Minutes from the CAIGE BMS Working group meeting
Friday 5th November 3.00 - 4.20pm

Attendees: - Julie Nicol, Robin Wilson, Amit Singh, Meiqin Lu, Scott Sydenham, Chris Moore, William Eusebio, Chris Lisle and Brett Lobsey.
Apologies: - Richard Trethowan, Mark Dieters

Julie – Introduced Chris Lisle (CBB, UoW) who will be managing the CAIGE data analysis moving forward. Also, William Eusebio introduced himself and the work that his does on the BMS database.

Points discussed during the meeting:

1. **Data Compilation and BMS** –
   Robin - There was an issue with categorical data while using graphical query tool for data compilation. This issue has been fixed by BMS development team on staging server and soon the necessary changes will be reflected on the production server. There was a memory size limit which was encountered while including the several studies in the tool and this has been increased. This will give more control on filtering and data compilation process on BMS.
   Conclusion – moving forward will be able to perform data compilation using BMS tools. (As the data is shared from the breeders it will be uploaded onto the CAIGE BMS)

2. **Loading of CG data on BMS**
   Julie – Introduced the point to load some valuable CG (CIMMYT & ICARDA) data on CAIGE BMS. Some useful data (w.r.t Australian Breeders) can be uploaded on BMS to help breeders in taking breeding decisions. We along with breeders can identify such data for putting on BMS.
   Robin – We can put a cut down version of those data and provide a link of CIMMYT and ICARDA database if they permit us. Moving to the next phase of CAIGE, it is good to have CG data on BMS and we can use data query tool to narrow down our search on BMS.
   Meiqin – We wish to have some extra or specific information on some lines but need to consider the time and effort required behind this. It should simple or quick to access.
   Scott – Supported the Meiqin thoughts and expressed to have more and more data on hand will help us with finding better parent.
   ACTION STEP: the addition of CG data to this year's CAIGE Yield Trial entries in BMS will be achieved. Ongoing work will explore the possibility to link CG data bases with BMS.

3. **A summary of the key points from CAIGE consultations around BMS and data analysis.**
   The BMS would like to integrate some of MET analysis data visualization tools developed by CBB (University of Wollongong).

   Julie – In discussion with Brian and his team, on CAIGE data analysis, we will have single year MET analysis and Over year MET analysis. CBB is consistently refining and improving the methods of analysis and developing some interactive tools. It would be logical to put some of tools in BMS to facilitate interpretation of data.
   Robin – indicated that BMS has an associative package called BREEDING VIEW, mainly being designed for developing countries for doing single site and multi-site analysis. Data prepared from BMS goes into this package and analysed.

The point was that the interaction with the Data Visualization tool could happen in the same way as with BREEDING VIEW whereby the data is prepared in the BMS to enter the Data visualization tool, and then possibly upload the results.
Scott – Commented that the more tool available will be better.
Chris – We have developed data visualisation apps which can be easily integrated on BMS.
ACTION STEP: Amit, Robin, Julie and Chris to have a discussion to see how the CBB visualization tools can be integrated into BMS moving forward.

4. **Disease Capture with new project – in relation to BMS**.
Moving forward in new CAIGE project, the pathology collaborators have agreed to provide the disease rating following the NVT format along with raw data.
Julie – In new CAIGE, our focus will be more on standardizing all CAIGE data including pathological data.

5. **Capturing of meta data on CAIGE Yield Trials**.
There are lots of missing information (example harvesting dates) and non-uniformity in reporting of such data from the collaborators. Moving forward, we need to discuss on the nature of metadata need to be collected from all yield trials. I have listed some such data based on breeder's comments and suggestions.
    a. Sowing and Harvesting dates
    b. Average rainfall of the site including moisture condition
    c. Soil type & soil treatments
    d. Paddock/plot history
    e. Summary of any specific production constraints (For example biotic stress or abiotic stress or other climatic factors) during the trial.

Chris – Commented that it is great idea to have these data, it helps us to filter the data and finding what we are looking for.
Scott – He is ready to help with any historical data if required on their CAIGE trials.
Julie – In new CAIGE more stress will be put on collecting data in a standardised format. We have proposed to use Predicta B assay for soil borne pathogens on each CAIGE trial. This gives us additional information on trials.
Brett – Why don't we can filter the historical data on BMS and form disease core set?

6. **Discussion on disease Core set –**
Julie – We are collecting disease core set from pathological collaborators. Once we have the set completed for each crop (BW, DW & Barley). The CAIGE database management team will populate all available CAIGE data and CG data on them to be accessioned and stored by Brett (AGG, Tamworth).

7. **COPs progress (with a view to having these available with the next Data Analysis).**
William and Robin et al have done a sterling job in fixing the pedigrees to work out the COPs matrix especially for bread wheat. We are liaising people to fix the truncated pedigrees of durum and barley for both pedigree-based analysis and generating COPs matrixes. The generation of COPs matrix is in progress and will be updated during the meeting.

Robin – William has run the software for COPs for Bread wheat and result have been extracted. We have worked solidly on BW and Barley to fix the truncated pedigrees and still not completed. We will work for DW and Barley once the pedigree issues are fixed completely.

William – Demonstrated and presented the result file of COPs analysis for BW and expressed concern over many data points due to lack pedigree information. He was confident to fix this problem of missing pedigree string with help of old IWIS data base and run the software again. We have requested Sandra (previous CAIGE database manager) to share the database.
Robin – The COP presented is William's effort and it is considerable with 2.3M lines of data providing associations between lines. There are still gaps to fill but we will get there.

ACTION STEP: Plans are afoot to have the COPs available for this year yield trial entries.

In closing Julie indicated that a new CAIGE for 5 years is currently under negotiation with GRDC and once this is finalised, we will have a clear plan for what is including moving forward. She indicated that database management, analysis and streamlining are key features (aside from the core of the project). The next meeting is planned for most likely February considering the field harvest is underway and Christmas is around the corner.

Amit thanked all participants and the meeting was closed at 4.20pm.