Leukemia* 2007, **21**(5):926-935.

127. Alves R, Fonseca AR, Gonçalves AC, Ferreira-Teixeira M, Lima J, Abrantes AM, Alves V, Rodrigues-Santos P, Jorge L, Matoso E *et al*: **Drug transporters play a key role in the complex process of Imatinib resistance in vitro**. *Leukemia research* 2015, **39**(3):355-360.
128. Faca V, Pitteri SJ, Newcomb L, Glukhova V, Phanstiel D, Krasnoselsky A, Zhang Q, Struthers J, Wang H, Eng J *et al*: **Contribution of Protein Fractionation to Depth of Analysis of the Serum and Plasma Proteomes**. *Journal of proteome research* 2007, **6**(9):3558-3565.
129. Righetti P, Castagna A, Antonioli P, Boschetti E: **Prefractionation techniques in proteome analysis: The mining tools of the third millennium**. *Electrophoresis* 2005, **26**:297-319.
130. Blattmann P, Henriques D, Zimmermann M, Frommelt F, Sauer U, Saez-Rodriguez J, Aebersold R: **Systems Pharmacology Dissection of Cholesterol Regulation Reveals Determinants of Large Pharmacodynamic Variability between Cell Lines**. *Cell systems* 2017, **5**(6):604-619.e607.
131. Zhang X, Walsh T, Atherton JJ, Kostner K, Schulz B, Punyadeera C: **Identification and Validation of a Salivary Protein Panel to Detect Heart Failure Early**. *Theranostics* 2017, **7**(18):4350-4358.
132. Bjelosevic S, Pascovici D, Ping H, Karlaftis V, Zaw T, Song X, Molloy MP, Monagle P, Ignjatovic V: **Quantitative Age-specific Variability of Plasma Proteins in Healthy Neonates, Children and Adults**. *Molecular & cellular proteomics : MCP* 2017, **16**(5):924-935.
133. Liu Y, Chen J, Sethi A, Li QK, Chen L, Collins B, Gillet LC, Wollscheid B, Zhang H, Aebersold R: **Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers N-acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness**. *Molecular & cellular proteomics : MCP* 2014, **13**(7):1753-1768.
134. Arya S, Emri E, Synowsky SA, Shirran SL, Barzegar-Befroei N, Peto T, Botting CH, Lengyel I, Stewart AJ: **Quantitative analysis of hydroxyapatite-binding plasma proteins in genotyped individuals with late-stage age-related macular degeneration**. *Experimental eye research* 2018, **172**:21-29.
135. Batkulwar K, Godbole R, Banarjee R, Kassaar O, Williams RJ, Kulkarni MJ: **Advanced Glycation End Products Modulate Amyloidogenic APP Processing and Tau Phosphorylation: A Mechanistic Link between Glycation and the Development of Alzheimer's Disease**. 2018, **9**(5):988-1000.
136. Schillaci O, Fontana S: **Exosomes from metastatic cancer cells transfer amoeboid phenotype to non-metastatic cells and increase endothelial permeability: their emerging role in tumor heterogeneity**. 2017, **7**(1):4711.
137. Lewandowska AE, Macur K, Czaplewska P, Liss J, Lukaszuk K, Oldziej S: **Qualitative and Quantitative Analysis of Proteome and Peptidome of Human Follicular Fluid Using Multiple Samples from Single Donor with LC-MS and SWATH Methodology**. *Journal of proteome research* 2017, **16**(8):3053-3067.
138. Ortea I, Rodriguez-Ariza A, Chicano-Galvez E, Arenas Vacas MS, Jurado Gamez B: **Discovery of potential protein biomarkers of lung adenocarcinoma in bronchoalveolar lavage fluid by SWATH MS data-independent acquisition and targeted data extraction**. *Journal of proteomics* 2016, **138**:106-114.
139. Lin Q, Lim HS, Lin HL, Tan HT, Lim TK, Cheong WK, Cheah PY, Tang CL, Chow PK, Chung MC: **Analysis of colorectal cancer glyco-secretome identifies laminin beta-1 (LAMB1)**

as a potential serological biomarker for colorectal cancer. *Proteomics* 2015, **15**(22):3905-3920.
140. Korwar AM, Vannuruswamy G, Jagadeeshaprasad MG, Jayaramaiah RH, Bhat S, Regin BS, Ramaswamy S, Giri AP, Mohan V, Balasubramanyam M *et al*: **Development of Diagnostic Fragment Ion Library for Glycated Peptides of Human Serum Albumin: Targeted Quantification in Prediabetic, Diabetic, and Microalbuminuria Plasma by Parallel Reaction Monitoring, SWATH, and MSE**. *Molecular & cellular proteomics : MCP* 2015, **14**(8):2150-2159.
141. Tan HT, Lim TK, Richards AM, Kofidis T, Teoh KL, Ling LH, Chung MC: **Unravelling the proteome of degenerative human mitral valves**. *Proteomics* 2015, **15**(17):2934-2944.
142. Chaudhari PR, Charles SE, D'Souza ZC, Vaidya MM: **Hemidesmosomal linker proteins regulate cell motility, invasion and tumorigenicity in oral squamous cell carcinoma derived cells**. *Experimental cell research* 2017, **360**(2):125-137.
143. Liu Y, Borel C, Li L, Muller T, Williams EG, Germain PL, Buljan M, Sajic T, Boersema PJ, Shao W *et al*: **Systematic proteome and proteostasis profiling in human Trisomy 21 fibroblast cells**. 2017, **8**(1):1212.
144. Chen L, Li J, Guo T, Ghosh S, Koh SK, Tian D, Zhang L, Jia D, Beuerman RW, Aebersold R *et al*: **Global Metabonomic and Proteomic Analysis of Human Conjunctival Epithelial Cells (IOBA-NHC) in Response to Hyperosmotic Stress**. *Journal of proteome research* 2015, **14**(9):3982-3995.
145. Orellana CA, Marcellin E, Schulz BL, Nouwens AS, Gray PP, Nielsen LK: **High-antibody-producing Chinese hamster ovary cells up-regulate intracellular protein transport and glutathione synthesis**. *Journal of proteome research* 2015, **14**(2):609-618.
146. Netherton JK, Hetherington L, Ogle RA, Velkov T, Baker MA: **Proteomic analysis of good- and poor-quality human sperm demonstrates that several proteins are routinely aberrantly regulated**. *Biology of reproduction* 2018, **99**(2):395-408.
147. Nemeth J, Vongrad V, Metzner KJ, Strouvelle VP, Weber R, Pedrioli P, Aebersold R, Gunthard HF, Collins BC: **In Vivo and in Vitro Proteome Analysis of Human Immunodeficiency Virus (HIV)-1-infected, Human CD4+ T Cells**. *Molecular & cellular proteomics : MCP* 2017, **16**(4 suppl 1):S108-s123.
148. Chang RY, Etheridge N, Nouwens AS, Dodd PR: **SWATH analysis of the synaptic proteome in Alzheimer's disease**. *Neurochemistry international* 2015, **87**:1-12.
149. Martins-Marques T, Anjo SI, Pereira P, Manadas B, Girao H: **Interacting Network of the Gap Junction (GJ) Protein Connexin43 (Cx43) is Modulated by Ischemia and Reperfusion in the Heart**. 2015, **14**(11):3040-3055.
150. Villeneuve LM, Stauch KL, Fox HS: **Proteomic analysis of the mitochondria from embryonic and postnatal rat brains reveals response to developmental changes in energy demands**. *Journal of proteomics* 2014, **109**:228-239.
151. McCarthy DJ, Smyth GK: **Testing significance relative to a fold-change threshold is a TREAT**. *Bioinformatics (Oxford, England)* 2009, **25**(6):765-771.
152. Xiao Y, Hsiao T-H, Suresh U, Chen H-IH, Wu X, Wolf SE, Chen Y: **A novel significance score for gene selection and ranking**. *Bioinformatics (Oxford, England)* 2014, **30**(6):801-807.
153. Dalman MR, Deeter A, Nimishakavi G, Duan Z-H: **Fold change and p-value cutoffs significantly alter microarray interpretations**. *BMC Bioinformatics* 2012, **13**(2):S11.
154. Narasimhan M, Kannan S, Chawade A, Bhattacharjee A, Govekar R: **Clinical biomarker discovery by SWATH-MS based label-free quantitative proteomics: impact of criteria**

for identification of differentiators and data normalization method. *Journal of Translational Medicine* 2019, **17**(1):184.
155. Warsch W, Kollmann K, Eckelhart E, Fajmann S, Cerny-Reiterer S, Hölbl A, Gleixner KV, Dworzak M, Mayerhofer M, Hoermann G *et al*: **High STAT5 levels mediate imatinib resistance and indicate disease progression in chronic myeloid leukemia**. *Blood* 2011, **117**(12):3409-3420.
156. Brachet-Botineau M, Deynoux M, Vallet N, Polomski M, Juen L, Herault O, Mazurier F, Viaud-Massuard MC, Prie G, Gouilleux F: **A Novel Inhibitor of STAT5 Signaling Overcomes Chemotherapy Resistance in Myeloid Leukemia Cells**. *Cancers (Basel)* 2019, **11**(12).
157. Gleixner KV, Schneeweiss M, Eisenwort G, Berger D, Herrmann H, Blatt K, Greiner G, Byrgazov K, Hoermann G, Konopleva M *et al*: **Combined targeting of STAT3 and STAT5: a novel approach to overcome drug resistance in chronic myeloid leukemia**. *Haematologica* 2017, **102**(9):1519.
158. Berger A, Sexl V, Valent P, Moriggl R: **Inhibition of STAT5: a therapeutic option in BCR-ABL1-driven leukemia**. *Oncotarget* 2014, **5**(20):9564-9576.
159. Nelson EA, Walker SR, Weisberg E, Bar-Natan M, Barrett R, Gashin LB, Terrell S, Klitgaard JL, Santo L, Addorio MR *et al*: **The STAT5 inhibitor pimozide decreases survival of chronic myelogenous leukemia cells resistant to kinase inhibitors**. *Blood* 2011, **117**(12):3421-3429.
160. Uhart M, Bustos DM: **Human 14-3-3 paralogs differences uncovered by cross-talk of phosphorylation and lysine acetylation**. *PloS one* 2013, **8**(2):e55703-e55703.
161. Mancini M, Veljkovic N, Corradi V, Zuffa E, Corrado P, Pagnotta E, Martinelli G, Barbieri E, Santucci MA: **14-3-3 ligand prevents nuclear import of c-ABL protein in chronic myeloid leukemia**. *Traffic* 2009, **10**(6):637-647.
162. Bulavin DV, Fornace AJ, Jr.: **p38 MAP kinase's emerging role as a tumor suppressor**. *Adv Cancer Res* 2004, **92**:95-118.
163. Donnelly SM, Paplomata E, Peake BM, Sanabria E, Chen Z, Nahta R: **P38 MAPK contributes to resistance and invasiveness of HER2- overexpressing breast cancer**. *Current medicinal chemistry* 2014, **21**(4):501-510.
164. García-Cano J, Roche O, Cimas FJ, Pascual-Serra R, Ortega-Muelas M, Fernández-Aroca DM, Sánchez-Prieto R: **p38MAPK and Chemotherapy: We Always Need to Hear Both Sides of the Story**. *Front Cell Dev Biol* 2016, **4**:69-69.
165. Parmar S, Katsoulidis E, Verma A, Li Y, Sassano A, Lal L, Majchrzak B, Ravandi F, Tallman MS, Fish EN *et al*: **Role of the p38 mitogen-activated protein kinase pathway in the generation of the effects of imatinib mesylate (STI571) in BCR-ABL-expressing cells**. *The Journal of biological chemistry* 2004, **279**(24):25345-25352.
166. Kohmura K, Miyakawa Y, Kawai Y, Ikeda Y, Kizaki M: **Different roles of p38 MAPK and ERK in STI571-induced multi-lineage differentiation of K562 cells**. *J Cell Physiol* 2004, **198**(3):370-376.
167. Giafis N, Katsoulidis E, Sassano A, Tallman MS, Higgins LS, Nebreda AR, Davis RJ, Platanias LC: **Role of the p38 mitogen-activated protein kinase pathway in the generation of arsenic trioxide-dependent cellular responses**. *Cancer research* 2006, **66**(13):6763-6771.
168. Newby LK, Marber MS, Melloni C, Sarov-Blat L, Aberle LH, Aylward PE, Cai G, de Winter RJ, Hamm CW, Heitner JF *et al*: **Losmapimod, a novel p38 mitogen-activated protein kinase inhibitor, in non-ST-segment elevation myocardial infarction: a randomised phase 2 trial**. *Lancet (London, England)* 2014, **384**(9949):1187-1195.

169. O'Donoghue ML, Glaser R, Aylward PE, Cavender MA, Crisp A, Fox KAA, Laws I, Lopez-Sendon JL, Steg PG, Theroux P *et al*: **Rationale and design of the LosmApimod To Inhibit p38 MAP kinase as a TherapeUtic target and moDify outcomes after an acute coronary syndromE trial**. *American heart journal* 2015, **169**(5):622-630.e626.

Narasimhan *et al. J Transl Med* (2019) 17:184
https://doi.org/10.1186/s12967-019-1937-9

Journal of Translational Medicine

RESEARCH

Open Access

Clinical biomarker discovery by SWATH-MS based label-free quantitative proteomics: impact of criteria for identification of differentiators and data normalization method

Mythreyi Narasimhan[1,2], Sadhana Kannan[1], Aakash Chawade[3], Atanu Bhattacharjee[4] and Rukmini Govekar[1,2*]

Abstract

Background: SWATH-MS has emerged as the strategy of choice for biomarker discovery due to the proteome coverage achieved in acquisition and provision to re-interrogate the data. However, in quantitative analysis using SWATH, each sample from the comparison group is run individually in mass spectrometer and the resulting inter-run variation may influence relative quantification and identification of biomarkers. Normalization of data to diminish this variation thereby becomes an essential step in SWATH data processing. In most reported studies, data normalization methods used are those provided in instrument-based data analysis software or those used for microarray data. This study, for the first time provides an experimental evidence for selection of normalization method optimal for biomarker identification.

Methods: The efficiency of 12 normalization methods to normalize SWATH-MS data was evaluated based on statistical criteria in 'Normalyzer'—a tool which provides comparative evaluation of normalization by different methods. Further, the suitability of normalized data for biomarker discovery was assessed by evaluating the clustering efficiency of differentiators, identified from the normalized data based on p-value, fold change and both, by hierarchical clustering in Genesis software v.1.8.1.

Results: Conventional statistical criteria identified VSN-G as the optimal method for normalization of SWATH data. However, differentiators identified from VSN-G normalized data failed to segregate test and control groups. We thus assessed data normalized by eleven other methods for their ability to yield differentiators which segregate the study groups. Datasets in our study demonstrated that differentiators identified based on p-value from data normalized with Loess-R stratified the study groups optimally.

Conclusion: This is the first report of experimentally tested strategy for SWATH-MS data processing with an emphasis on identification of clinically relevant biomarkers. Normalization of SWATH-MS data by Loess-R method and identification of differentiators based on p-value were found to be optimal for biomarker discovery in this study. The study also demonstrates the need to base the choice of normalization method on the application of the data.

Keywords: Proteomic, LC–MS, SWATH, Normalization, p-value, Fold change

*Correspondence: rgovekar@actrec.gov.in
[1] Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai 410210, India
Full list of author information is available at the end of the article

Background

Liquid chromatography-mass spectrometry (LC–MS) based quantitative proteomic profiling has substantially contributed to identification of disease biomarkers for improved diagnosis/better prognostication or to monitor response to therapy [1–3]. This is achieved through assessment of the ability of differentiators, identified by quantitative proteomics, to segregate the comparison groups distinctly by cluster analysis—an essential feature of biomarkers. Success of this process depends not only on selection of appropriate clinical samples and sample processing strategies, but also on mass spectrometry based-factors such as the depth of LC–MS profile and occurrence of instrumental or non-instrumental errors in the data. Therefore, use of mass spectrometers with capabilities for in-depth profiling and data processing strategies which reduce biases, errors and optimize the desired outcome is necessary.

The recent feature in MS—Sequential window acquisition of all theoretical fragment-ion spectra (SWATH) provides in depth profiling by data independent acquisition (DIA) [4]. It is preferred for profiling clinical samples, as in data or information-dependent acquisition (IDA) data from low expressers is lost permanently [5]. SWATH not only provides for fragmentation of almost all ions but also for re-interrogation of data, after detection capabilities are improvised to identify more number of proteins [4]. These features are conducive to profiling of clinical samples which are available in amounts insufficient for enrichment and are unavailable for reanalysis. A testimony to this is the wide use of SWATH-MS in clinical proteomics after its discovery in 2012. PubMed results show that 44% (20/45) of the SWATH-MS studies on clinical samples published till date are aimed at biomarker discovery or therapeutic target identification.

However, a feature in quantification by SWATH-MS, if overlooked, can hinder biomarker identification. Unlike labelled quantification by IDA wherein all samples for relative quantification are run together, in label-free quantification by SWATH, each sample from the comparison group is run individually in MS. This increases the probability of both systematic and random error. Intervention to reduce these variations by 'normalization' is thus a prerequisite to subsequent analysis of SWATH data for identification of differentiators. The data from reported SWATH-MS studies is normalized using either methods provided by the MS instrument-based software or those used to normalize microarray data [6–9]. As the source of systematic bias differs between MS and microarray, it is essential to experimentally validate the appropriate normalization strategy for SWATH data.

The present study was undertaken to experimentally identify an appropriate normalization method for SWATH-MS data. The statistical tool 'Normalyzer', which compares the efficiency of diverse methods to normalize 'omics' data based on statistical criteria [10], was used to achieve the same. Fu et al. [11] in their study to identify the optimal analysis chain have reported total ion current normalization as optimal for SWATH-MS data based on statistical end-point. Further, considering the wide application of SWATH-MS in biomarker identification, in this study we have supplemented the statistical evaluation with biologically relevant criteria of precise stratification of comparison groups by cluster analysis. Towards this (a) Normalization of data was assessed using 'Normalyzer' to identify the optimal method of normalization based on statistical criteria (b) from the data normalized by different methods in Normalyzer, differentiators between comparison groups were identified based on p-value, fold change and combination of both. The potential of these differentiators to segregate comparison groups distinctly, was assessed by cluster analysis. Detection of optimal method for normalization of SWATH-MS data and optimum criteria for identification of differentiators would have an impact on biomarker discovery.

Methods

The scheme of experiments employed to identify the normalization strategy optimum for SWATH-MS data is depicted in Fig. 1. It involves:

A. Inclusion of a quantitatively defined dataset from public domain, generated from hybrid of peptides from three different sources mixed in defined proportions, to serve as a 'reference set'. Generation of datasets using K562 cells for quantitation by SWATH-MS, referred to as 'study set' which includes one set with smaller number of samples and two sets with larger number of samples. Further, inclusion of two datasets from public domain comprising of larger sample size, to serve as 'validation set' to confirm the findings in the study set.
B. SWATH-MS analysis of reference, study and validation set.
C. Normalization of SWATH-MS data obtained from reference, study and validation sets using methods in Marker view software and Normalyzer and identification of optimum method of normalization based on statistical criteria.
D. Identification of differentiators from this normalized data based on criteria of p value, fold-change and both, followed by cluster analysis of these differentiators.

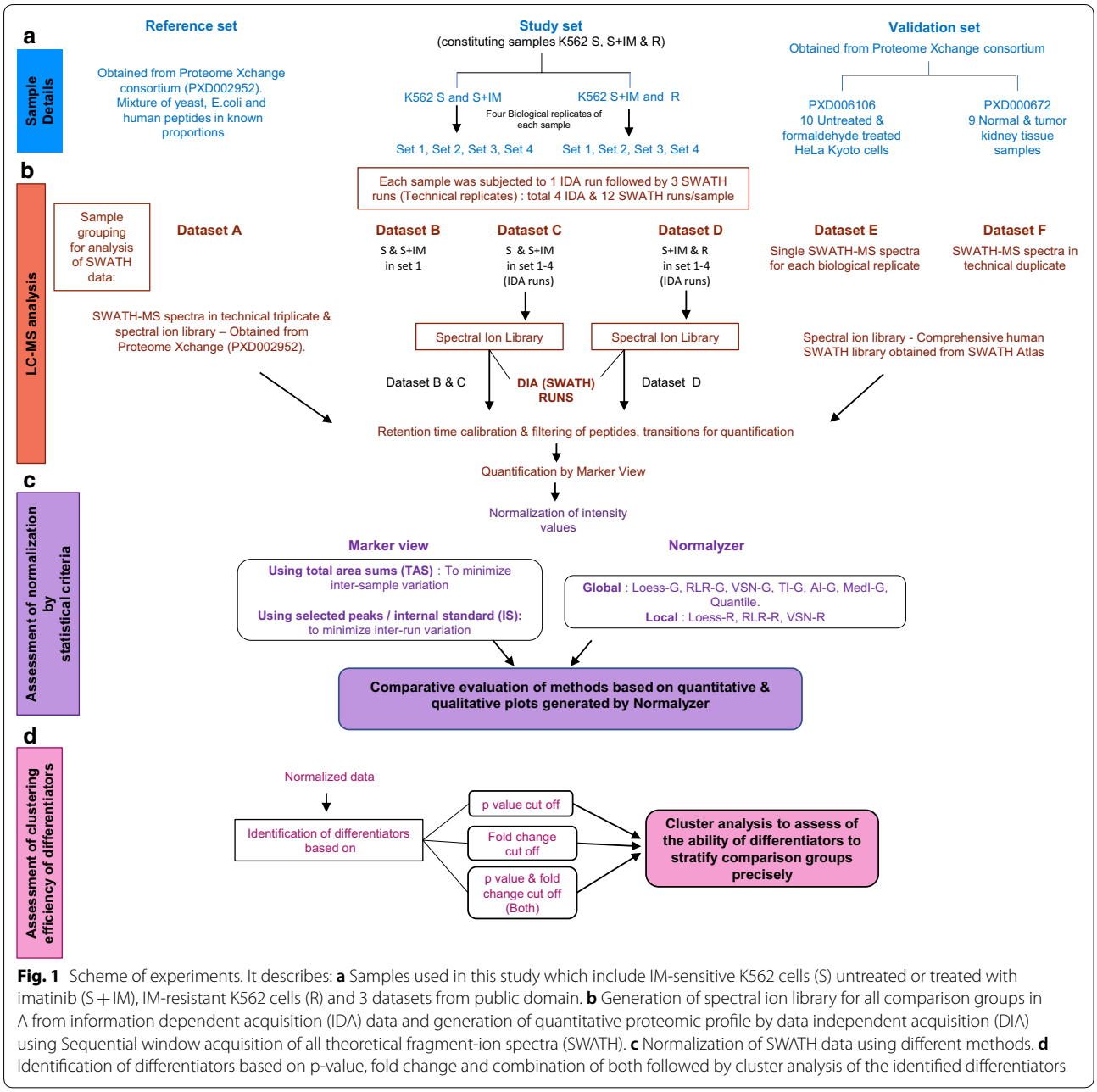


Fig. 1 Scheme of experiments. It describes: **a** Samples used in this study which include IM-sensitive K562 cells (S) untreated or treated with imatinib (S + IM), IM-resistant K562 cells (R) and 3 datasets from public domain. **b** Generation of spectral ion library for all comparison groups in A from information dependent acquisition (IDA) data and generation of quantitative proteomic profile by data independent acquisition (DIA) using Sequential window acquisition of all theoretical fragment-ion spectra (SWATH). **c** Normalization of SWATH data using different methods. **d** Identification of differentiators based on p-value, fold change and combination of both followed by cluster analysis of the identified differentiators

Details of samples for SWATH-MS analysis

Details of samples used for SWATH-MS analysis are summarized in Fig. 1a. The reference set was obtained from data published by Navarro et al. [12] wherein samples were prepared by mixing known proportions of constituent proteome (i.e. with known fold-difference in quantities). Samples with a hybrid of human, yeast and *E. coli* peptides referred to as HYE124 had differences in relative proportions of the constituent peptides and served as control (65% w/w human, 30% w/w yeast, 5% w/w *E. coli* peptides) and test (65% w/w human, 15% w/w yeast, 20% w/w *E. coli* peptides). SWATH runs of these samples in technical triplicate and their corresponding spectral ion library deposited in Proteome Xchange consortium (identifier-PXD002952), was used for SWATH data analysis.

A 'study set' was generated in our laboratory using K562, an erythroleukemic cell line (generous gift from Dr. Tadashi Nagai, Jichi Medical University, Tochigi, Japan). It was maintained in RPMI 1640 medium supplemented with 10% FBS and 1% antibiotic (Gibco, Thermo Fisher Scientific, USA). K562 harbours BCR/ABL

oncogene which encodes a constitutively active tyrosine kinase, whose activity is inhibited by the small molecular inhibitor imatinib (IM). Inhibition of BCR/ABL activity by imatinib is known to cause quantitative changes in the proteome of K562 cells [13, 14].

Thus IM-sensitive K562 cells (S) untreated or treated with imatinib (S + IM) and IM-resistant K562 cells (R) were analyzed. For S cells, treatment with imatinib was carried out at 0.75 µM concentration for 12 h, a condition observed to inhibit BCR/ABL activity without compromising on cell viability (data not shown). R cells were always maintained in medium containing 0.75 µM imatinib. SWATH-MS profiles were generated for four biological replicates of S, S + IM and R, each run-in triplicate (Fig. 1a, b).

The 'validation set' constituted SWATH data deposited by Tan et al. [15]. and Guo et al. [16] in Proteome Xchange consortium with identifiers PXD006106 and PXD000672, respectively. SWATH runs of ten biological replicates of HeLa Kyoto cells untreated (UT) and treated with formaldehyde (FA) were obtained from PXD006106 while duplicate SWATH runs of normal (N) and tumorous (T) kidney tissue samples from nine patients were obtained from PXD000672.

Preparation of K562 lysates for LC–MS analysis

To prepare whole cell lysate, 1×10^6 cells were suspended in 100 µl SDS buffer (10% glycerol, 2% SDS, 5% β-mercaptoethanol and 62.5 mM tris pH 6.8), boiled for 10 min and centrifuged at 13,000×*g* for 15 min. The supernatant was collected and acetone precipitated with 1 ml chilled acetone to remove detergents. Protein pellet thus obtained was denatured by resuspending in 6 M urea and protein concentration was determined by Bradford assay [17]. 10 µg protein was subjected to in-solution trypsin digestion. Briefly, the denatured proteins were reduced by incubating with 200 mM dithiothreitol (DTT) for 1 h at room temperature. It was followed by alkylation with 200 mM iodoacetamide (IAA) for 1 h in dark. Before trypsin digestion, urea concentration was adjusted to 0.6 M using 1 mM $CaCl_2$. In-solution digestion was carried out by adding proteomic grade trypsin (Sigma Aldrich, USA) in the ratio of 1:50 trypsin: protein (w/w) and incubated for 16 h at 37 °C. Peptides were then desalted using C18 spin columns (Pierce, Thermo Fisher Scientific, USA), dried in a speed vac and reconstituted with 0.1% Formic acid (FA) in water to get a final concentration of 0.5 µg/µl.

LC–MS/MS data acquisition for the study set

Each sample in the study set was spiked with 1 pmol/µl of digested *Escherichia coli* β-galactosidase (β-gal) peptides (Sciex, USA), before injection, which served as internal standard. The samples were then injected into Eksigent ekspertTM nano-LC 400 with cHiPLC® system, with trap column (200 µmX 0.5 mm) and analytical column (75 µmX 15 cm), both packed with 3 µl ChromXp C18 (120 Å). For reverse phase HPLC, 0.1% FA in water and 0.1% FA in acetonitrile (ACN) served as solvent A and B respectively. A gradient elution of 225 min, with increasing percentage of mobile phase B was used to elute the peptides at a flow rate of 300 nl/min. Eluate from the column was analyzed in a positive ion mode on Triple TOF 5600 + (Sciex, USA) mass spectrometer.

Each sample was subjected to 1 IDA run for spectral ion library generation followed by 3 DIA (SWATH) runs, which served as technical replicates. Thus, with four biological replicates, K562 S, S + IM and R cells had 4 IDA runs and 12 SWATH runs each (Fig. 1b). IDA mode involved a survey scan over a mass range of 350–1250 m/z and MS/MS scan over 200–1800 m/z for top 30 precursor ions with rolling collision energy, 50 mDa mass tolerance and accumulation time of 250 ms for MS and about 50 ms for MS/MS.

For DIA-SWATH acquisition, the instrument was tuned to a looped product ion mode. A sequential isolation window width of 25 m/z (with 1 m/z overlap) covering a mass range of 350–1250 m/z was set, resulting in 36 overlapping windows. The accumulation time was 50 ms for MS scan and 80 ms for MS/MS scan, thereby making a total cycle time of about 3 s. The conditions used to generate data by Navarro et al. [12]. Guo et al. [16] and that used to generate data experimentally in this study were comparable, while data generated by Tan et al. [15] used 64 variable wide precursor ion selection window. Further, samples in the reference set and validation set were spiked with indexed retention time (iRT) peptides for retention time calibration while those in the study set were spiked with *E. coli* β-gal peptides.

Generation of spectral ion library for the study set

The reference set from Navarro et al. [12] was referred to as Dataset A. The data acquired from S, S + IM and R sets were further grouped for comparison into datasets (Fig. 1b and Table 1). Only one out of the four sets of S and S + IM each, was considered as dataset B while all four together as dataset C. All four sets of S + IM and R were included in dataset D. The validation sets from Tan et al. [15] and Guo et al. [16] were referred to as dataset E and F respectively.

A common spectral ion library was generated for datasets B and C while a separate library was created for dataset D. The spectral ion library for datasets B, C and D was generated by pooling the IDA runs of the corresponding biological replicates and analysing in Protein Pilot software v4.5 (Sciex, USA) with paragon algorithm,

Table 1 Details of datasets

Datasets	Source	Constituents	Purpose—in this study
Dataset A	Pride ID—PXD002952	3 samples of 65% human, 30% yeast, 5% *E. coli* peptides (Control) 3 samples of 65% human, 15% yeast, 20% *E. coli* peptides (Test)	Reference set—a well-defined dataset with predictable quantification
Dataset B	In vitro experiments carried out in this study	3 samples of K562/S cells (Control) 3 samples K562/S + IM cells (Test)	Study set to check comparability of observations in defined (A) versus undefined (B) datasets
Dataset C	In vitro experiments carried out in this study	12 samples of K562/S cells (Control) 12 samples K562/S + IM cells (Test)	Larger dataset (C) to check the application of observations from small dataset (B)
Dataset D	In vitro experiments carried out in this study	12 samples of K562/S + IM cells (Control) 12 samples K562/R cells (Test)	Larger dataset (D) to check the consistency of observations in independent large datasets
Dataset E	Pride ID—PXD006106	10 samples of untreated HeLa Kyoto cells (Control) 10 samples of formaldehyde treated HeLa Kyoto cells (Test)	Validation set to check the consistency of observations in independent large datasets
Dataset F	Pride ID—PXD000672	18 non-tumorous kidney tissue samples (Control) 18 tumorous kidney tissue samples (Test)	Validation set to check the consistency of observations in independent large datasets

to obtain protein identities. The parameters used were as follows: Cysteine alkylation—IAA, digestion—trypsin and no special factor was chosen. The search effort was set to 'thorough ID' and false discovery rate (FDR) analysis was enabled. Proteins identified with 1% FDR were considered. The search was carried out against UniProt database (November 2016 release) containing human proteins as well as *E. coli* β-gal. The result (.group) file thus generated served as the spectral ion library. For dataset A the spectral ion library deposited by Navarro et al. [12], generated by pooling individual libraries of constituent human, yeast and *E. coli* peptides, was used. For datasets E and F comprehensive human SWATH library with about 10,000 proteins deposited in SWATH Atlas by Rosenberger et al. [18] was used.

SWATH data analysis

Spectral alignment and targeted data extraction of the swath runs of all six datasets were carried out in Peak View 2.2 software using MS/MS ALL with SWATH acquisition microapp (Sciex, USA). Proteins from spectral ion library identified with 1% FDR were first imported into Peak View 2.2 software. Retention time calibration was carried out using iRT peptides for datasets A, E and F and *E. coli* β- gal peptides for datasets B-D. Processing settings were used to filter the ion library, where up to 6 peptides per protein and 6 transitions per peptide with peptide confidence threshold of 99% and FDR of 1%, were chosen for quantification. Modified peptides were excluded from extraction. Extracted ion chromatogram (XIC) window was set to 5 min for datasets A, B, C, E, F and 15 min for dataset D with XIC width of 50 ppm. The MS/MS extracted peak areas from the filtered results were exported to Marker View software v1.3 (Sciex, USA) for quantification. The marker view output raw data file with list of proteins and their peak areas were used for further analysis.

Normalization of SWATH Data

The raw data of all datasets was processed and analyzed in Normalyzer (Fig. 1c), wherein it was log2 transformed and then normalized globally (G) or locally (R) using 10 statistical methods. Global normalization is carried out without consideration of affiliation of the sample such as replicate, control group, test group, etc. [10]. In SWATH-MS since each sample is run individually, errors can arise irrespective of their origin. Thus, in the present study global normalization methods were included. However, since the study focuses on identification of normalization method conducive to biomarker identification, retention of distinguishing features of the comparison groups was necessary while normalizing the data. This was achieved by including local normalization methods for analysis [10]. The normalization methods include locally estimated scatterplot smoothing (Loess-R, Loess-G) which assumes non-linear relationship between the bias in the data and magnitude of protein intensity; robust linear regression (RLR-R, RLR-G) which assumes that the bias in data is linearly dependent of the magnitude of the measured protein intensity; variance stabilization normalization (VSN-R, VSN-G) which aims at making the sample variances nondependent from their mean intensities and bringing the samples onto the same scale; quantile normalization which forces the distribution of the samples to be the same; total intensity (TI-G), average

intensity (AI-G) and median intensity (MedI-G) normalization methods wherein intensity of each variable is divided by sum of intensities, mean of sum of intensities, median intensities of all variables respectively [8, 10].

Marker view v1.3 along with quantitation also provides options for sample normalization using either total area sums (TAS) wherein total area of all peaks in a sample is considered or using area of the selected peaks or internal standard (IS). In this study spiked iRT peptides and trypsin digest of *E. coli* β- gal served as an internal standard for dataset A, E, F and datasets B–D respectively. In TAS as well as IS normalization, the peak areas of each sample were normalized by multiplying with its corresponding scale factor. The scale factor for TAS method was obtained by dividing the average of total area of all samples by the total area of each sample while for IS method the average area of internal standard of all samples was divided by the area of internal standard of each sample. Data normalized by the above two methods i.e. TAS and IS was log2 transformed before running through Normalyzer, to generate the evaluation report.

The normalization efficiency of all 12 methods was assessed through 'Normalyzer' quantitatively by pooled intragroup coefficient of variation (PCV) and qualitatively by relative log expression (RLE) plot as reported in earlier studies [8, 10].

Identification of differentiators from normalized data and cluster analysis

Differentiators were identified from the data of all datasets normalized by 12 methods based on the criteria of p-value, fold-change and a combination of both (Fig. 1d). To obtain p-value, log2 transformed data, normalized by different normalization methods from comparison groups were assessed by Student's t-test using IBM SPSS statistics 21. Differences in protein intensities with p-value ≤ 0.05 were considered statistically significant and chosen as differentiators. The fold change difference in protein levels was calculated from the peak area values and a cut-off of 1.5-fold change was applied. Further, the efficiency of differentiators obtained from data normalized using the 12 methods to segregate the comparison groups was assessed by cluster analysis. The peak areas of differentiators identified using p-value (≤ 0.05), fold change (1.5 fold) and combination of both were used as inputs for cluster analysis (Fig. 1d) in Genesis software v.1.8.1. Hierarchical clustering was performed with the following parameters: Agglomeration rule – Average linkage WPGMA and Calculation parameters – Cluster experiments.

Results

Identification and quantitation of proteins by SWATH-MS

In this study, each of the four biological replicates of K562 S, S + IM and R, underwent one IDA run for the generation of spectral ion library followed by three DIA runs for SWATH-MS analysis, thereby resulting in a total of 4 IDA and 12 DIA runs for K562 S, S + IM and R each. Samples with improper chromatogram were eliminated from analysis sets leaving 11 runs each in S and R in datasets C and D respectively (Fig. 1b). In dataset F, there were 2 technical replicates for each sample. Upon spectral alignment and filtering of ion library, 4404, 1450, 1757, 1808, 7057 and 5316 proteins that fulfilled the criteria (described in Methods under 'SWATH data analysis') were further used for quantification of datasets A, B, C, D, E and F respectively. Quantities of the identified proteins were further assessed for variation.

Assessment of variation in un-normalized data

The quantified log2 transformed 'un-normalized' data of each dataset was evaluated based on RLE plot, which assesses the inter- and intra-group alignment of the replicates qualitatively. In RLE plot, samples should be aligned around zero. Any deviation would indicate discrepancies in the data [10]. Among the datasets constituted of single set of samples, alignment around zero was seen in all the representative samples of dataset A (Fig. 2a) and 50% of those in dataset B (Fig. 2b). Datasets C (Fig. 2c), D (Fig. 2d) and F (Fig. 2f) comprising of multiple sets, showed considerable deviation from zero in replicates as well as between groups in RLE plots, indicating the need for normalization of SWATH-MS data.

Identification of optimum method for normalization using 'Normalyzer'

The efficiency of 12 different normalization methods to normalize datasets A–F, was assessed quantitatively and qualitatively in 'Normalyzer' using PCV and RLE plots respectively. PCV reflects the ability of a normalization method to decrease intragroup variation between technical and/or biological replicates [8]. The results indicated that, VSN-G-normalized data consistently showed lesser intra-group variation in all datasets compared to data normalized by other methods (Fig. 3I). Additionally, in datasets B–F VSN-R normalized data also reduced intra group variation. Further, qualitative assessment of the normalization methods with lowest PCV (VSN-G and VSN-R) by RLE plot indicated that only VSN-G showed good inter and intra group alignment among the replicates in all datasets (Fig. 3II). Thus, VSN-G was identified

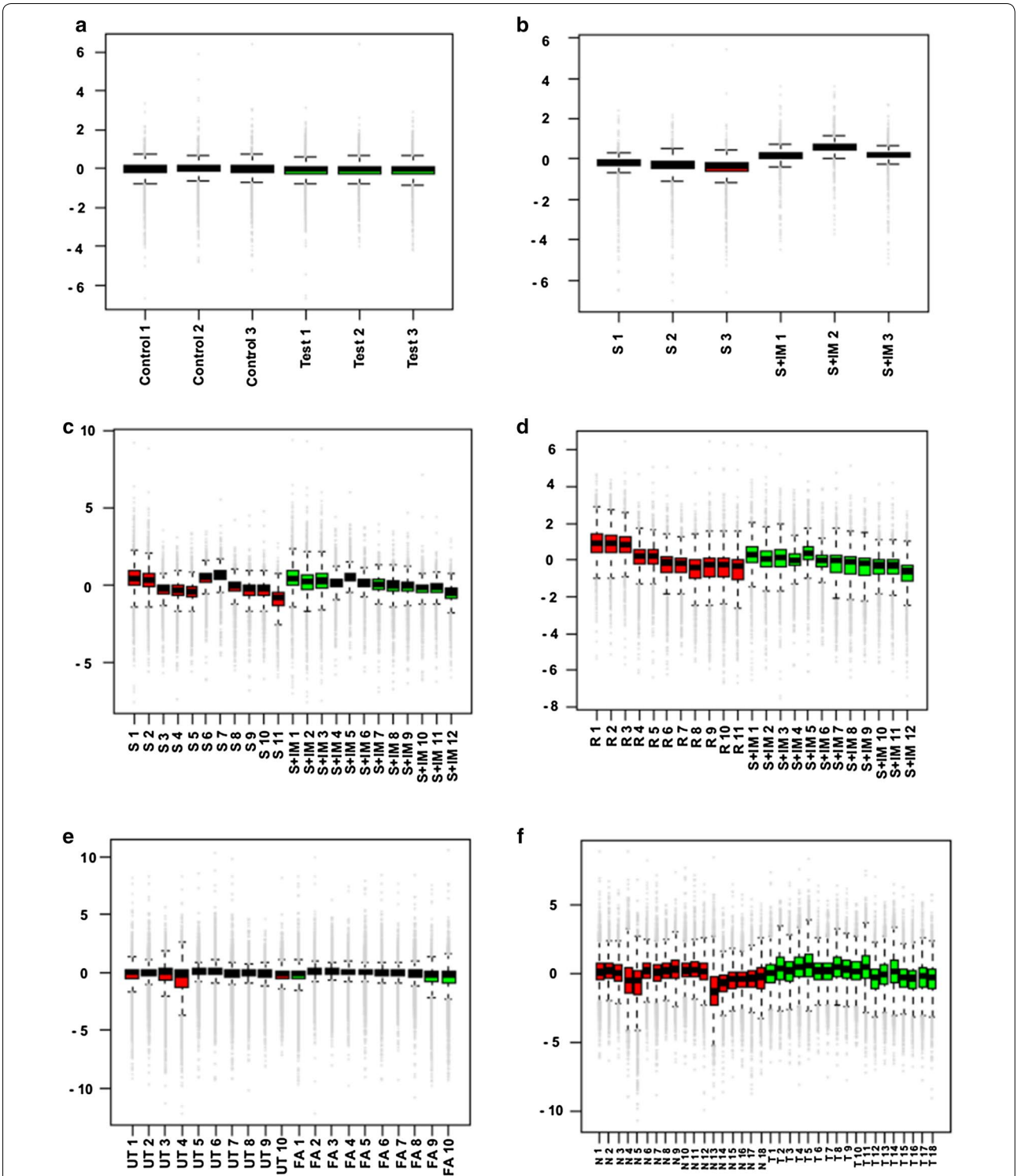


Fig. 2 Analysis of unnormalized SWATH data for datasets A–F (**a**–**f**) by RLE plot: Qualitative assessment of the spread of data shows that the test and control groups vary in their spread of values in all datasets except **a** and **e**

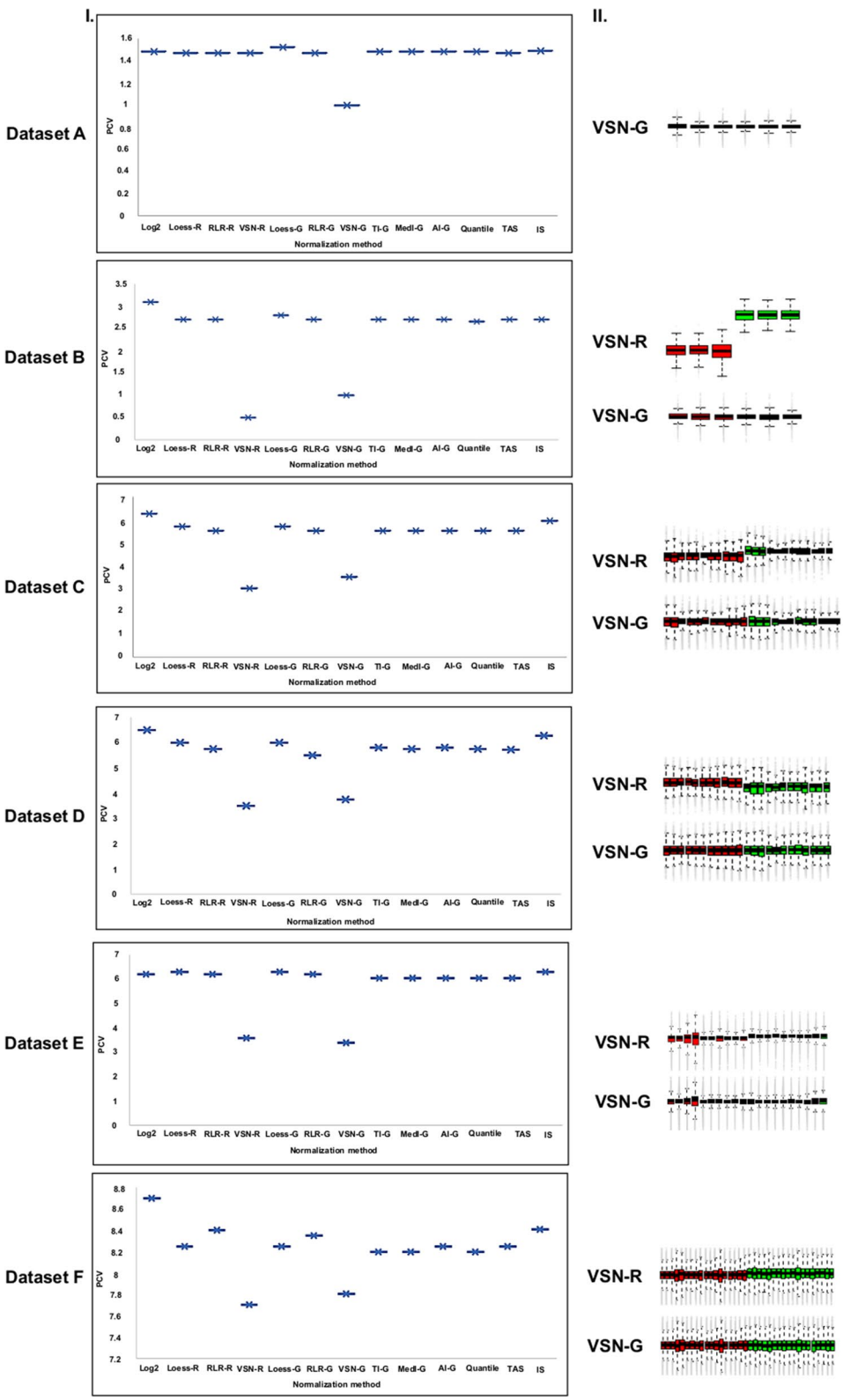


Fig. 3 **I** PCV plot: Quantitative assessment of twelve normalization methods indicates that VSN-G has less PCV in all datasets along with VSN-R in datasets B–F. **II** RLE plot: Qualitative analysis of methods with less PCV, by RLE plot revealed good inter group alignment only in VSN-G in all datasets

as the optimal normalization method using 'Normalyzer' based evaluation.

Assessment of VSN-G normalized data by cluster analysis

Differentiators identified from data normalized by VSN-G method based on p-value, fold change and a combination of both were subjected to cluster analysis. Differentiators identified by all three criteria could segregate the comparison groups appropriately in datasets A, B and D but not in dataset C, E and F (Fig. 4). Though VSN-G was identified as optimal normalization method based on PCV and RLE plots, the differentiators identified did not show consistent efficiency in clustering. In order to understand the contribution of VSN-G normalization to improper clustering of datasets C, E and F, differentiators identified by all three criteria, from data normalized with the remaining eleven methods were assessed for their clustering ability. The aim was to detect if any other normalization method could improve segregation of datasets C, E and F while retaining the efficient segregation of datasets A, B and D in VSN-G normalized data.

Assessment of data normalized by methods other than VSN-G by cluster analysis

As observed in VSN-G normalized data, clusters obtained from data normalized with the remaining eleven methods yielded improper clustering in datasets E and F. Thus the improper features of clusters i.e. formation of separate cluster by a few normal samples in datasets E and F; segregation of a pair of normal samples (N9 and N18) with tumor samples in dataset F was taken as a consistent feature across normalized data for these two datasets and was not applied to eliminate a cluster as imprecise. While retaining these features, clear segregation of the remaining control and test samples was considered as acceptable clustering efficiency of datasets E and F. Based on this relaxed criteria, it is seen in Fig. 5 (Detailed dendrograms for cluster analysis is given in Additional file 1) and Table 2 that differentiators identified based on p-value efficiently segregate the comparison groups for data normalized by majority of methods. On the other hand, differentiators identified based on fold change could not segregate the comparison groups in majority of the datasets. The ability of differentiators obtained from the combination of p-value and fold-change to segregate sets therefore could be attributed to the influence of p-value. Based on the above experimental evidence p-value is chosen as the criteria for differentiator identification in this study.

Of the 11 normalization methods assessed, differentiators identified based on p-value from data normalized by 3 methods (Loess-R, TI-G and AI-G) segregated the comparison groups precisely in all datasets (Fig. 5). These were further evaluated using more stringent criteria to identify the most optimal method for biomarker discovery. The criteria was to sub-cluster the technical replicates, of control and test groups, belonging to each biological replicate precisely in datasets C, D and F. Dataset E was not subclustered as each sample was run only once [15]. A scoring system was used to achieve this, wherein the ability to segregate control and test groups was given a score of 2. In dataset F, for every control which segregated separately from the major control cluster, a negative score of 1 was given. Thereafter for every correct subgrouping of the technical replicates of control and test, a score of 1 was given. The total score was calculated as score for precise clustering (2) + score of −1 for each control which clustered separately from the major control cluster in dataset F (not applicable to other datasets) + score for co-segregation of technical replicates in test and control (1)(Fig. 6).

As mentioned earlier, the efficiency of biomarkers lies in their ability to accurately stratify the heterogenous groups in a given population. It is evident from Fig. 6 that differentiators obtained from Loess-R normalized data could not only stratify the comparison groups precisely, but also had maximum sub-stratification score in the three large datasets assessed, thereby indicating its suitability for biomarker discovery.

Discussion

This study has addressed two previously unattended issues in analysis of quantitative SWATH-MS data, especially relevant to clinical proteomics—(i) experimental demonstration of ideal method of data normalization which does not diminish the vital features of the data necessary for segregation of comparison groups (ii) experimental verification of criteria for identification of differentiators. Carefully chosen sets of samples, mimicking the biological and experimental variations which can influence the data were included in the study. The reference set from public domain (dataset A) represented a quantitatively defined set wherein the differences in relative proportions of the constituents between samples made fold-differences in protein quantities predictable. The study set (datasets B–D) on the other hand represented the heterogeneity inherent to biological samples as in S cells and those contributed by extraneous manipulations such as treatment with imatinib in S + IM and R cells. The validation set (dataset E and F) were analyzed to confirm the findings obtained in the earlier sets. While analysis of single sets in dataset A and B allowed to evaluate differences between quantitatively defined (dataset A) and undefined set (dataset B), multiple sets in dataset C, D, E and F allowed for evaluation of the consistency of observations within quantitatively undefined

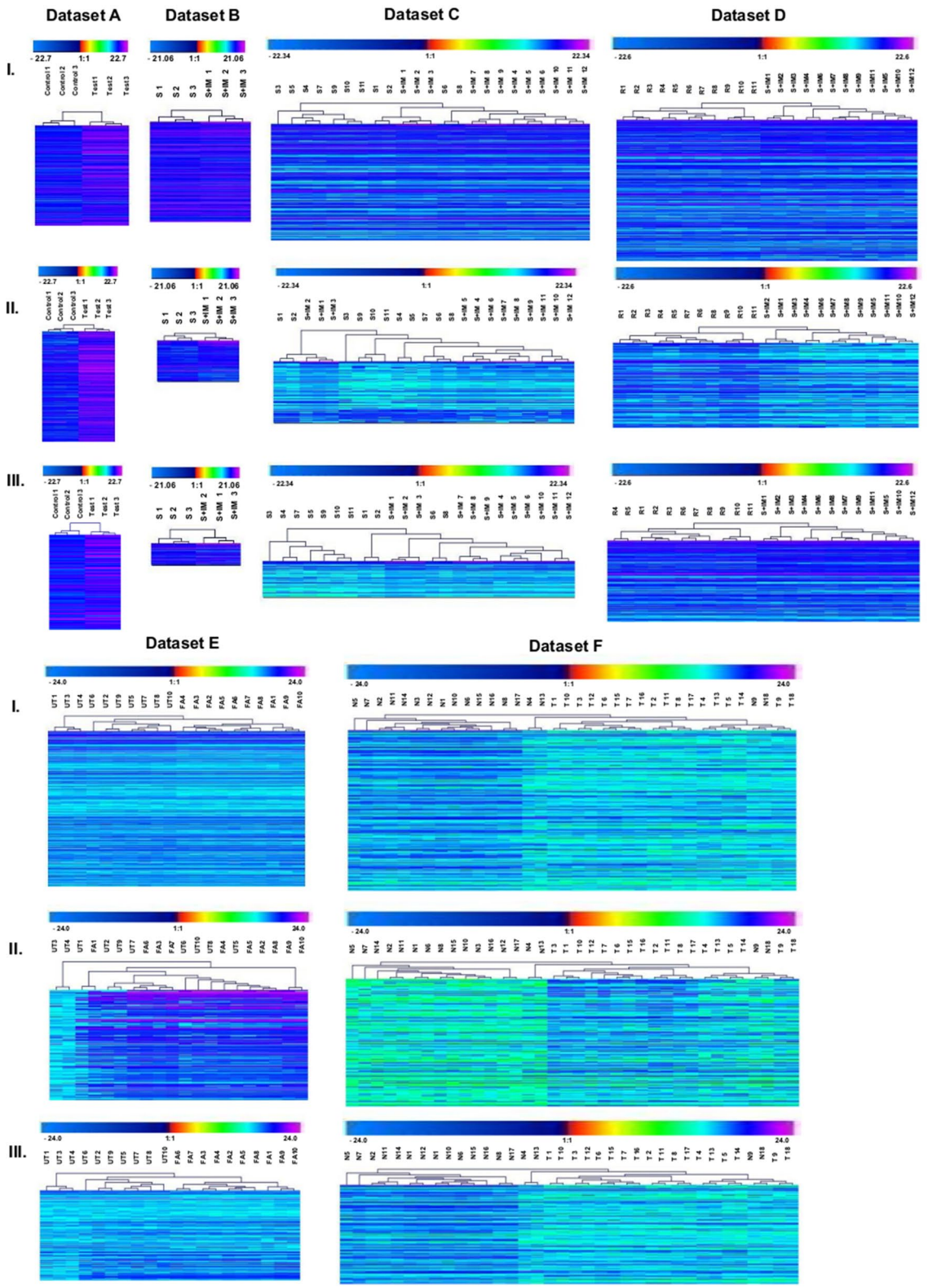


Fig. 4 Hierarchical clustering of differentiators obtained from VSN-G normalized data based on **I**- p-value, **II**- fold change, **III**- p- value together with fold change

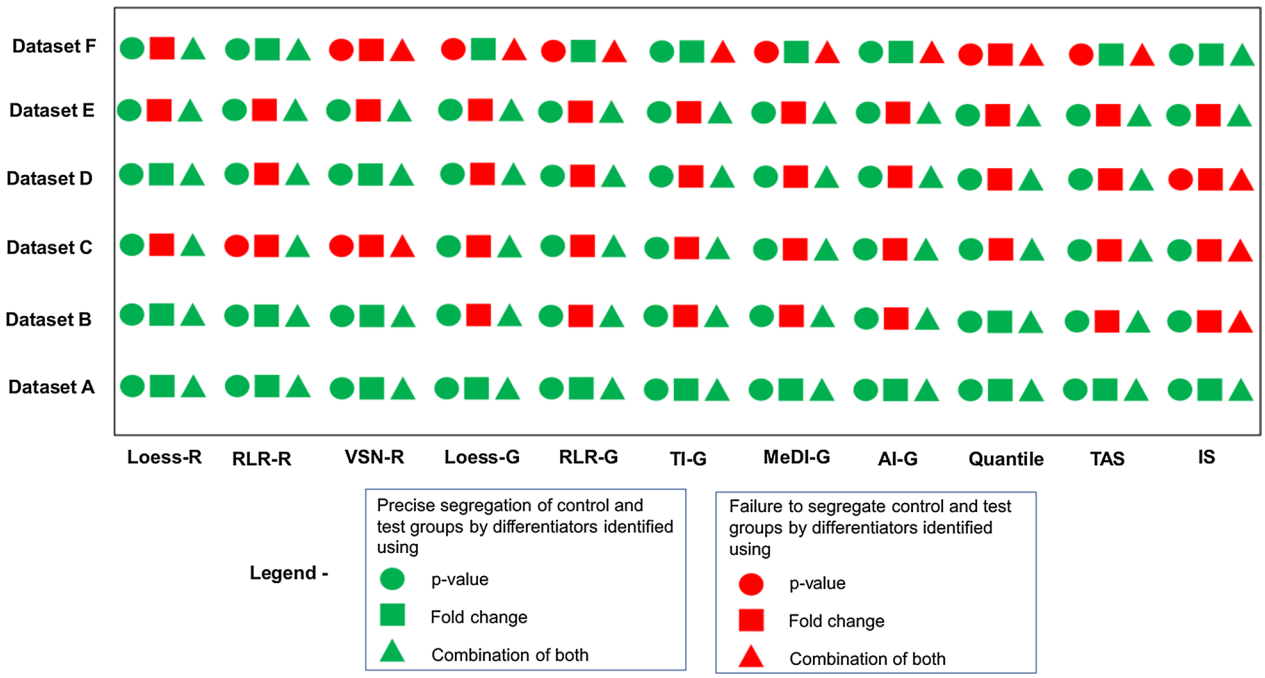


Fig. 5 Ability of differentiators to cluster the study groups distinctly

Table 2 Clustering efficiency of differentiators identified based on p-value, fold change and combination of both, from data normalized by 12 methods

Datasets	Clustering efficiency		
	p-value	Fold change	Both
A	100% (12/12 methods)	100% (12/12 methods)	100% (12/12 methods)
B	100% (12/12 methods)	42% (5/12 methods)	92% (11/12 methods)
C	75% (9/12 methods)	0% (0/12 methods)	75% (9/12 methods)
D	92% (11/12 methods)	25% (3/12 methods)	92% (11/12 methods)
E	100% (12/12 methods)	0% (0/12 methods)	100% (12/12 methods)
F	42% (5/12 methods)	66.6% (8/12 methods)	25% (3/12 methods)

sets. Further the reference, study and validation sets differed in the depth of spectral ion library, the choice of calibrants for retention time calibration as well as peak intensities. They are the prototypes of variation of methodologies observed in the reported literature. Analysis of these samples was designed to identify a strategy for normalization of SWATH-MS data which is applicable universally.

The differences in datasets throw light on certain valuable aspects of experimental design. Detection of greater number of proteins in dataset A, E and F as compared to B, C and D could be attributed to deeper spectral ion library. For dataset A, library was generated by pooling individual libraries of constituent human, yeast and *E. coli* peptides while for datasets E and F extensive human protein library was used. It could also be due to use of iRT peptides for retention time calibration in dataset A, E and F which allows high-quality spectral library generation. Retention time calibration of datasets B, C and D had been carried out using spiked peptides of *E. coli* beta galactosidase which span a limited range of retention times. This represents the studies where retention time calibration has been carried out using highly conserved and abundant endogenous peptides or spike-in peptides other than iRT [19, 20].

The extent of variation among un-normalized datasets, when assessed by RLE plot, showed a progressive increase from dataset A to D and F (Fig. 2). This increase

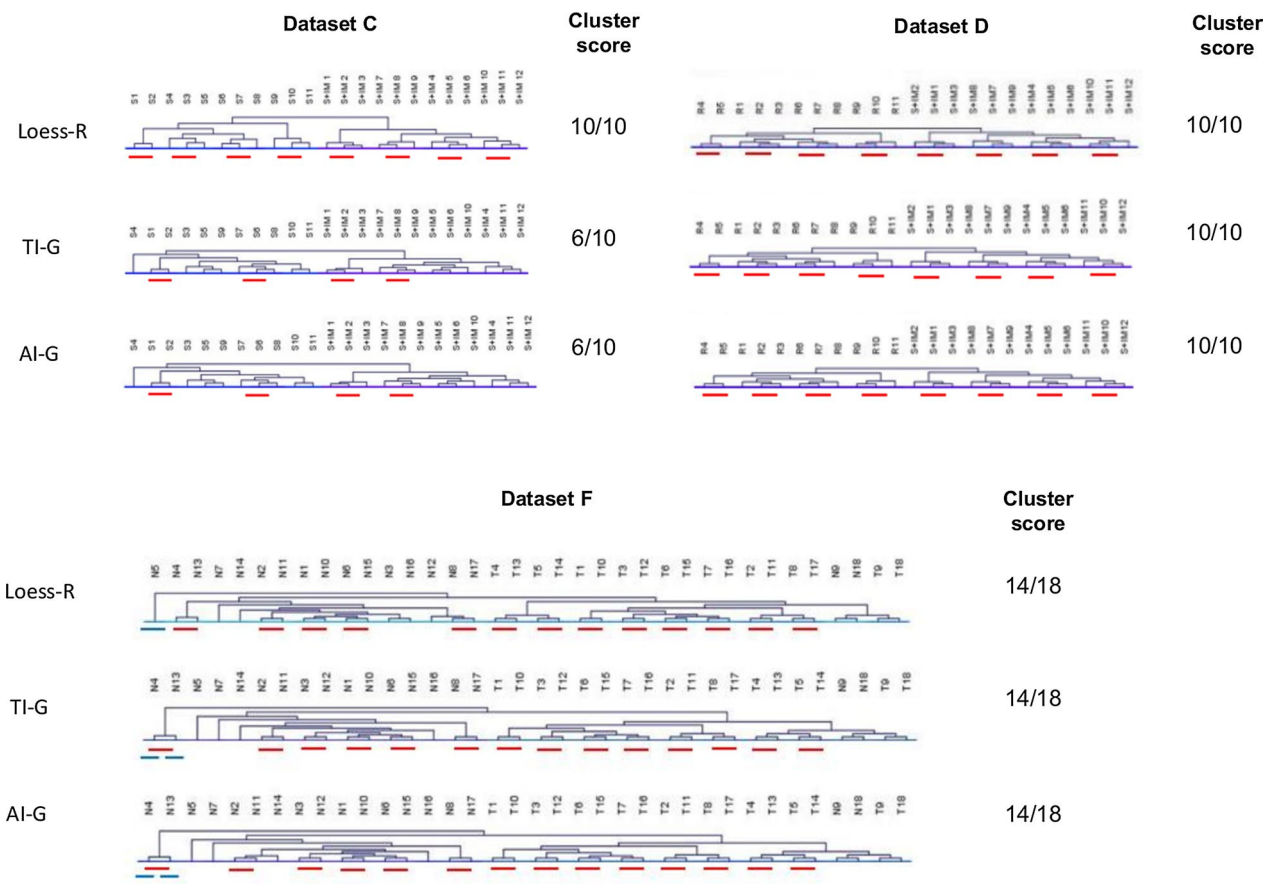


Fig. 6 Evaluation of efficiency of p-value based differentiators to sub-stratify the technical replicates. Proper clustering of test and control groups is given a score of 2 and proper sub-clustering of technical replicates of each set indicated by red line, is given a score of 1. In dataset F, a score of -1 is given to each control which formed a cluster outside the major control or test cluster, indicated by blue line

in variation could be attributed to increase in sample heterogeneity, as dataset A, generated by addition of defined proportions of constituents, was less heterogenous and had more precise quantification. Dataset B on the other hand, was not a defined set and thus would exhibit variations inherent to any biological system. In datasets C and D, the probability of variation increased as the heterogeneity increased due to inclusion of greater number of samples. Dataset E which also involved cell lines as in datasets C and D, showed least variation which reflects precision in experimentation but is not commonly observed due to experimental errors. Dataset F included human samples which are inherently heterogenous. Such variations are a commonplace in clinical samples and reflect in the assessment of un-normalized data based on RLE plot. These observations emphasized the need for normalization of SWATH-MS data.

In most of the previously reported SWATH-MS studies, data has been normalized by TAS [21–29], median [15, 30–32], TI [33–36], quantile [37–39] and IS [40, 41] methods. In this study, to identify the optimum normalization method, datasets A–F were normalized using 10 normalization methods from Normalyzer and 2 methods from Marker View software, which included the above mentioned methods used in previous studies. Their normalization efficiency in Normalyzer was evaluated based on PCV and RLE plots. VSN-G was found to efficiently normalize not only dataset A with minimum variation but also datasets B-F with considerable variation (Fig. 3). This indicates that VSN-G could have a broad applicability for normalization of SWATH-MS data. VSN normalization has also been reported to efficiently normalize data generated by DIA using LTQ orbitrap [8]. Considering wide use of SWATH-MS for biomarker identification, the utility of VSN-G normalized data for biomarker discovery was assessed based on its ability to yield differentiators which segregate the comparison groups precisely.

Differentiators could be identified by comparing quantities of proteins in comparison groups based on p value, fold change or combination of both. PubMed search

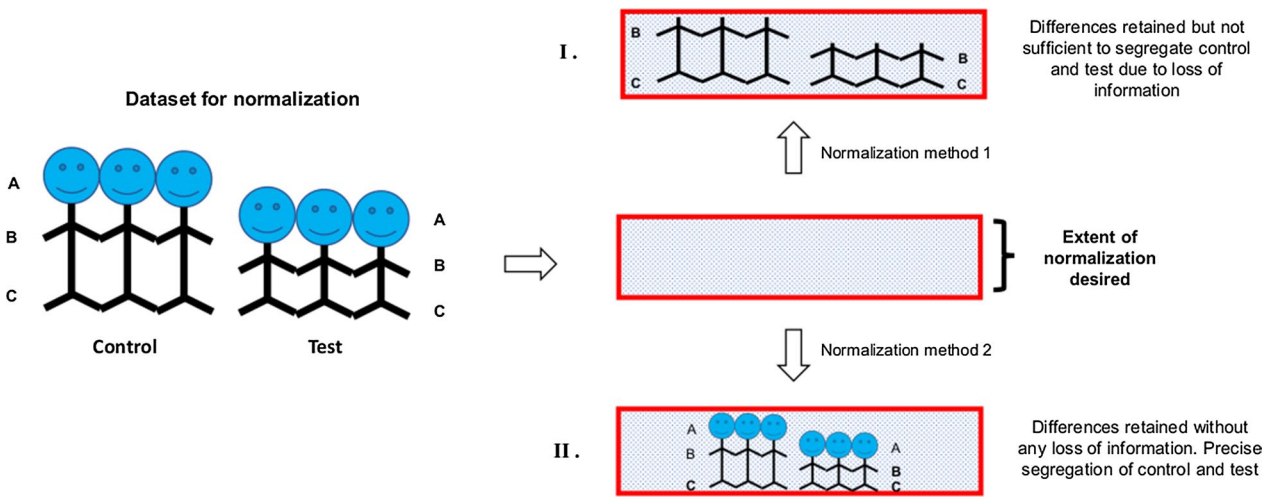


Fig. 7 Should application dictate the choice of normalization method for SWATH-MS data?

revealed that in over 134 human SWATH related publications in a span of 5 years (2012-Mar 2018), 45 studies (33.6%) were aimed at identifying differentiators between control and test groups. Out of these 45 studies, we observed that 14 studies (31.1%) used statistical significance (p-value), 6 studies (13.3%) used fold change and 25 studies (55.6%) used both (p value and fold change) as criteria to identify differentiators. While experimental evidence for either choosing the criteria for differentiator identification (p-value, fold change or combination of both) [42, 43] or their cut-off values [44] is available for transcriptomic data, no such studies are reported for MS data. Our study provides the first experimental evaluation of choice of criteria to identify differentiators from SWATH-MS data for biomarker discovery based on their ability to segregate comparison groups- an essential feature of biomarkers, by cluster analysis. Cluster analysis revealed that differentiators identified based on p-value, from data normalized by 12 methods, could segregate the comparison groups with maximum efficiency in 5 out of 6 datasets (Fig. 5). Hence p-value was chosen as the criteria for differentiator identification in this study.

VSN-G, though was identified as optimal normalization method based on PCV and RLE plots, the differentiators identified did not efficiently cluster the comparison groups in all datasets (Fig. 4), thereby raising question on its suitability for biomarker discovery. Loess-R, a method ranked lower in 'Normalyzer' based evaluation, on the other hand, yielded differentiators with maximum clustering as well as sub-clustering efficiency in all datasets assessed (Fig. 6), thereby making it suitable for biomarker discovery by SWATH-MS. This may be due to the differences in the assumptions made for normalization by these methods. The perceived treatment of data is depicted in Fig. 7. VSN-G aims at making the sample variances non-dependent on their mean intensity thereby bringing the samples onto the same scale. This assumption remarkably reduces the intensity differences between samples so as to achieve optimum normalization. However, the reduction in intensity differences is not conducive to identification of differentiators and in turn segregation of comparison groups (Fig. 7I) Loess normalization on the other hand, probably retains the differences between intragroup protein intensities by assuming non-linear relationship between biases in the data and the magnitude of protein intensity—a feature essential for segregation of comparison groups (Fig. 7II).

We thus propose that apart from statistically recommended criteria for evaluation of methods for normalization, a biologically relevant criteria like precise stratification of data should be assessed before a normalization method is used for biomarker identification from SWATH-MS data.

Conclusion

This study for the first time has identified VSN-G as method for optimum normalization of SWATH-MS data based on statistical criteria. Acknowledging the extensive use of this technology for biomarker discovery this study has also identified the normalization strategies conducive to this application. In the process, p-value based identification of differentiators has been demonstrated to be most suitable for biomarker discovery from SWATH-MS data. While VSN-G normalization was not found conducive to biomarker discovery in this study, Loess-R normalization was observed to retain features of the data

necessary to yield differentiators which could segregate the comparison groups efficiently. The probable effect of two normalization methods on the data which are responsible for these observations are depicted in Fig. 7. The study has thus demonstrated the need to base the choice of normalization method on the application of the data.

Additional file

Additional file 1. Dendrograms representing hierarchical clustering of control and test groups based on differentiators obtained by p-value, fold change or both in both datasets.

Abbreviations

LC–MS: liquid chromatography Mass spectrometry; SWATH: Sequential window acquisition of all theoretical fragment ion spectra; DIA: data independent acquisition; IDA: information/data dependent acquisition; IM: imatinib; K562 S: imatinib sensitive cells; K562 S + IM: sensitive cells treated with imatinib; K562 R: cells resistant to imatinib; DTT: dithiothreitol; IAA: iodoacetamide; FA: formic acid; β-gal: *E. coli* beta galatosidase; ACN: acetonitrile; iRT: indexed retention time; FDR: false discovery rate; XIC: extracted ion chromatogram; R: local normalization; G: global normalization; RLR: robust linear regression; VSN: variance stabilizing normalization; TI: total intensity; AI: average intensity; MedI-G: median intensity; TAS: total area sums; IS: internal standard; PCV: pooled intragroup coefficient of variation; RLE: relative log expression.

Acknowledgements

We acknowledge the support of mass spectrometry facility, ACTREC, Navi Mumbai.

Authors' contributions

MN has carried out experimental work, data analysis and initial writing of manuscript and has incorporated subsequent changes suggested by co-authors, AC has carried out data processing on 'Normalyzer'. AB and SK have carried out statistical analysis and RG has conceptualized the project and executed the same. All authors read and approved the final manuscript.

Funding

This work was funded by Intramural Research Grant of Tata Memorial Centre.

Availability of data and materials

The proteomics data generated for the study set in the current study have been deposited to the Proteome Xchange Consortium via the PRIDE [45] partner repository with the dataset identifier PXD009686. The reference dataset deposited by Navarro et al. [12], used for analysis in this study, has been obtained from Proteome Xchange Consortium (Identifier—PXD002952).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Kharghar, Navi Mumbai 410210, India. [2] BARC Training School Complex, Homi Bhabha National Institute, Anushakti Nagar, Mumbai 400094, India. [3] Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden. [4] Section of Biostatistics, Centre for Cancer Epidemiology, Tata Memorial Centre, Kharghar, Navi Mumbai 410210, India.

Received: 28 March 2019 Accepted: 24 May 2019
Published online: 31 May 2019

References

1. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool opportunities and potential limitations. Mol Cell Proteomics. 2004;3(4):367–78.
2. Luo Y, Mok TS, Lin X, Zhang W, Cui Y, Guo J, Chen X, Zhang T, Wang T. SWATH-based proteomics identified carbonic anhydrase 2 as a potential diagnosis biomarker for nasopharyngeal carcinoma. Sci Rep. 2017;7:41191.
3. Ralhan R, DeSouza LV, Matta A, Tripathi SC, Ghanny S, Gupta SD, Bahadur S, Siu KM. Discovery and verification of head-and-neck cancer biomarkers by differential protein expression analysis using iTRAQ labeling, multidimensional liquid chromatography, and tandem mass spectrometry. Mol Cell Proteomics. 2008;7(6):1162–73.
4. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics. 2012;11(6):O111–016717.
5. Domon B, Aebersold R. Options and considerations when selecting a quantitative proteomics strategy. Nat Biotechnol. 2010;28(7):710–21.
6. Karpievitch YV, Dabney AR, Smith RD. Normalization and missing value imputation for label-free LC–MS analysis. BMC Bioinform. 2012;13(16):S5.
7. Callister SJ, Barry RC, Adkins JN, Johnson ET, Qian W-J, Webb-Robertson B-JM, Smith RD, Lipton MS. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. J Proteome Res. 2006;5(2):277–86.
8. Valikangas T, Suomi T, Elo LL. A systematic evaluation of normalization methods in quantitative label-free proteomics. Brief Bioinform. 2018;19(1):1–11.
9. Kultima K, Nilsson A, Scholz B, Rossbach UL, Fälth M, Andrén PE. Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides. Mol Cell Proteomics. 2009;8(10):2285–95.
10. Chawade A, Alexandersson E, Levander F. Normalyzer: a tool for rapid evaluation of normalization methods for omics data sets. J Proteome Res. 2014;13(6):3114–20.
11. Fu J, Tang J, Wang Y, Cui X, Yang Q, Hong J, Li X, Li S, Chen Y, Xue W, et al. Discovery of the consistently well-performed analysis chain for SWATH-MS based pharmacoproteomic quantification. Front Pharmacol. 2018;9:681.
12. Navarro P, Kuharev J, Gillet LC, Bernhardt OM, MacLean B, Rost HL, Tate SA, Tsou CC, Reiter L, Distler U, et al. A multicenter study benchmarks software tools for label-free proteome quantification. Nat Biotechnol. 2016;34(11):1130–6.
13. Xiong L, Zhang J, Yuan B, Dong X, Jiang X, Wang Y. Global proteome quantification for discovering imatinib-induced perturbation of multiple biological pathways in K562 human chronic myeloid leukemia cells. J Proteome Res. 2010;9(11):6007–15.
14. Arvaniti K, Papadioti A, Kinigopoulou M, Theodorou V, Skobridis K, Tsiotis G. Proteome changes induced by imatinib and novel imatinib derivatives in K562 human chronic myeloid leukemia cells. Proteomes. 2014;2(3):363–81.
15. Tan SLW, Chadha S, Liu Y, Gabasova E, Perera D, Ahmed K, Constantinou S, Renaudin X, Lee M, Aebersold R, et al. A class of environmental and endogenous toxins induces BRCA2 haploinsufficiency and genome instability. Cell. 2017;169(6):1105–18.
16. Guo T, Kouvonen P, Koh CC, Gillet LC, Wolski WE, Rost HL. Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. Nat Med. 2015;21(4):407–13.
17. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72:248–54.
18. Rosenberger G, Koh CC, Guo T, Rost HL, Kouvonen P, Collins BC, Heusel M, Liu Y, Caron E, Vichalkovski A, et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. Sci Data. 2014;1:140031.
19. Parker SJ, Rost H, Rosenberger G, Collins BC, Malmstrom L, Amodei D, Venkatraman V, Raedschelders K, Van Eyk JE, Aebersold R. Identification

of a set of conserved eukaryotic internal retention time standards for data-independent acquisition mass spectrometry. Mol Cell Proteomics. 2015;14(10):2800–13.
20. Hou G, Lou X, Sun Y, Xu S, Zi J, Wang Q, Zhou B, Han B, Wu L, Zhao X, et al. Biomarker discovery and verification of esophageal squamous cell carcinoma using integration of SWATH/MRM. J Proteome Res. 2015;14(9):3793–803.
21. Ortea I, Rodriguez-Ariza A, Chicano-Galvez E, Arenas Vacas MS, Jurado Gamez B. Discovery of potential protein biomarkers of lung adenocarcinoma in bronchoalveolar lavage fluid by SWATH MS data-independent acquisition and targeted data extraction. J Proteomics. 2016;138:106–14.
22. Chaudhari PR, Charles SE, D'Souza ZC, Vaidya MM. Hemidesmosomal linker proteins regulate cell motility, invasion and tumorigenicity in oral squamous cell carcinoma derived cells. Exp Cell Res. 2017;360(2):125–37.
23. Arya S, Emri E, Synowsky SA, Shirran SL, Barzegar-Befroei N, Peto T, Botting CH, Lengyel I, Stewart AJ. Quantitative analysis of hydroxyapatite-binding plasma proteins in genotyped individuals with late-stage age-related macular degeneration. Exp Eye Res. 2018;172:21–9.
24. Batkulwar K, Godbole R, Banarjee R, Kassaar O, Williams RJ, Kulkarni MJ. Advanced glycation end products modulate amyloidogenic APP processing and tau phosphorylation: a mechanistic link between glycation and the development of Alzheimer's disease. 2018;9(5):988–1000.
25. Schillaci O, Fontana S. Exosomes from metastatic cancer cells transfer amoeboid phenotype to non-metastatic cells and increase endothelial permeability: their emerging role in tumor heterogeneity. Sci Rep. 2017;7(1):4711.
26. Lewandowska AE, Macur K, Czaplewska P, Liss J, Lukaszuk K, Oldziej S. Qualitative and quantitative analysis of proteome and peptidome of human follicular fluid using multiple samples from single donor with LC–MS and SWATH methodology. J Proteome Res. 2017;16(8):3053–67.
27. Lin Q, Lim HS, Lin HL, Tan HT, Lim TK, Cheong WK, Cheah PY, Tang CL, Chow PK, Chung MC. Analysis of colorectal cancer glyco-secretome identifies laminin beta-1 (LAMB1) as a potential serological biomarker for colorectal cancer. Proteomics. 2015;15(22):3905–20.
28. Korwar AM, Vannuruswamy G, Jagadeeshaprasad MG, Jayaramaiah RH, Bhat S, Regin BS, Ramaswamy S, Giri AP, Mohan V, Balasubramanyam M, et al. Development of diagnostic fragment ion library for glycated peptides of human serum albumin: targeted quantification in prediabetic, diabetic, and microalbuminuria plasma by parallel reaction monitoring, SWATH, and MSE. Mol Cell Proteomics. 2015;14(8):2150–9.
29. Tan HT, Lim TK, Richards AM, Kofidis T, Teoh KL, Ling LH, Chung MC. Unravelling the proteome of degenerative human mitral valves. Proteomics. 2015;15(17):2934–44.
30. Netherton JK, Hetherington L, Ogle RA, Velkov T, Baker MA. Proteomic analysis of good- and poor-quality human sperm demonstrates that several proteins are routinely aberrantly regulated. Biol Reprod. 2018;99(2):395–408.
31. Nemeth J, Vongrad V, Metzner KJ, Strouvelle VP, Weber R, Pedrioli P, Aebersold R, Gunthard HF, Collins BC. In vivo and in vitro proteome analysis of human immunodeficiency virus (HIV)-1-infected, human CD4+ T cells. Mol Cell Proteomics. 2017;16(4 suppl 1):S108–23.
32. Chang RY, Etheridge N, Nouwens AS, Dodd PR. SWATH analysis of the synaptic proteome in Alzheimer's disease. Neurochem Int. 2015;87:1–12.
33. Blattmann P, Henriques D, Zimmermann M, Frommelt F, Sauer U, Saez-Rodriguez J, Aebersold R. Systems pharmacology dissection of cholesterol regulation reveals determinants of large pharmacodynamic variability between cell lines. Cell systems. 2017;5(6):604–19.
34. Zhang X, Walsh T, Atherton JJ, Kostner K, Schulz B, Punyadeera C. Identification and validation of a salivary protein panel to detect heart failure early. Theranostics. 2017;7(18):4350–8.
35. Bjelosevic S, Pascovici D, Ping H, Karlaftis V, Zaw T, Song X, Molloy MP, Monagle P, Ignjatovic V. Quantitative age-specific variability of plasma proteins in healthy neonates, children and adults. Mol Cell Proteomics. 2017;16(5):924–35.
36. Liu Y, Chen J, Sethi A, Li QK, Chen L, Collins B, Gillet LC, Wollscheid B, Zhang H, Aebersold R. Glycoproteomic analysis of prostate cancer tissues by SWATH mass spectrometry discovers *N*-acylethanolamine acid amidase and protein tyrosine kinase 7 as signatures for tumor aggressiveness. Mol Cell Proteomics. 2014;13(7):1753–68.
37. Liu Y, Borel C, Li L, Muller T, Williams EG, Germain PL, Buljan M, Sajic T, Boersema PJ, Shao W, et al. Systematic proteome and proteostasis profiling in human Trisomy 21 fibroblast cells. Nat Commun. 2017;8(1):1212.
38. Chen L, Li J, Guo T, Ghosh S, Koh SK, Tian D, Zhang L, Jia D, Beuerman RW, Aebersold R, et al. Global metabonomic and proteomic analysis of human conjunctival epithelial cells (IOBA-NHC) in response to hyperosmotic stress. J Proteome Res. 2015;14(9):3982–95.
39. Orellana CA, Marcellin E, Schulz BL, Nouwens AS, Gray PP, Nielsen LK. High-antibody-producing Chinese hamster ovary cells up-regulate intracellular protein transport and glutathione synthesis. J Proteome Res. 2015;14(2):609–18.
40. Martins-Marques T, Anjo SI, Pereira P, Manadas B, Girao H. Interacting network of the gap junction (GJ) Protein Connexin43 (Cx43) is modulated by ischemia and reperfusion in the heart. Mol Cell Proteomics. 2015;14(11):3040–55.
41. Villeneuve LM, Stauch KL, Fox HS. Proteomic analysis of the mitochondria from embryonic and postnatal rat brains reveals response to developmental changes in energy demands. J Proteomics. 2014;109:228–39.
42. McCarthy DJ, Smyth GK. Testing significance relative to a fold-change threshold is a TREAT. Bioinformatics. 2009;25(6):765–71.
43. Xiao Y, Hsiao T-H, Suresh U, Chen H-IH, Wu X, Wolf SE, Chen YA. Novel significance score for gene selection and ranking. Bioinformatics. 2014;30(6):801–7.
44. Dalman MR, Deeter A, Nimishakavi G, Duan Z-H. Fold change and p-value cutoffs significantly alter microarray interpretations. BMC Bioinformatics. 2012;13(2):S11.
45. Vizcaíno JA, Csordas A, del Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, et al. 2016 update of the PRIDE database and its related tools. Nucleic acids Res. 2016;44(1):447–56.
46. Loke MF, Ng CG, Vilashni Y, Lim J, Ho B. Understanding the dimorphic lifestyles of human gastric pathogen Helicobacter pylori using the SWATH-based proteomics approach. Sci Rep. 2016;6:26784.
47. Lewandowska AE, Macur K, Czaplewska P, Liss J, Lukaszuk K, Oldziej S. Human follicular fluid proteomic and peptidomic composition quantitative studies by SWATH-MS methodology. Applicability of high pH RP-HPLC fractionation. J Proteomics. 2018;191:131–42.

Publisher's Note