Plan for DivIII paper in light of Bonnema article, “Modern hybrid to future.”

Population that this research covers *(Why didn’t they cite Stansell?)*

Focused on the range of commercial broccoli traits.

First paper was broader.

Which part of this population is covered by Cai? Which is not?

We don’t have much of the LHL lineage, but lots in the AIL. We have better info on structure in the latter.

They did not provide code or raw data, so it is hard to check or to rerun with better reference.

Population structure.

What can we say about relationships and descent?

*We can say much more. See below*

Phenotype

* What variation can be detected with GWAS? *None detectable but there is noise that cannot all be rejected.*
* Which traits are fully confounded with population structure? *Can’t tell.*
* Can we partition variance in phenotypes among G, E and GxE? *Do we have E?*

Utility

Which accessions would one use to diversify genetics of without messing up fixed domestication loci?

*We can answer that question. There is not much from the Cai-Bonnema paper.*

The big advantage of this paper is that they have enough individuals to do GWAS. That is presumably coming later (they acknowledge help with field work). 1

**Reference genome.** A big shortcoming of the Bonnema paper is that they used the Lv genome, which not accurate and does not have SNPs of importance for detecting variation within broccoli. Zach has been unable to get the broccoli data to line up with that genome, and others have discouraged him from trying.

Improving the broccoli tree. We can do a lot better on the broccoli tree, especially the modern part.

Can they rerun the broccoli part of the tree using the HDEM or Parkin (or long-read Cauliflower) reference genome to recreate the AIL on a better framework? That re-run should find more relevant SNPs and make a more accurate tree. Alternatively, we can request access to their data, but that would be a lot of work.

Citation for HDEM broccoli reference genome:

[Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F. C., Falentin, C., ... & Deniot, G. (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. Nature plants, 4(11), 879.](https://www.nature.com/articles/s41477-018-0289-4)

(Emailed Chévre to find out what genotype the dh line HDEM is from.)

The better genome should resolve the true SNPs better and allow a less aggressive imputation of missing data.

Should we be using a IBS heatmap instead? The modern materials under study are not descended from common ancestors but have a lot of backcrossing and remixing. Use the loci with the most effect on the grouping. Does the grouping reflect breeding program at all?

Other topics to address in this paper.

**Reduced diversity regions.** Do regions of reduced diversity occur within this sample as well?

Do regions of enhanced diversity occur in this sample as well? If so, are they associated with any of the phenotypic diversity? (GWAS-like. Yes, no, cannot be determined).

**Increased diversity.** How much genetic variation has been introgressed from outside Cavolo Broccolo Ramoso Calabrese?

**GWAS.** We should show what GWAS reports. The conclusion will likely be a variation of “a reliable signal is not possible within a population this genetically uniform.” Even though it represents the full phenotypic diversity that allows us to describe variation within commercial-quality broccoli it does not have enough individuals to detect associations reliably. Since the genetic diversity is already well represented, adding more individuals will mean adding close relatives and the genetic information will be mostly redundant.

**Alboglabra.**  For the future, investigating alboglabra-derived broccoli could have great value. B1118 and B135 make a broccoli-like head but are genetically distinct. Alboglabra lines are close to broccoli and cauliflower in the tree. More than would be expected given the many sprouting headed types. They are generally white-flowered.

For Sandra Branham. She is doing WGS in 300 oleracea. She will need more to do GWAS. We might be able to indicate good candidates to add. An advantage of WGS over GBS is that you can add individuals over time, they don’t all have to be in one run.

**Center of Origin.** They identify GBR as the origin of the basal broccolis. We suspect that is because they came from Wellesbourne, not that they were British landraces. We should check that. It is worth saying somewhere that this identification is wrong.

The collards also seem to be misclassified.

**Their Table S1. Listing the accessions that were identified as ancestral**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Accession Number** | **New\_ Number** | **Type** | **Collection** | **Genus** | **Species** | **Variety** | **Origin** | **Accession Name** | **Gene- bank** | **Notes** |
| TKI596 | Br\_ge\_ssa46 | Broccoli spring-summer-autumn | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 8631 | HRI | winter cauliflower |
| TKI603 | Br\_ge\_un17 | Broccoli unknown | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 10702 | HRI |  |
| TKI1072 | Br\_ge\_un5 | Broccoli unknown | genebank | *Brassica* | *oleracea* | *italica* | GBR | PI 662529 | GRIN |  |
| TKI1074 | Br\_ge\_wi18 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | PI 662525 | GRIN |  |
| TKI1075 | Br\_ge\_wi19 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | PI 662526 | GRIN |  |
| TKI1076 | Br\_ge\_wi20 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | PI 662524 | GRIN |  |
| TKI1077 | Br\_ge\_wi21 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | PI 662528 | GRIN |  |
| TKI1078 | Br\_ge\_wi22 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | PI 662672 | GRIN |  |
| TKI592 | Br\_ge\_wi25 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 3577 | HRI |  |
| TKI593 | Br\_ge\_wi26 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 6223 | HRI |  |
| TKI607 | Br\_ge\_wi29 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 3516 | HRI |  |
| TKI608 | Br\_ge\_wi30 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 3543 | HRI |  |
| TKI609 | Br\_ge\_wi31 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 3546 | HRI |  |
| TKI610 | Br\_ge\_wi32 | Broccoli winter | genebank | *Brassica* | *oleracea* | *italica* | GBR | HRIGRU 3566 | HRI |  |

Their ID Their genebank # PI->HRI Name

|  |  |  |  |
| --- | --- | --- | --- |
| Br\_ge\_ssa46 | HRIGRU 8631 | 8631 | SICILIAN PURPLE LAKE BI74008 |
| Br\_ge\_un17 | HRIGRU 10702 | 10702 | Red Lion Donor: BI88538 |
| Br\_ge\_un5 | PI 662529 | 3570 | White Sprouting Improved EEC 0890-PCC 1982 |
| Br\_ge\_wi18 | PI 662525 | 3540 | Late Purple Sprouting EEC 0890-PCC 1982 |
| Br\_ge\_wi19 | PI 662526 | 3542 | Purple Sprouting Late EEC 0890-PCC 1982 |
| Br\_ge\_wi20 | PI 662524 | 3515 | Purple Sprouting Xmas EEC 0890-PCC 1982 |
| Br\_ge\_wi21 | PI 662528 | 3546 | White Sprouting Early EEC 0890-PCC 1982 |
| Br\_ge\_wi22 | PI 662672 | 3511 | Xmas Purple Sprouting EEC 0890-PCC 1982 |
| Br\_ge\_wi25 | HRIGRU 3577 | 3577 | Purple Cape Donor: EEC 0890-PCC |
| Br\_ge\_wi26 | HRIGRU 6223 | 6223 | Donor Heather at Winchester 1986 |
| Br\_ge\_wi29 | HRIGRU 3516 | 3516 | Purple Sprouting Early Donor: EEC 0890-PCC |
| Br\_ge\_wi30 | HRIGRU 3543 | 3543 | Purple Sprouting Late Improved. EEC 0890-PCC |
| Br\_ge\_wi31 | HRIGRU 3546 | 3546 | White Sprouting Early. EEC 0890-PCC |
| Br\_ge\_wi32 | HRIGRU 3566 | 3566 | Late White Sprouting. EEC 0890-PCC |

The ones labeled EEC 0890-PCC 1982 are from the Southern Italy collection Branca made and deposited

<https://link.springer.com/chapter/10.1007/978-3-642-14871-2_2>

Since 1982, several expeditions have been carried out by Professor Gómez-Campo from the UPM and his collaborators in order to rescue and collect Mediterranean populations of wild *Brassica* species. These missions were supported by the International Board for Plant Genetic Resources (IBPGR), later International Plant Genetic Research Institute (IPGRI), and now Bioversity International and were performed in the Mediterranean coast of Spain, Italy, Greece, and Tunisia and along the Atlantic coast of northern Spain, France, and the UK. As a result, different wild *B. oleracea* species with a chromosome number of *n* = 9 (including Atlantic *B. oleracea*) were collected. Four wild *B. oleracea*-related species were found in Sicily (*B. rupestris*, *B. incana*, *B. villosa*, and *B. macrocarpa*). Gómez-Campo and Gustafsson ([1991](https://link.springer.com/chapter/10.1007/978-3-642-14871-2_2#ref-CR34_2)) described the accessions collected in detail and the new locations found. According to the IPGRI policy, each sample was split into three parts, which were stored at the UPM (Spain), the University of Tohoku (Sendai, Japan) and also at seed banks of those countries, where the collection was done (Izmir, Turkey; Thessaloniki, Greece, Bari, Italy; Porquerolles, France; Kew, UK). Recently, two new expeditions have been carried out by the UPM team. The first one targeted the northern coast of Spain (Gómez-Campo et al. [2005](https://link.springer.com/chapter/10.1007/978-3-642-14871-2_2#ref-CR36_2)) and the second one was focused on the northeastern coast of Spain in search of new localities and seeds of *B. montana* (Gómez-Campo et al. [2007](https://link.springer.com/chapter/10.1007/978-3-642-14871-2_2#ref-CR38_2)).