The mosaic oat genome: what's in it for you?

Nick Tinker, Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, 960 Carling Ave., Ottawa, ON K1A 0C6; nick.tinker@agr.gc.ca.

Oat might be "Nature's original, benevolent Frankenfood".

But please don't quote me on that without the following context. Many millennia ago, before Mary Shelley frightened us about rogue modified organisms, natural selection invented an organism with a monstrously large and complicated set of 21 chromosome pairs. From a simple, barley-like common ancestor with seven chromosome pairs, a set of wild oat species evolved, each with a genome that we call either "A", "C", or "D". The genomes of these wild oats looked like barley chromosomes that came out of a blender, each with large and small chromosome chunks that were chopped up and rearranged differently. Then the C and D genomes decided to marry, forming a new wild ancestor (similar to a species called *Avena insularis*) with 14 chromosome pairs. Feeling left out, an A genome ancestor (similar to a species called *Avena longiglumis*) joined in the fun, making the kinky m*énage à trois* with 21 chromosomes that we call cultivated oat *(Avena sativa)*. Strangely, those original chromosome chunks continued to rearrange themselves, making oat look more like *three barley genomes in a blender*. And it seems that the oat chromosome blender never stopped, because even two different modern oat cultivars can have chromosomes with different rearrangements. Frankenfood indeed!

So, come up to the lab, and see what's on the slab!

This story is about breakthroughs in our understanding of the oat genome, as reported in a recent paper in the journal Nature (Kamal*, et al.*, 2022) and in a companion paper (Tinker, *et al.*, 2022). As I will explain, this story is about more than just "another bunch of genome sequences". It is about understanding those sequences, solving mysteries that have challenged oat researchers for decades, and setting the stage for new, transformative research that will benefit everyone who cultivates oat, processes oat, or eats and drinks oat products. We called oat a "mosaic" rather than a "Frankenstein" (but I needed to get your attention). The mosaic metaphor is better. Those of us who have spent many years mapping the oat genome already knew that oat was a mosaic, but, until now, we did not have the tools to see the pictures in that mosaic, or to understand how that mosaic evolved. Now we have dusted off those ancient and mysterious tiles to reveal the complex and beautiful patterns that make up the 21 oat chromosomes - chromosomes that contain more than 80,000 genes, working together to orchestrate the healthy food-plant that we love so much.

Barriers and bonanzas.

The complete sequence and detailed annotation of all genes in the oat genome opens new doors to discovery. A genome browser maintained by the GrainGenes database at the USDA (https://wheat.pw.usda.gov/jb/?data=/ggds/oat-sang) allows anyone to access the DNA sequence of any part of the enormous oat genome, identify what genes are there, and determine if a gene in that region affects a trait. Tools like this have been available for many years to researchers studying model plant species like Arabidopsis, as well as those studying crops such as corn, soybean, and canola, and those crops have benefited from early research on gene

function and gene variation. Now it is time for oat researchers to leverage that groundwork from other species.

This catch-up/leap-frog game was demonstrated nicely by several biological discoveries described in the Nature paper. These include a thorough characterization of genes affecting oat β-glucan fiber, the discovery of a gene that increases a protective wax coating, and new molecular evidence supporting the safety of oat in gluten-free diets. Such discoveries are merely the first examples of exciting work that will follow. Based on work in other species with fully characterized genomes, we can expect a bonanza of discoveries in oat at the molecular level. It is fairly easy to predict that this will include genes and mechanisms that determine heading date, pathogen resistance, and levels of major nutrients, such as oil, protein, starch, and fibre. It is likely that other oat-specific traits will also be unravelled - traits like antioxidant levels, hull content, or even flavour.

In effect, the largest barrier in oat research has finally been removed. The starting point for molecular-level discovery is now freely available in public databases for all who know how to use them. With the removal of this barrier, many projects can now be proposed, funded, and completed within the typical three- to five-year research cycle. This is something that funding agencies will recognize, but it is also something that young, ambitious students and postdocs will recognize. Research projects in oat genomics will suddenly go from being risky, niche endeavours to being plausible, attractive, and sexy! One might even suggest that the oat field is filled with low-hanging fruit.

But barriers still remain. Probably the biggest barrier is this: to use a new gene or trait, you need to find it in existing oat germplasm. The future does hold possibilities for going beyond natural variation, but in the short- to medium-term, synthetic biology such as gene editing may be restricted to the testing or validation of hypotheses. However, this use may be exceptionally important in oat. This is due, in part, to natural breeding barriers that Mother Nature has put in front of us. Let me explain.

In our companion study (Tinker, *et al.* 2022), we sought to characterize gene regions (QTL) with major effects on some of the most important oat traits, including heading date, height, lodging, groat percentage, and macronutrient content. We performed this study in five different populations created from nine different parents that differed widely in their adaptation and performance. One result was that we saw a highly significant correlation between the positions of these QTL and genes (usually from other species) that were known to affect these traits. While these results can be used directly in molecular breeding programs, they also provide a new, publicly available starting point for discovering and proving underlying genetic mechanisms.
So, what are these natural breeding barriers? Put simply, barriers are caused by chromosome rearrangements that differ between parents. When sister-chromosomes from different parents of a cross don't have the same order of genes, those chromosomes may refuse to pair up, or they may pair up in unusual ways. This can prevent the normal series of crossovers that breeders rely on to recombine the best traits from both parents (see Fig.1). Such barriers are often present when breeders make crosses with wild species, but, in the case of oat, we showed that different chromosome rearrangements may be common even among adapted, cultivated oat lines. We

were able to observe two of these barriers in our crosses (Tinker, *et al.* 2022), but we and other colleagues have observed evidence for chromosome rearrangements in previous work as well.

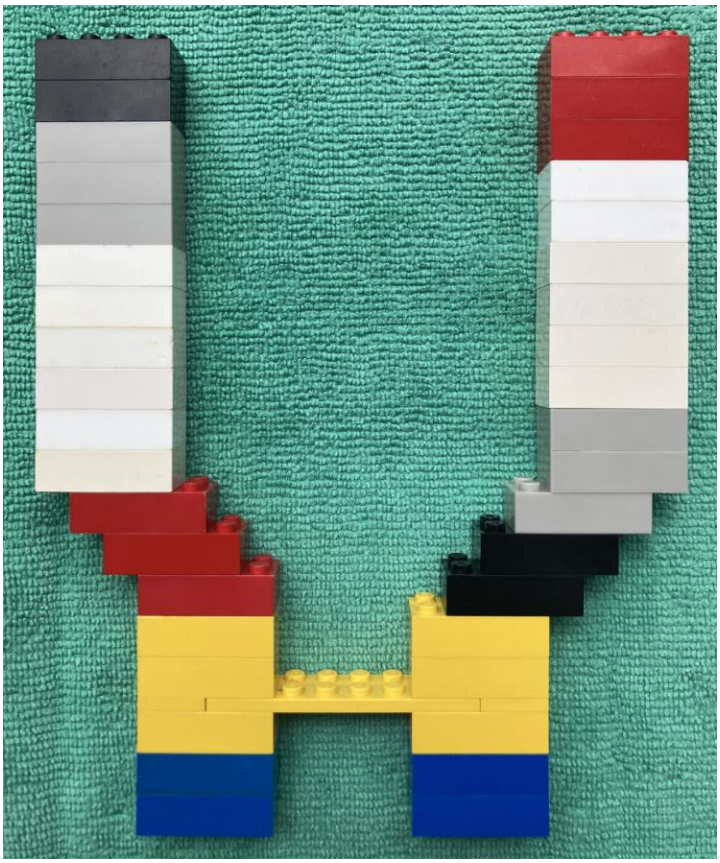

Figure 1. Hypothetical Lego-rendering of chromosome 7D at meiosis in an F1 hybrid between two oat parents. Different coloured blocks represent approximations of the "mosaic" ancestral orthologues (groups of genes whose order is unchanged throughout the ancestral oat family) as described by Kamal, *et al.* (2022). On the left is a "normal" chromosome 7D from oat variety 'Hidalgo'. On the right is chromosome 7D from variety 'TX07CS-1948', with a large inverted section. The bridge between the yellow blocks is an example of a crossover that can occur at a location where the chromosomes pair normally. Crossovers are prevented in the other regions because the inverted gene order prevents pairing. The result is that breeders cannot recombine genes from these two different parents if they are in the red, white, black, or grey regions.

While these rearrangements will sometimes present barriers, they may also be used to breeders' advantage by preserving the order of genes that work well together; for example, if breeders do manage to assemble a "better package of genes" within a natural chromosome rearrangement, then nature will help to preserve that package in future crosses. Knowing the positions of rearrangements may also help to guide gene introgression from wide crosses. Characterizing all such rearrangements will be part of our forthcoming "oat pan genome", and knowing their exact locations and the genes that they contain will soon become just another tool for building better oats.

Who dunnit?

Like a good detective story, the mystery of the mosaic oat genome was solved by visionary investigators, by hard-working deputies, and by helpful bystanders. Both papers contain detailed lists of authors and their contributions, but here I want to expand on the personal side of the story, and also to recognize that the story goes way beyond these important contributions. The main plot started to unfold six years ago when CropTailor and the ScanOats Industrial Research Centre in Sweden hatched plans to sequence the oat cultivar 'Sang'. At the time, genome sequencing was prohibitively expensive, but the ScanOats sequencing project (led by Nick Sirijovski) persevered, and eventually obtained a draft genome sequence and pursued single gene mapping studies on an EMS population developed by CropTailor. Over the following

years, ScanOats worked closely with researchers at other institutions to develop specialized bioinformatics analyses and to build biological stories. A team at IPK Germany, led by Martin Mascher, assembled the draft sequences into validated full-scale chromosomes, and compared these to other genomes. This work was guided by "Hi-C" analysis, as well as by the locations of high-density markers on the oat consensus map (Bekele, et al. 2018), building on many years of hard work by the oat community. A team at the Helmholtz Center Munich, led by Manuel Spannagl, conducted detailed annotations of genes and many other bioinformatics analyses of genome structure and gene families. An Australian team of cereal proteogenomics researchers, led by Michelle Colgrave at the CSIRO and Edith Cowan University, performed detailed analyses of oat seed storage proteins, substantiating the safety of oats in gluten-free diets and demonstrating that proteomics is ready to take oat research to the next level. Along the way, we at Agriculture and Agri-Food Canada, and our colleagues Rick Jellen and Jeff Maughan at Brigham Young University, engaged to contribute and interpret the genome sequences of the wild oats *Avena insularis* and *Avena longiglumis*, which were key to unraveling oat genome evolution. With early access to the Sang genome, my team was able to initiate the reference-based QTL analysis that we reported in the companion paper (Tinker, *et al.* 2022), and also to perform detailed recombination analysis that helped to demonstrate the presence and effects of chromosome rearrangements.

The Nature paper and its extensive set of supplementary materials contains several biological stories, and many clever bioinformatic and graphical interpretations that will guide oat research for years to come. For this we thank many students, postdocs, and research associates, whose contributions begin at the top of the author list. For those unfamiliar with author lists on major collaborative papers, the contributions of principal investigators (PIs) are prioritized from the bottom of the author list. For me, this was my first opportunity to contribute to a Nature paper, and the first time I coordinated a joint publication. The rigor of the Nature review process was eye-opening for me, and I can attest that the PIs earned their spots on the author list! The time between submission and publication was long, but it was anything but leisurely, with many late-night deadlines for submissions, revisions, and rebuttals along the way.

Finally, I want to acknowledge the many oat researchers who are NOT on these papers, but whose work was also instrumental in getting to this place. I can't name everyone, but they include almost every oat worker that I know. They include those who helped build the foundations of oat genome mapping, those who are devoted to managing and disseminating data and germplasm, breeders who paused their important work to contribute germplasm and insight, cytogeneticists and taxonomists whose skill and patience were key to understanding the oat family tree, chemists and nutritionists who figured out why oat is so healthy, pathologists who patiently parse host-plant interactions, agronomists and physiologists who make oat a competitive crop, and farmers, government, and industry personnel who put faith and investment into science. I also want to recognize those colleagues with overlapping projects who may feel temporarily setback by these publications - I have been in your shoes. Your work is important, and I know you will find a way to make it even more important now. So, let the fun begin!

References:

Bekele, W.A., Wight, C.P., Chao, S., Howarth, C.J., and Tinker, N.A. (2018) Haplotype based genotyping-by-sequencing in oat genome research. Plant Biotechnology Journal, **16**: 1452-1463. https://doi.org/10.1111/pbi.12888

Kamal, N., Tsardakas Renhuldt, N., Bentzer, J., Gundlach, H., Haberer, G., Juhász, A., Lux, T., Bose, U., Tye-Din, J.A., Lang, D., van Gessel, N., Reski, R., Fu, Y.-B., Spégel, P., Ceplitis, A., Himmelbach, A., Waters, A.J., Bekele, W.A., Colgrave, M.L., Hansson, M., Stein, N., Mayer, K.F.X., Jellen, E.N., Maughan, P.J., Tinker, N.A., Mascher, M., Olsson, O., Spannagl, M., and Sirijovski, N. (2022) The mosaic oat genome gives insights into a uniquely healthy cereal crop. Nature. https://doi.org/10.1038/s41586-022-04732-y

Tinker, N.A., Wight, C.P., Bekele, W.A., Yan, W., Jellen, E.N., Tsardakas Renhuldt, N., Lux, T., Spannagl, M., and Mascher, M. (2022) Genome analysis in *Avena sativa* reveals hidden breeding barriers and opportunities for oat improvement. Communications Biology, **5**: 474. https://doi.org/10.1038/s42003-022-03256-5