


























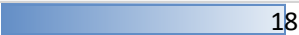
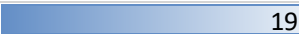





## MBGP Brassica Information System Questionnaire 2018/19

Preliminary report 7/1/2019

**Total respondents = 39** (all identified as University/academic apart from one industry)

Argentina		1
Australia		8
China		1
France		5
Germany		2
India		5
Korea (south)		1
Poland		2
Spain		1
The Netherlands		1
United Kingdom		7
USA		5
<b>User domains (1 or more per person)</b>		
Structural genomics		15
Functional genomics/transcriptomics		26
Quantitative and statistical genetics		21
Breeding		19
Agronomy/soil nutrition		5
Physiology		10
Biochemistry		5
Plant pathology/entomology		7
End-use quality		4
Phenomics		8
Proteomics		1
Metabolomics		7
Bioinformatics/computational biology/data m		11
Other		5

<b>Level of Curation - what do you consider to be desirable and realistic objectives for MBGP</b>		
Data inventory - categorized set of links to repositories and 'supplementary datasets		19
Data warehouse - collection of datasets with persistent DOIs		18
Indexed databases with single portal (as developed for Wheat IS: <a href="https://urgi.versail">https://urgi.versail</a>		19
Set of MBGP data registries with indexes of key identifiers for different entities/data		12
Integrated data system allowing navigation between genome and phenotypic trait		16
Other		2

*Comment (other):* I regard 3 & 5 as desirable, the other options are more realistic in a 5-year timeframe but could be regarded as intermediate steps?

## MBGP Brassica Information System Questionnaire 2018/19

Preliminary report 7/1/2019

<b>DATA you would contribute/use :</b>		
<b>Genetic Resources (passport data)</b>	<b>Contribute data</b>	<b>Use data</b>
Bi-parental populations	14	19
Multi-parent populations	4	14
Diversity collections/Association panels	13	22
Landraces	7	17
Cultivars	6	19
Wild relatives	6	18
Mutants, insertion and gene-edited lines	3	15
Mutant populations (eg TILLING)	3	14
Other	1	0
<b>Genomic Sequences</b>	<b>Contribute data</b>	<b>Use data</b>
Bi-parental populations	10	25
Gene coding annotations	7	21
Proteins	1	15
Non-coding RNA	4	15
Molecular markers	9	23
Transposable elements	2	13
Segmental duplications	4	15
Simple repeats	1	11
Syteny	5	19
CHIPseq	0	10
Orthologues, paralogues, gene families	7	22
Epigenome (DNA methylation)	4	16
Other		1
<b>Sequence polymorphism scores</b>	<b>Contribute data</b>	<b>Use data</b>
Bi-parental populations	3	10
SNP	13	24
Indel	6	17
PAV (Presence-Absence variation)	9	20
CNV (Copy Number Variation)	5	20
<b>Maps</b>	<b>Contribute data</b>	<b>Use data</b>
Bi-parental populations	2	5
Genetic linkage	12	22
QTL	16	22
GWAS Manhattan plots	12	22
Physical (eg BAC tiling)	1	7
<b>Transcriptome and other 'Omics</b>	<b>Contribute data</b>	<b>Use data</b>
Bi-parental populations	1	7
RNAseq	14	21
qPCR	7	10
Gene Atlas (expression in tissue/stage)	5	20
Metabolite profiles (metabolomics)	8	13
Protein profiles (proteomics)	2	10
Ion composition (ionomics)	3	10

## MBGP Brassica Information System Questionnaire 2018/19

Preliminary report 7/1/2019

<b>Data exchanges with other databases/portals</b>	
Ensembl Plants	20
Brassica Information Portal	23
Cyverse (iPlant)	11
Elixir	2
NCBI / ENA	18
TAIR / IAIC (Arabidopsis portal)	21
URGI	6
www.brassica.info (MBGP portal)	24
http://brassicadb.org/brad/ (BRAD)	20
Other	4

Other: <https://brassibase.cos.uni-heidelberg.de/> and <http://www.brassicagenome.net/> (2)

*Comment:* Particularly for existing Brassica platforms I would prefer to discuss how/what would be exchanged and whether/how a BIS might incorporate existing structures. Portals which are less well maintained may become redundant, others with good ideas but which are not yet overflowing with data (BIP?) might be enriched by more exchange?

<b>In which services would you be interested ?</b>	
Data browsing	27
Database integration (link data together)	21
Ready access to experimental meta-data	20
Navigate trait to genome (GWAS/QTL)	25
Download data files	23
Download community analysis tools	17
BLAST or other alignment servers	23
Genome viewers	23
Synteny / collinearity viewer	23
Paralogue catalogue / finder	21
Complex query capability	12
Analysis workflows / pipelines	17
Computing capacity	10
Other	1 <i>Trait to plant accession</i>
<b>Which tools or data standards would you use ?</b>	
MBGP look-up data REGISTRIES for key iden	22
MIAPPE (for phenotyping)	16
BraTO (Brassica Trait Ontology)	20
MBGP standardised gene-model per genome	23
Look up table of legacy gene-models	16
Others	0

## MBGP Brassica Information System Questionnaire 2018/19


Preliminary report 7/1/2019

### **Restricted data access.** Would you be interested in such functionality?




(no if wished to ensure all data submitted to be public)

Yes (purpose)		5
No		29

### **Reasonable embargo period?**

Immediate		1
1 month		4
6 months		15
12 months		8
Any other consideration		3

*Other = upon release by the authors; Immediately upon acceptance of publication (if relevant)*

<b>Other comments/suggestions</b>		
Suggested funding sources/mechanisms for		3
None		31
Other Comments		1

Suggestions for BIS:

- *H2020 CORNET: Must include at least two EU countries (e.g. DE/UK/NL and/or FR?).*
  - *PRO: Application process relatively simple for EU and funding quota is unusually high at ~50%!*
  - *CON: Must include interested 5 small/medium industry partners in each participating country so very hard to raise the numbers, would need lobby work to bring in breeders, biotech companies, IT startups, etc...*
- *Sponsorship from industry players with potential interest in better access to large-scale public datasets, pangenomes etc., e.g. for use in deep learning applications (Bayer, BASF, Syngenta, KWS, Limagrain, NPZ, Rijk Zwaan)*
- *Feel free to contact [urgi-contact@inra.fr](mailto:urgi-contact@inra.fr) to get the source code of the WheatIS portal that have been developed to be generic for all plant species.*

Report compiled from Qualtrics summary data by Graham King, SCU, Jan 7<sup>th</sup> 2019