

A brief report of 6thworkshop on CCPM

CCPM-6 “Conversation with Experimentalists”

February 24, 2016

This workshop “CCPM-6 February2016”consisting of one day workshops at two places IIIT Hyderabad and JNU Delhi on 24th February, 2016 both online and offline, was organized to familiarized the progressive prototype of the LIMS platform to its potential users.

Announcement and Invitations with program outline were sent well in advance to most of the PI’s and Experimentalist. Reminder mails were also sent (**Appendix 1 a – d**). Announcements and workshop details were also displayed on the CCPM portal

Each potential participant was contacted and given an option to choose a mode (online / offline) as well as place (IIIT Hyderabad/ JNU Delhi) of their convenience for workshop participation through the workshop portal.

A total of 66 participants, including four Faculty members and several research group members of PIs, registered at both venues (JNU and IIIT-H) combined. Six registrants attended online, whereas all others attended the workshop physically. All the researchers attended both the morning as well as the afternoon sessions. Among morning session attendees there were several students who were exposed to the area of Metabolomics and to the CCPM portal.

Program Details

Venue B6 – 309 @ IIIT-H

Inaugural Session

10:00 – 10:10 AM Welcome addresses

Prof. P J Narayanan (Director, IIIT-H)

Prof. Vasudeva Varma (Dean R&D, IIIT-H)

Morning Session

10:10 – 10:30 AM Introduction to CCPM version 3.4

- Prof. Abhijit Mitra

10:30 – 11:15 AM Conversation with Experimentalist

- Coordinated by CCPM Team

11:15 – 11:30 AM Tea Break

11:30 – 12:30 PM Lecture by Prof. RP Sharma on

“Metabolic shifts during trans-differentiation of tomato plant”

12:45 – 1:45 PM Lunch

Afternoon Session

1:45 – 3:15 PM Tutorial Presentation

Dr. Sangeeta Kumari

Demo related to functional features of CCPM v3.4

(a) Published projects (by Dr. Sangeeta Kumari)

(b) Unpublished Projects (by Kapil Sharma)

3:15 – 5:00 PM Physical/Online workshop

(Participants to upload their data and use portal functionalities for analysis)

Morning Session

The morning session was conducted also over google hangout (For online participants) and contents are available at <https://www.youtube.com/watch?v=ksPeu4CCvBA>

Inauguration and introduction: Prof. P J Narayanan (Director, IIIT-H) & Prof. Vasudeva Varma (Dean R&D, IIIT-H) both of them lauded the initiative. IIIT-H is very happy to have a project like this. CCPM project is a good example of Computer Science and Science group working together: more such project must be coming in future. Interdisciplinary projects are where we find lot more productivity. Prof Abhijit Mitra gave Introduction to CCPM version 3.4.

Conversation with Experimentalist

After the brief overview of the why and what of the project "Computational Core for Plant Metabolomics", Prof Mitra introduced the conversation by outlining the context by indicating that the CCPM portal is ready with most of the major functionalities. During the last workshop CCPM-5 and also in the last DBT review meeting, it was felt that

- 1. The software now needs testing with real data from the experimentalists*
- 2. Constant dialogue should evolve between the portal team and the users*

Accordingly all bug fixes, and some of the enhancements, as suggested during the last workshop (rest being in the final stages of development for v4.x) were incorporated. In December 2015, this revamped version was released to all concerned DBT experimentalists, who were requested to create their PI accounts in the portal, and to upload their own raw data for testing. This "Conversation" is a follow up, primarily to address the 2nd point. The objective of this dialogue was not only to collate inputs based user experiences with the current version, but also to understand more about requirements and expectations from the final version of this facility.

As experimentalists we had Dr. R.P Sharma and Dr. T Sarada at IIIT-H. There were others who joined from the JNU venue, along with Dr. Indira Ghosh, the Coordinating PI of Project CCPM. The conversation was coordinated from IIIT-H by Prof. Mitra along with Dr. Nita Parekh, Dr. Vikram Pudi and Dr. Sangeeta Kumari.

Prof R.P.Sharma started the conversation by mentioning that excellent modules have been developed to deposit and analyze the data. He also discussed the workflow of the

experimentalists, highlighting the complexities starting with defining of harvesting of plant material from different conditions and different time of harvesting, starting from day one to succeeding time. He pointed out that conditions of harvest, protocol vary from sample to sample eg. if one is analyzing primary metabolites or tomato cuticle or secondary metabolites, protocols for extraction of tissue will vary. Based on his experience with tomato metabolomics, he felt that the portal will be helpful in storing, visualizing and analyzing his data, and that the different modules had the potential to match the workflow involved.

He felt that the portal should be tested in real situation and validated regarding its ability to associate different types of data (GC/LC) with their protocols. A major problem arises because high - end machines come with inbuilt or incorporated software for processing data. Most of the machines don't allow the interactive inputs from users, thus making it difficult for users to resort to other analysis tools of their choice. We can get a TIC file in GC-MS in 30 min. with retention time and retention index, but at end of the day we have probably 100 files from different samples, with their controls and replicates, which need to be taken care of and analyzed and linked with each other. This is where the process becomes complicated because most commercial machines will not handle it. CCPM can be very useful to handle these types of situations.

Another problem is in matching data with metabolites from NIST and GOLM library available, eliminate false positives, and, finally integrating the data to answer research needs. This problem is not for all groups, some groups have developed their own tools and software's, but they don't share with each other, and there is a serious problem when one is working in isolation. He mentioned that his group faced this problem earlier since they started getting data far ahead of CCPM being able to launch the functional platform. So they used software which are available in market or from colleagues to analysis their data. He was happy to note that CCPM has slowly caught up and has the potential of assisting them with their work.

Dr Vikram spoke next to explain the technology used in implementing the different functionalities of CCPM. He mentioned about the bug fixes additional features incorporated since the last workshop. Many more features are in the pipeline and will be available in the next release. Major updates were on interactive plots, publication quality images are given. Default option is not publication quality we have given the users this option to check for publication quality images, customize the statistical part by adding more parameters.

Discussions followed and Dr. Vikram Pudi and Dr. Nita Parekh provided explanations on both technology and algorithm issues. The contents of the ensuing discussions are in **Appendix – 2a**

Lecture by Prof. RP Sharma on “Metabolic shifts during trans-differentiation of tomato plant”

Tomato seedlings of different mutant and wild type have been used for this study. He first compared the callus initiation and mass among mutants with wild type. The increase in callus mass in mutants was higher than the WT. We went for primary metabolite analysis and phytohormone profiling in the callus of mutants vs WT. We observed that during callus initiation the levels of all the intermediates corresponding to TCA cycle, Glycolysis, fatty acids cycles, and cell wall degradation were higher in mutant callus than WT callus as well as the transcript profiling of some important gene related to these pathways also showed correlation with

metabolites. Also the remarkable change was observed in phytohormone levels during callus initiation.

On the other hand, during callus differentiation the levels of primary metabolites and transcripts did not altered significantly in mutants compared to WT. Also the phytohormone levels were not altered during callus differentiation.

Questions and answers are summarized in **Appendix – 2b**

Afternoon Session

Tutorial and Hands on session

(Also conducted over Skype for online participants)

- The tutorial was presented by Dr. Sangeeta. The printout of the tutorial was given to the physical participants as hand out. The soft copy of the tutorial is available in the portal.
- The following Demos related to functional features of CCPM v3.4 were presented:
 - (a) Published projects (by Dr. Sangeeta Kumari)
 - (b) Unpublished Projects (by Kapil Sharma)
- Handouts for DEMO were distributed for all the participants. (**Appendix – 2c**)

The hands on session was for participants to try out all functionalities on either of the workshop portals at JNU and IIIT-H respectively. It was announced that the workshop portal will be kept alive for at least 15 days after the workshop for participants to continue their trials. Extensive and detailed discussions went on during the handson session and covered topics such as related to internal standards, biological replicates/technical replicates, METLIN links for peak annotation, data variance among groups, annotation of peaks detected, etc.

- Questions and answers and the ensuing discussions are summarized in **Appendix – 2d**

Follow up action plan to contribute further to the project's objectives

The suggestion and feedback from the participants are noted will be incorporated into the CCPM portal.

Conclusions:

The one day workshops at both the places were successful. Valuable suggestions received will enable us to improve the prototype development.

Appendix 1a

Workshop Announcement



Computational Core for Plant Metabolomics



Sponsored by DBT Govt. of India

Announcing CCPM-6 **Conversation with Experimentalist** **Workshop on February 24, 2016** *Conducted jointly by JNU Delhi and IIIT-H*

We are pleased to announce the sixth one day workshop (CCPM-6), based on CCPM v3.4 < metabolomics.iiit.ac.in >, entitled as 'Conversation with Experimentalist' to be conducted concurrently (both online as well as physical hands on) at two venues: JNU Delhi & IIIT Hyderabad on Wednesday, 24th February 2016. All the presentations and lectures will also be streamed live for online participants.

There will be no registration fees, and local hospitality will be taken care of. However, participants are requested to fill in the online registration form, available at the workshop portal, <metabolomics.iiit.ac.in/workshopregistration>, for arrangements and logistics. The portal has all requisite information about the workshop. All participants are welcome to the workshop venues (Hyderabad/Delhi) according to their convenience. Participants, who cannot make it physically, may join online using Google Hangouts.

Depending on your convenience for hands on, you may either register as physical or as online participant. Especially participants from Hyderabad or Delhi are strongly encouraged to participate physically at their respective workshop venues. Those who will be attending the workshop physically are hereby requested to bring their personal laptops. Online participants are also requested to provide us with their SKYPE account details and test it beforehand.

Please mark your calendars and save the date to attend the CCPM-6 Workshop 'Conversation with Experimentalists' on 24th February 2016. A detailed program schedule will be following soon.

All the necessary information about the workshop are available in the portal <metabolomics.iiit.ac.in/workshopregistration>. We look forward to your active participation in the workshop.

With Best Regards
Sincerely,

Dr. Indira Ghosh
School of Computational & Integrative Sciences
Jawaharlal Nehru University
New Delhi 110067
Phone: +91-9971287771; 011-2673 8707
Email: indira0654@gmail.com

Dr. Abhijit Mitra,
Center for Computational Natural Sciences and
Bioinformatics, IIIT Hyderabad
Gachibowli, Hyderabad - 500032
Phone: +91-40-6653 1156(O)/ 1201(R)
Email: abi_chem@iiit.ac.in

Appendix 1b

Workshop Invitation



Computational Core for Plant Metabolomics



Sponsored by DBT Govt. of India

CCPM-6

Conversation with Experimentalist

Workshop on February 24, 2016

Conducted jointly by JNU Delhi and IIIT-H

As you are aware, “Computational Core for Plant Metabolomics (CCPM)” is a collaborative project being implemented jointly by IIIT-H and JNU Delhi. Funded by the Department of Biotechnology (DBT), Government of India, the goal of the project is to create a national computational facility to support collaborative initiatives in Plant Metabolomics. The implementation involves development of a LIMS platform integrated with databases and tools required for storage, retrieval and analysis of metabolomics data as per the International Metabolomics Standards Initiative (MSI) guidelines.

In this connection, you must have by now received the announcement of the 6th Workshop (CCPM-6) – “Conversation with Experimentalist” at IIIT Hyderabad, CEH Auditorium scheduled on February 24th, 2016 jointly conducted by JNU Delhi and IIIT-H. As a build up to this workshop, an updated version of CCPM v3.4 had been released in December 2015. This was based on suggestions received during the previous workshop CCPM-5, from DBT experts and users, who were requested to test the updated version with live data. With nearly all modules already incorporated in this version, we have planned workshop CCPM-6 as “Conversation with Experimentalist” so that their suggestions can be incorporated, along with new functionalities in the upcoming release version.

This will be a one day workshop. The morning session will have lectures and conversation with experimentalist. In case you have not already done so, we request you to depute members of your group to create ‘User’ logins and begin testing the CCPM portal. This would help us in having a fruitful conversation session. In the afternoon, we will have online and hands on sessions at two different places JNU Delhi and IIIT-H, respectively. The detail program is appended below.

We cordially invite you all to participate in this workshop, along with one or two of your project personnel. You may also participate online by going to this link www.youtube.com/watch?v=ksPeu4CCvBA where the morning seminar session will be broadcasted. The participants coming to attend the hand on session are requested to come with laptops. They are also requested to come with their own raw data to test the functionality of the CCPM portal. This will enable us to incorporate your valuable inputs for further improvements in the CCPM portal. For accessing the afternoon demo session, online participants will be requested to connect on Skype (CCPM_IIIT). Important: Link for registration <metabolomics.iiit.ac.in/workshopregistration>.

For further details on the workshop, or for any assistance in registration or usage, please contact:

JNU Delhi: Gaurav Sharma

IIIT Hyderabad: Sushma Venkatesh

Email: gauravsharma.aiims@gmail.com

Email: sushma0140@gmail.com

Phone: +91 9013677887

Phone: +91 9908729299

Your presence in this workshop and your feedback is very important for the success of the project. We look forward to have you with us on February 24, 2016.

With Best Regards
Sincerely,

Dr. Indira Ghosh
School of Computational & Integrative Sciences
Jawahar Lal Nehru University
New Delhi 110067
Phone: +91-9971287771; 011-2673 8707
Email: indira0654@gmail.com

Dr. Abhijit Mitra,
Center for Computational Natural Sciences and
Bioinformatics, IIIT Hyderabad
Gachibowli, Hyderabad - 500032
Phone: +91-9866406028, 040-6653 1201(R)
Email: abi_cham@iiit.ac.in

Appendix 1c

Follow up Mail



Computational Core for Plant Metabolomics



Sponsored by DBT Govt. of India

CCPM-6

Conversation with Experimentalist

Workshop on February 24, 2016

Conducted jointly by JNU Delhi and IIIT-H

Dear Experts and PIs,

We hope you have received our earlier announcement and invitation letter, attached for your reference, please find. We cordially invite you all to participate, with your teams, in this workshop. If you are not able to join us personally, please depute one or more of your students, research fellows or any other team member for attending the workshop.

In case you have not already done so, please preregister for the workshop at the earliest, giving details of the venue of your participation (JNU, Delhi or IIIT Hyderabad), and mode (Online or physical). This is important for us to work out the logistics. The detailed instructions for attending online are highlighted below:

Workshop details and preregistration	metabolomics.iiit.ac.in/workshopregistration
Morning Session: Portal overview, Conversations and Lecture	Please log on to: www.youtube.com/watch?v=ksPen4CCvbA
Afternoon Session: Demo and Hands on session. (NB. Physical participants are requested to bring their laptops, loaded with their own data* for use in the hands on session. Online participants are requested to keep their own data* ready for upload)	For accessing the afternoon demo session, online participants are requested to connect on Skype (CCPM_IIIT). For the afternoon hands on session, please log on to metabolomics.iiit.ac.in/workshop NB. For accessing the workshop portal, it would be necessary to preregister following the preregistration link.
For any help please check our user manual file and tutorial	metabolomics.iiit.ac.in/workshop/wiki/usermanual < metabolomics.iiit.ac.in/workshop/wiki/tutorial >

*NB. LC – MS /GC MS analysis raw data (NetCDF/ mzXML/ mzData or mzML files) to access the functionality of the portal from preprocessing to metabolite identification and their statistical analysis.

The data that you upload in the workshop portal will be perfectly secure and visible/accessible only to you and your team.

For any queries please contact:

Gaurav Sharma (JNU), Mobile +91 9013677887, email: gauravsharma.aiims@gmail.com and
Sushma Venkatesh (IIIT H), Mobile +91 9908729299, email: sushma0140@gmail.com

We look forward to seeing you.

With Best Regards

Sincerely,

Dr. Indira Ghosh
School of Computational & Integrative Sciences
Jawaharlal Nehru University
New Delhi 110067
Phone: +91-9971287771; 011-2673 8707
Email: indira0654@gmail.com

Dr. Abhijit Mitra,
Center for Computational Natural Sciences and
Bioinformatics, IIIT Hyderabad
Gachibowli, Hyderabad - 500032
Phone: +91-40-6653 1156(O)/ 1201(R)
Email: abi_chem@iiit.ac.in

Appendix 1d

Program Details



Computational Core for Plant Metabolomics

Sponsored by DBT Govt. of India



Program (February 24 2016)
CEH Auditorium, IIT Hyderabad

Morning Session:

10:00 AM – 10:30 AM	Introduction about the CCPM version 3.4
10:30 AM – 11:15 AM	Conversation with Experimentalist
11:15 AM – 11:30AM	Tea Break
11:30 AM – 12:30 PM	Lecture by Prof RP Sharma on “Metabolic shifts during trans-differentiation of tomato plant”
12:45 PM – 1:45 PM	Lunch

Afternoon Session:

	Demo related to functional features of CCPM v3.4
	<ul style="list-style-type: none">• Published projects:<ul style="list-style-type: none">○ Browsing through published data and results○ Using analysis tools on published data• Unpublished projects:<ul style="list-style-type: none">○ Populating projects with sample data○ Using analysis tools on uploaded data
1:45 PM – 3:00 PM	
3:00 PM – 5:00 PM	Physical/Online workshop (Participants to upload their data and use portal functionalities for analysis)

Appendix 2a

Question came up from online participants during “Conversations”

Q. Please first discuss about the back-end programming?

A. Back-end is based on R scripts for preprocessing, normalization and statistical analysis which are from XCMS and Metaboanalyst. Portal is developed on web2py.

Q. The alignment while data preprocessing against what?

A. Question was not clear

Q. In the introduction as you mentioned that there is no limit to data input in ccpm so how much is the database size?

A. It has limit which can be extended if needed by requesting for extra space.

Q. Is there any specific method for normalization of metabolomics data for concerns ccpm database?

A. No there no specific method for normalization of metabolomics data for concerns ccpm database because it depend on the type of your data.

Q. In the flow chart provided in user manual for data per-processing step have mentioned alignment, so my question is against which library this alignment is performed?

A. Alignment is not based on any library but among peaks from different samples in same group. It is based on m/z values and retention time. Centwave algorithm is used which is most suitable for high resolution LC/{TOF, OrbiTrap, FTICR}-MS data in centroid mode. In the first phase of the method mass traces (characterized as regions with less than ppm m/z deviation in consecutive scans) in the LC/MS map are located. In the second phase these mass traces are further analyzed. Continuous wavelet transform (CWT) is used to locate chromatographic peaks on different scales.

Q. If you can please discuss the work flow of ccpm creation or development. Details or if you can advise some tutorial that can help me in development of any similar tool?

Q. Regarding Normalization what options it can provide and how they differ from each other. Not interested in algorithm?

A. User has the option to perform the following normalization methods:

Row Wise Procedures

Row – wise aims at adjusting the variance of different samples (to normalize each sample (Row) so that it is comparable to the other). User has the option of choosing ‘none’ to skip normalization, ‘by Sum’ and ‘by Median’. The median is preferred to average as a statistical measure since it is more robust.

Column Wise Procedures

Column-wise aims at adjusting the variance of different features executed on the columns of data (intensity) across all samples.

There are two components such as ‘Data transformation’ and ‘Data Scaling’

Data Transformations

Transformations are generally applied to convert multiplicative relations into additive relations, and to make skewed distributions symmetric. User has the option of choosing ‘None’ to skip, ‘Log transformation’ or ‘Inverse hyperbolic sine (IHS)’.

Log Transformation

It is the classic solution to normalizing skewed data. It adds symmetry to positive and negative

changes within a dataset as well as minimizes the impact of outlier entries. A drawback of the log transformation is that it is unable to deal with the value zero.

Metabolomics data from any given sample will always possess far more metabolites in the low-abundance range than the higher range, thus necessitating a log transform to eliminate the impact of the minority high-abundance metabolites.

Inverse hyperbolic sine (IHS)

Generally possesses the same characteristics as a classical log transform, such as large value suppression. The primary difference between the IHS and log transformations is that while a standard logarithm is undefined at zero, the IHS function is fully defined and thus will not remove zeros from the data during transformation.

Data Scaling

User has the option of choosing 'None' to skip, 'Auto scaling', 'Pareto scaling' and 'rangescaling'.

Auto Scaling

Mean-centered (each column of the table can be achieved a mean of "0" by subtracting the column mean from each value in the column) and divided by the standard deviation of each variable. Each column of the table can be scaled so that it has unit variance by dividing each value in the column by the standard deviation of the column. It results in every feature displaying a standard deviation of one, i.e. the data is transformed to standard units.

Advantage: All metabolites are equally important.

Disadvantage: Inflation of measurement errors.

Pareto Scaling

Mean-centered (each column of the table can be achieved a mean of "0" by subtracting the column mean from each value in the column) and divided by the square root of standard deviation of each variable) Pareto scaling is similar to auto scaling but the square root of the standard deviation is used as the scaling factor instead of standard deviation. Its normalizing effect is less intense, such that the normalized data stays closer to its original values. It is less likely to blow up noisy background and reduces the importance of large fold changes compared to small ones. However, very large fold changes may still show a dominating effect.

Advantage: Stays closer to original measurement.

Disadvantage: Sensitive to large fold changes.

Range Scaling

Mean-centered ((each column of the table can be achieved a mean of "0" by subtracting the column mean from each value in the column) and divided by the range of each variable. A disadvantage of range scaling with regard to the other scaling methods tested is that only two values are used to estimate the biological range, while for the standard deviation all measurements are taken into account. This makes range scaling more sensitive to outliers.

Advantage: All metabolites equally important. Biologically related scaling

Disadvantage: Inflation of measurement errors, sensitive to outliers.

Appendix 2b

Questions asked in Lecture of Prof R.P. Sharma

1. Dr. Sangeeta had asked regarding the mapping of correlations between the metabolites onto biological pathways to establish whether the metabolites are correlated just by random chance or if there is some biological relevance
2. Dr. Nita had asked how MetAlign and MSClust have been used to help in data reduction. Mr. Kapil Sharma to give the details. (Kapil explained Basically MSClust is meant for deconvolution or data reduction by m/z value and peak intensity while MetAlign is basically used for peak alignment and peak corrections.) She also asked about the meaning of positive and negative interactions between the metabolites that were shown in the slides (The positive interaction between the metabolites indicated the up regulation of the metabolites while negative interaction indicated the down regulation of particular metabolites.)
3. Dr. Sarada asked regarding the fact that the study is based on callus formation in a tissue culture. However, in the field condition, the metabolic shifts might be very different
4. Dr. Vinod had asked a question regarding distinguishing the metabolic shifts that happened during the callus formation and those during root and shoot differentiation.

Appendix 2c

Demo Handout

PREPROCESSING AND DATA ANALYSIS OF PUBLISHED PROJECT

Metabolic profile data of Arabidopsis four aerial tissues-Project 6

References: Matsuda F. et al., MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. Plant J. (2009). 57(3):555-77. doi: 10.1111/j.1365-313X.2008.03705.x.

Browse the data and metadata

The Groups to study chosen

Group 1 -Group 4

Comparison between the two groups

Group 1 [cauline leaf_tissue to Group 2 [flower_tissue]

Group 3 [Rosette leaf_tissue] to Group 4 [Internode_tissue]

Modules 1>Data preprocessing>Preprocessed Results>view results

Peak list link to Metlin database for Annotation purpose

A brief introduction of applicability of different normalization and statistical tools will be given during the demo session.

Normalization using column-wise, data transformation (log) and row wise normalization (showing different procedure.

Statistical analysis using different methods, different analysis and view results

GC-MS data preprocessing using CCPM portal and linking the preprocessed peaklist file to GOLM database for identification.

Flow chart for analysis of unpublished project entitled “Phytohormone analysis in tomato leaf tissue”

Data upload

- Sign in
- Create a Project or select a running project
- Click on metadata

View metadata

- Download the bulkupload template and filled in the required information of the sample
- Groups will be formed with sample number and names

Uploading raw data files

- For each sample, upload a raw file
- After uploading files, go to my task

Selecting groups for pre-processing

- Click on preprocessing
- Browse and download pre-processed results files
- Click on +New preprocessing task button
- Filled in study name, select single project
- Select a project and select set1 and set2
- Select instrument and select samples and submit
- After preprocessing, go through the preprocessed files, Retention time, TIC plots, Box plots and EIC plots
- Preprocessed files can be downloaded and go through the METLIN and GOLM links for peak annotation

Data pretreatment and statistical analysis

- Take peak list data and submit for normalization accordingly
- After normalization go for statistical analysis according to your data set like correlation, dendrogram, heat map and SVM etc.

Appendix 2d

Question asked in Hands On session

Q. Can internal marker or standard marker be used for normalization?

A. Right now there is no provision for normalization using internal standard. CCPM uses normalization by sum or median to normalize the data. CCPM can also transform your data using log transform or inverse hyperbolic sine. For Data Scaling CCPM provides Auto, Pareto and Range scaling.

Q. How do we annotate the peaks?

A. Metlin links are provided in diffreport for each feature that is detected in pre-processing and when clicked on the link the user gets a list of probable compound in that mass range. Diffreport is named as plot.tsv in results.zip.

Related to internal standard - If we could incorporate extra normalization option with the help of internal standard (which could be any stable chemical compound that is added during sample

preparation) then it would be better to understand the exact intensities of the peaks corresponds to the particular metabolites.

Related to biological replicates/technical replicates - Biological replicates indicates the samples which are extracted at same time (or occasionally different time) from the different plants/animals of a variety or mutant, maintained in a similar environmental condition. On the other hand technical replicates means samples that are aliquoted from a biological replicate.

Related to the METLIN links for peak annotation - During the results obtained by statistical analysis, the features defining the different pattern in SVM plots, If we have some option of clicking on the features dots this should directly indicate the peak number in the peaklist table, this would be better and time saving.

Data variance among groups:

Dr. Sarada Tetali from HCU during hand-on session asked how to get idea of the data variance among the groups. We showed her the option of box plot graph for each feature available at the preprocessing Tab. The box plot graph shows mean, median, max and min value for each feature.

Annotation of peaks detected

She also discussed about how peaks detected can be annotate in CCPM to which we had shown her that CCPM provides Metlin database links for each peak and when clicked on the link the user gets a list of probable compound in that mass range.