



Construction and analysis of pangenomic graphs in soft wheat from 14 public genomes

Djémilatou Ouandaogo, Vincent Pailler, Hélène Rimbert, Pierre Sourdille,
Pauline Lasserre-Zuber

► To cite this version:

Djémilatou Ouandaogo, Vincent Pailler, Hélène Rimbert, Pierre Sourdille, Pauline Lasserre-Zuber. Construction and analysis of pangenomic graphs in soft wheat from 14 public genomes. JOBIM 2025, Jul 2025, Bordeaux, France. , pp.1. hal-05194677

HAL Id: hal-05194677

<https://hal.inrae.fr/hal-05194677v1>

Submitted on 31 Jul 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Construction and analysis of pangenomic graphs in soft wheat from 14 public genomes

Djémilatou OUANDAOGO, Vincent PAILLER, Hélène RIMBERT, Pierre SOURDILLE and Pauline LASSERRE-ZUBER

Context

Bread wheat (*Triticum aestivum*) is a vital staple crop cultivated worldwide. However wheat has a large hexaploid (15 Gb, AABBDD, $2n = 6x = 42$) and complex genome with 85% of transposable elements (TE) and numerous structural variations (SV) that hampers genetic gain. To accelerate wheat breeding & better understand its genetic diversity, we constructed a wheat pangenome graph from 14 genomes using Minigraph-Cactus [1].

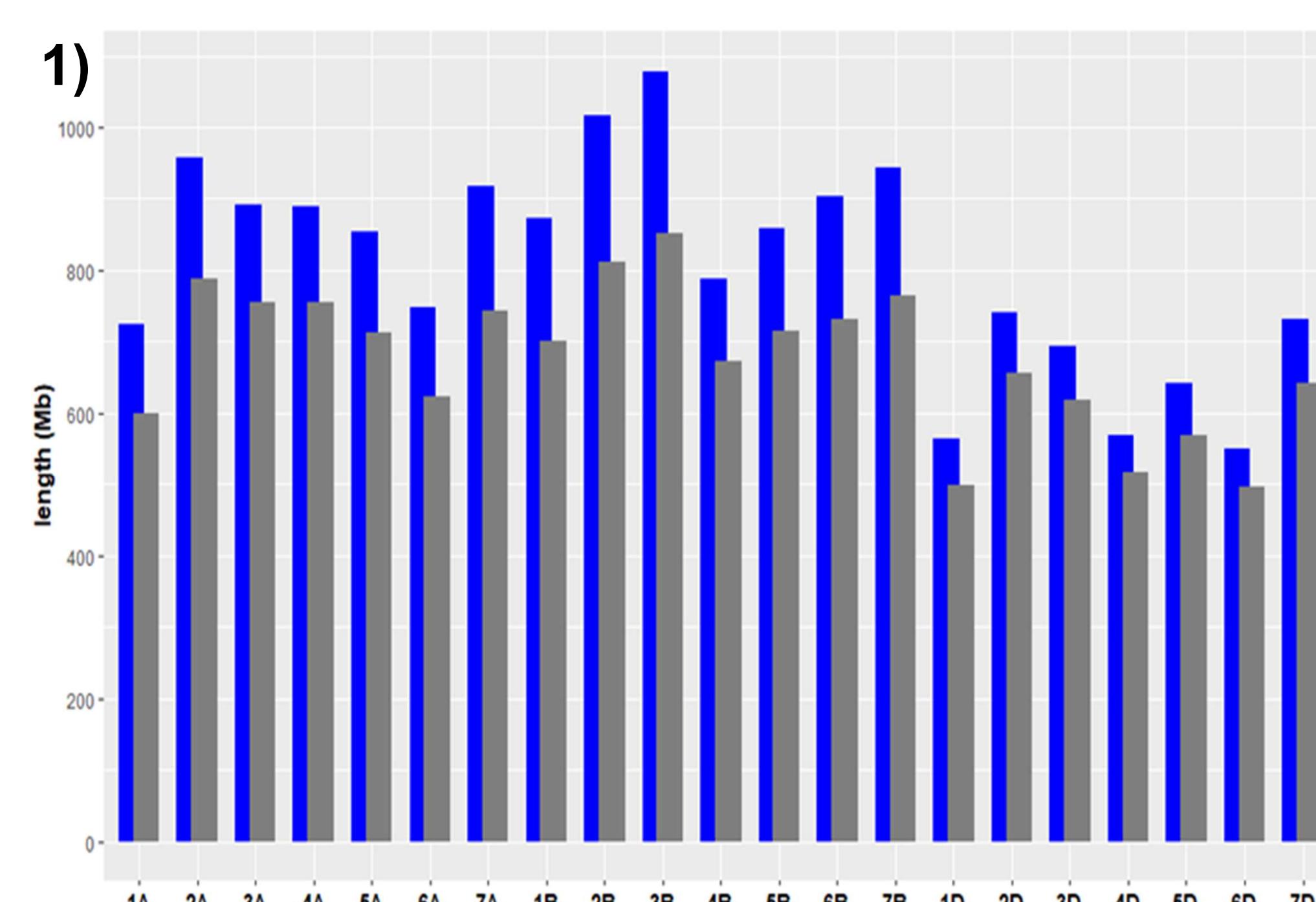
Material

- **13 public genomes:** Chinese Spring [2], Renan [3], Jagger, Julius, ArinaLrFor, LongReach_Lancer, Stanley, Landmark, Mace, Norin61, PI190962 (spelta), SY Mattis [4], zang1817 [5].
- + 1 genome assembled and annotated by our team: Courtot
- **public gene annotations**

Method

- **Minigraph-Cactus** [1]: pangenome graphs construction
- **vg**: VCF files and sub-graph extraction
- **GrAnnoT** [6]: gene annotation transfer
- **CLARITE_smk**: TE annotation [7] https://forge.inrae.fr/umr-gdec/clari-te_smk
- **tidyverse, ggplot2, UpSetR**: satistics and visualization

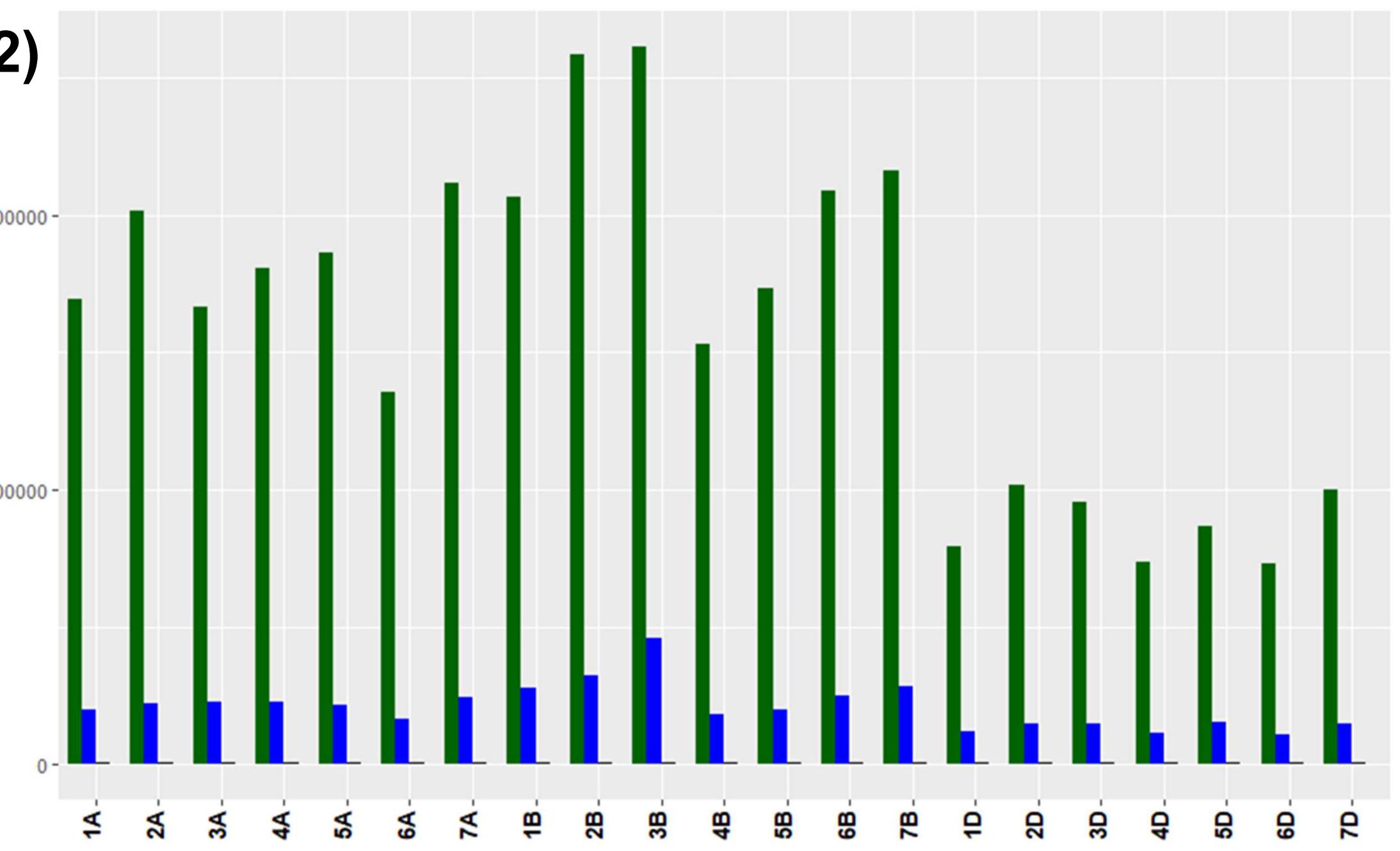
Results



2) Structural Variations (SV) count per chromosome by length (bp):

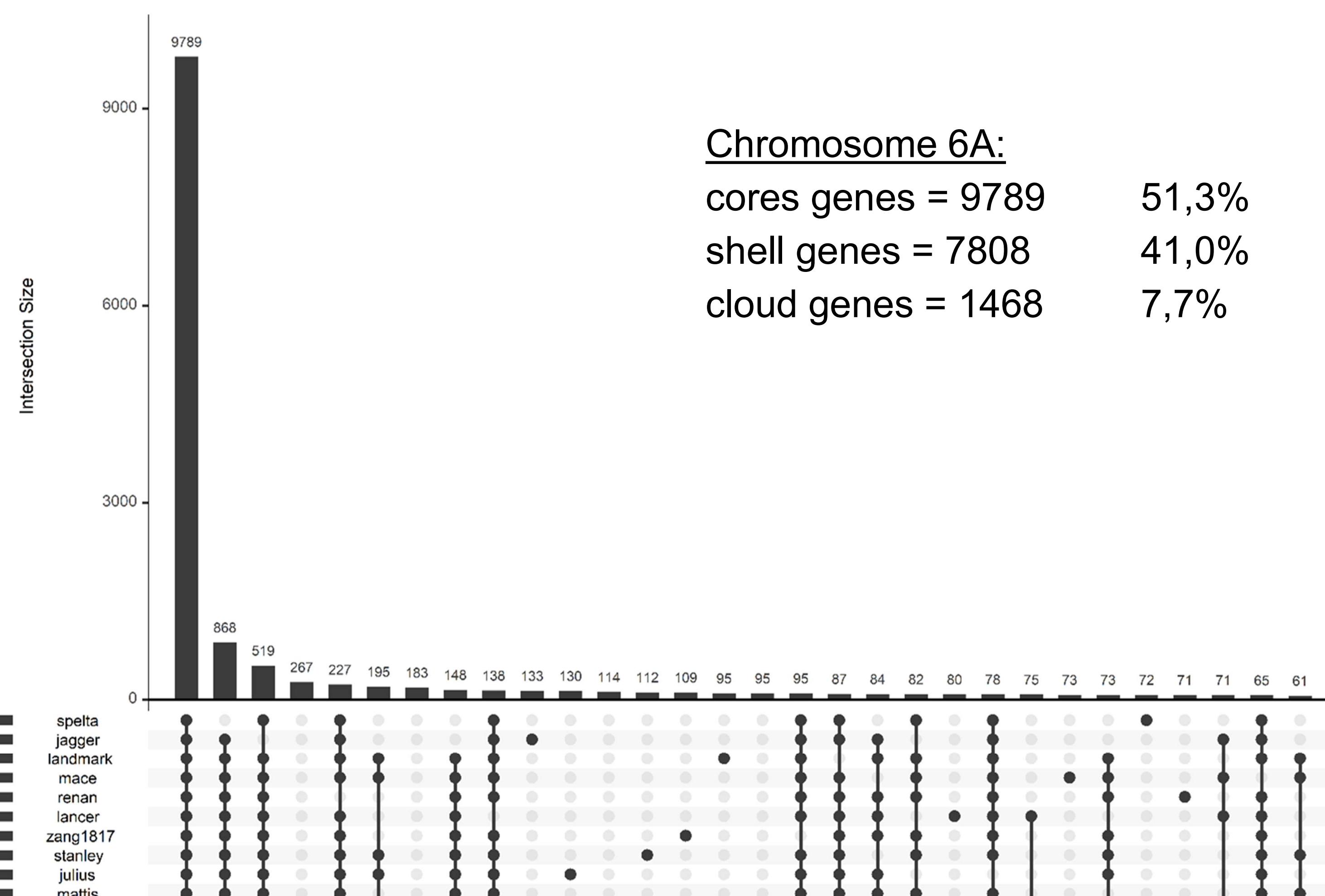
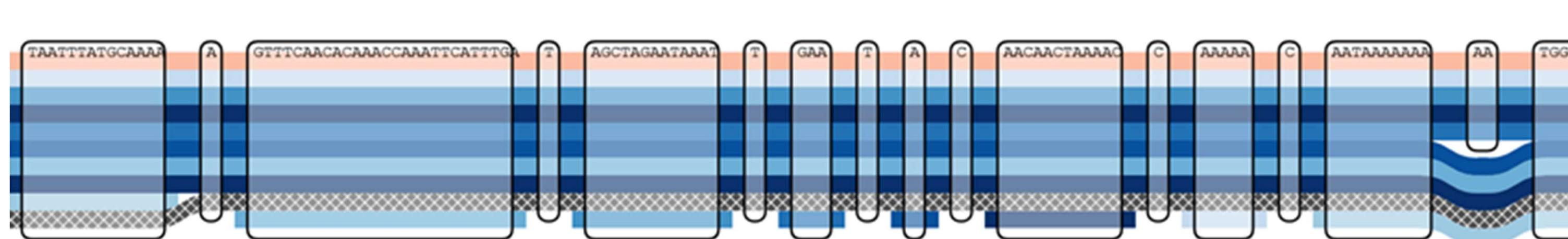
SVs were mostly SNP. 2B, 3B, and 7B chromosomes showed higher amounts of SVs, explained by higher sizes and repeats content of sub-genome B.

2)

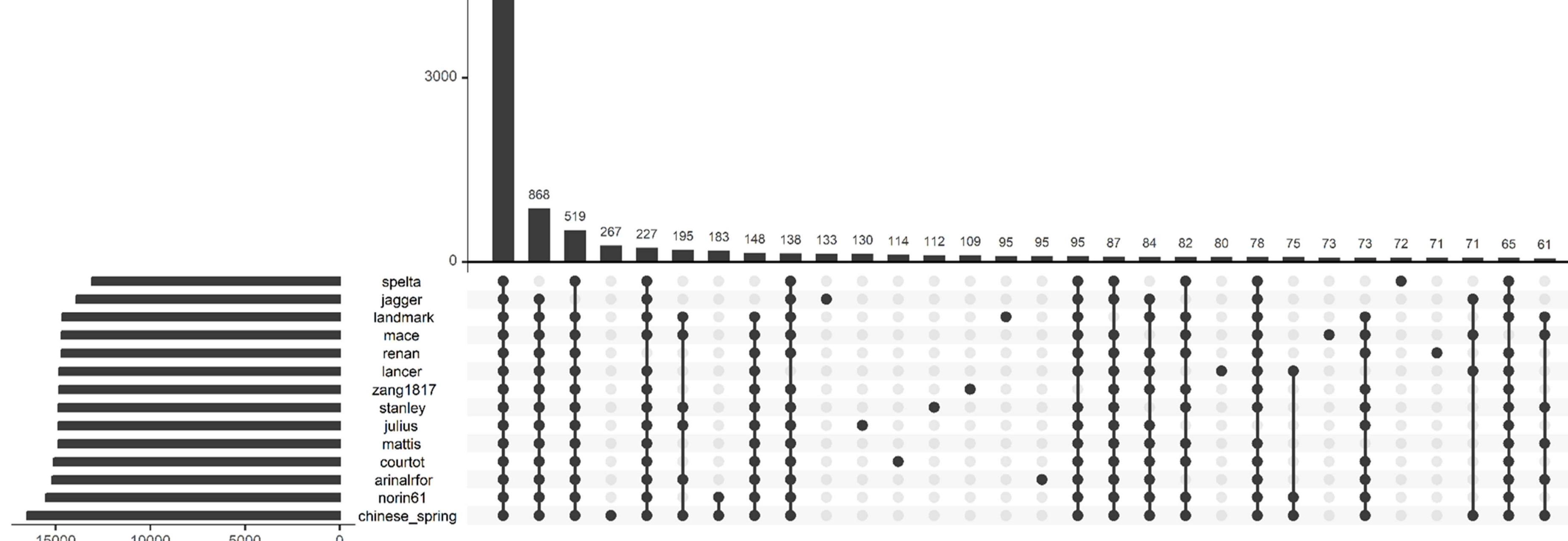
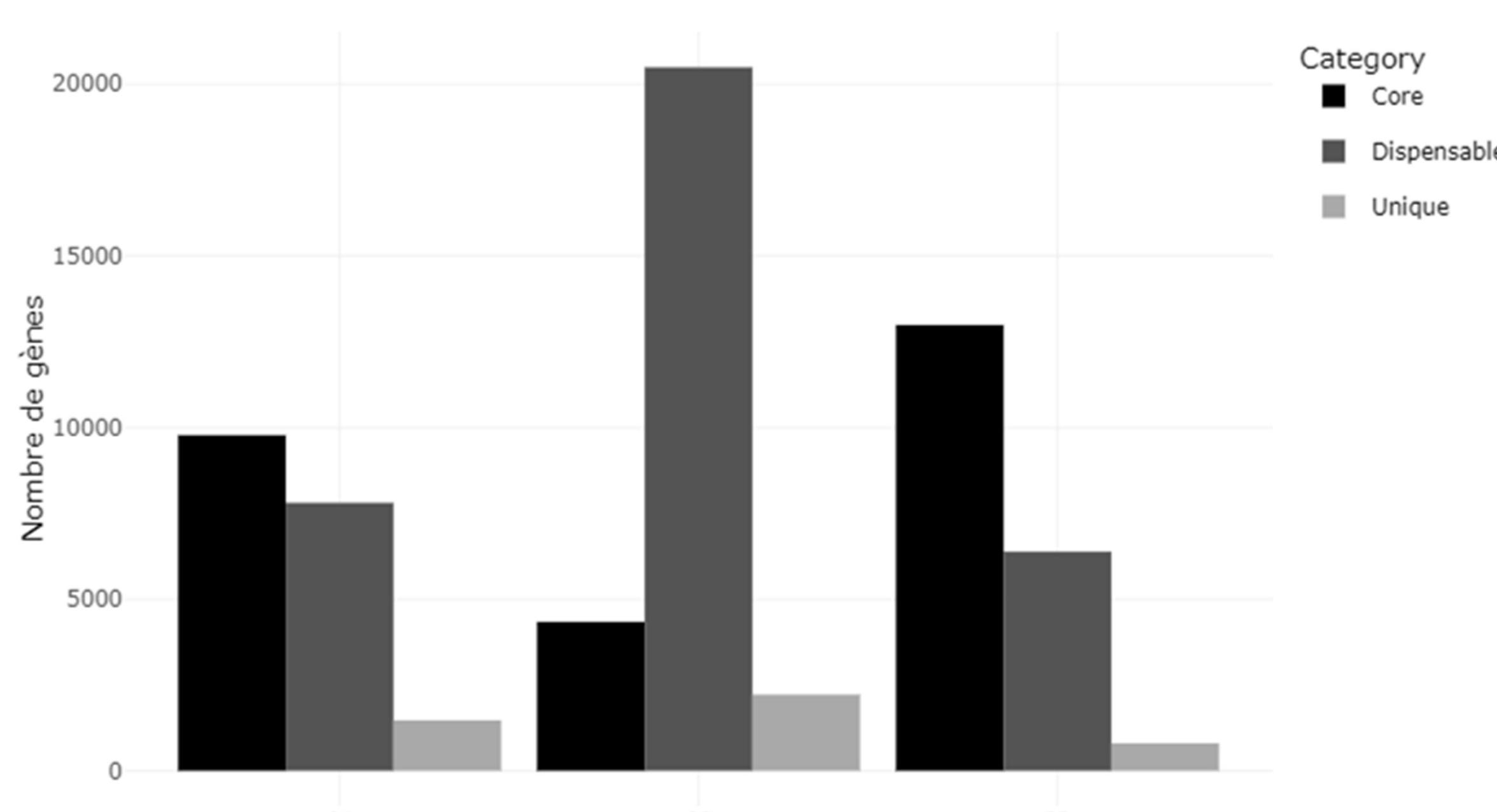


1) Chromosome sizes in linear genomes (REF) vs. Pangenome Graphs (PG):

Chromosomes in PG were consistently longer than those in REF (Chinese Spring), being more exhaustive by integrating variable sequences.



3) Gamma gliadin 1A locus sub-graph (SequenceTubeMap)



Perspectives and Conclusion

- Core, shell, and cloud genomes will be identified at the pangenome level and characterized for genes functions, notably shell and cloud genes that could be related to adaptive traits to climate change.
- Graph augmentation & sub-graph extraction will permit key agronomic loci analyses for genetic variability, such as for glutenin locus 1ABD (3)).
- TE will be analysed for their dynamic and detection of insertion variations will be applied using panREPET pipeline [8].
- Structural variants calling will be applied directly from the graph by aligning new sequencing data to supply association studies.
- Pangenome graphs offer a richer view of wheat genomic diversity than linear references, providing valuable insights for breeding.

