

# Chromosome-level genome assembly of the functionally extinct northern white rhinoceros (*Ceratotherium simum cottoni*)

Gaojianyong Wang<sup>1</sup>, Björn Brändl<sup>1,2</sup>, Camilo Jose Tomas Hernandez-Toro<sup>1</sup>, Jeanne Loring<sup>3</sup>, Marisa Korody<sup>4</sup>, Franz-Josef Müller<sup>1,2</sup>

1. Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany  
2. Department of Psychiatry and Psychotherapy, Christian-Albrechts University Kiel, Kiel, Germany

3. The Scripps Research Institute, La Jolla, USA  
4. San Diego Zoo Wildlife Alliance, Escondido, USA

## Abstract

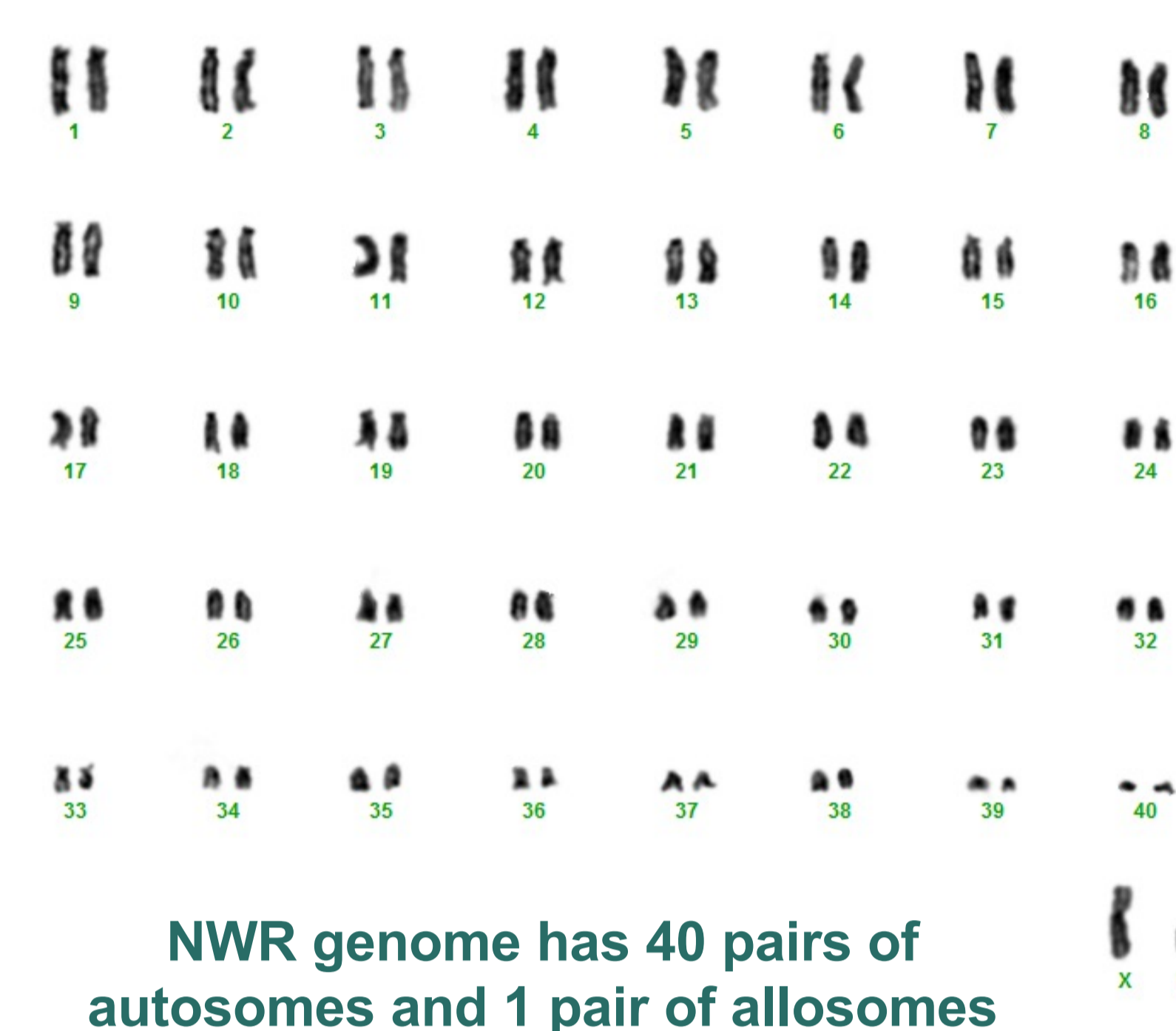
Traditional conservation methods have failed to save the northern white rhinoceros (NWR: *Ceratotherium simum cottoni*), leaving only two non-reproductive females. Scientists are exploring assisted reproductive technologies (ARTs), including artificial gametes produced from induced pluripotent stem cells (iPSCs) and somatic cell nuclear transfer, together with Southern white rhinoceros (SWR: *Ceratotherium simum simum*) surrogate mothers as potential methods to save functionally extinct NWR. The development of ARTs requires a high-quality reference genome. Fibroblast and iPSCs were cultured from a male NWR 'Angalifu' and deeply sequenced using four complementary platforms. We developed an assembly strategy to generate a chromosomal-level (2n = 82) NWR reference genome, including 40 acrocentric autosomes, the X and partially resolved Y chromosomes, as well as the mitochondria genome, that meets or exceeds the metrics proposed by the Vertebrate Genome Project (VGP). The high quality genome not only confirms the gene order and chromosome conservation between horse and rhino, but also shows that no chromosomal-level structure variants exists between NWR and SWR. Although gaps still exist in the generated NWR reference genome, recent development of ONT ultralong sequencing techniques has demonstrated that a gap-less telomere-telomere NWR genome assembly is highly achievable.

## I. Functionally extinct NWR

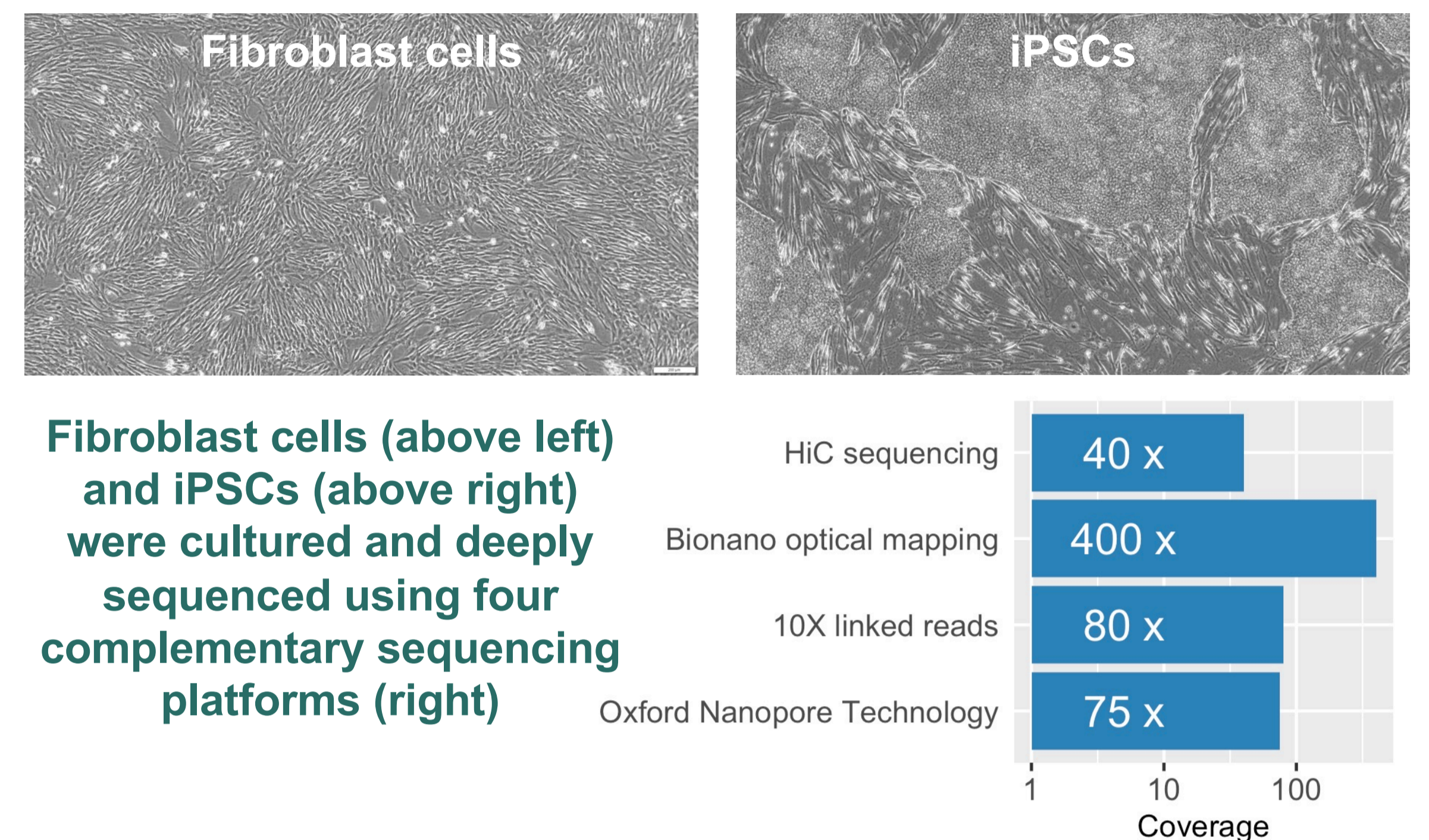


NWR is functionally extinct with only two non-reproductive females remaining

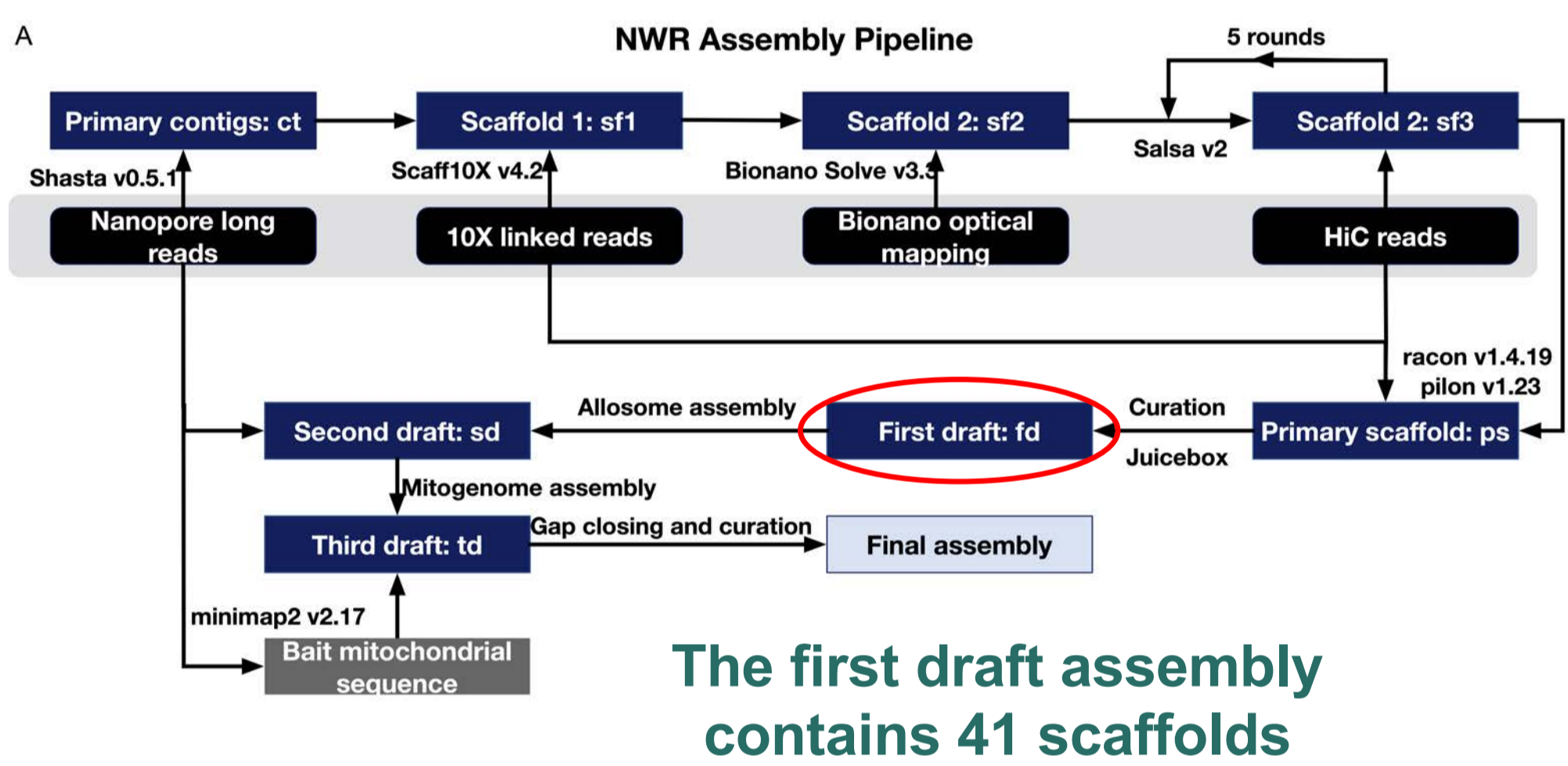
## II. Karyotype ("Angalifu" male)



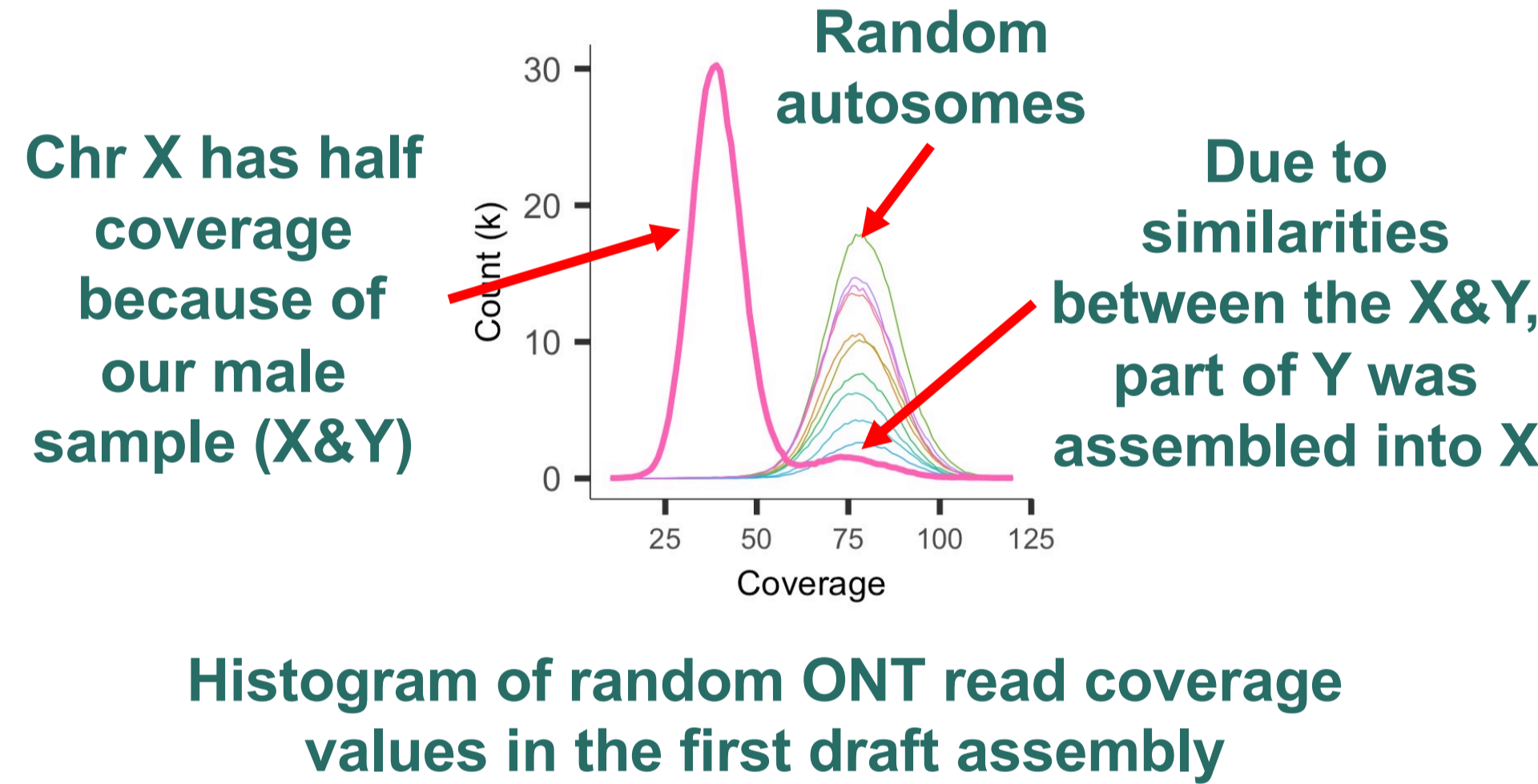
## III. Cells and sequencing



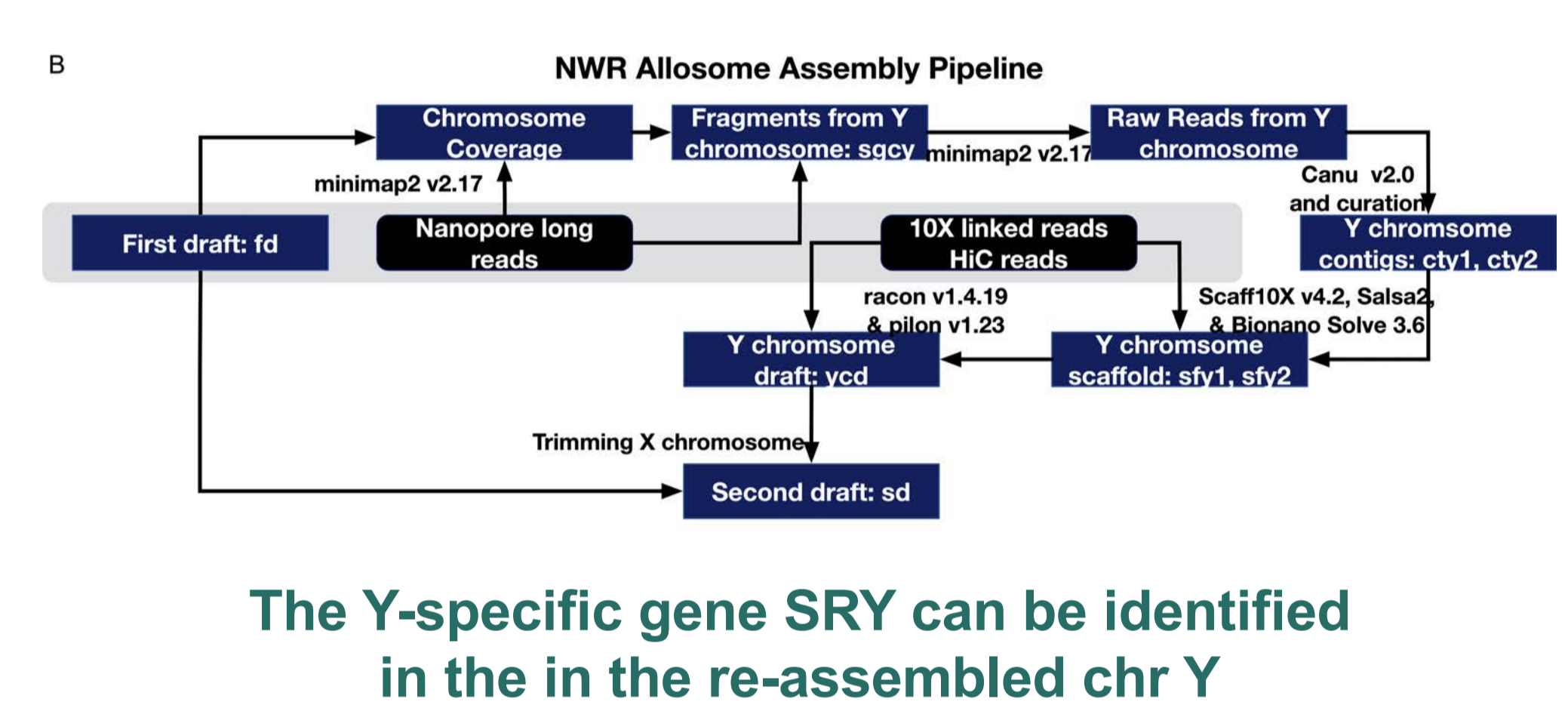
## IV. Main assembly pipeline



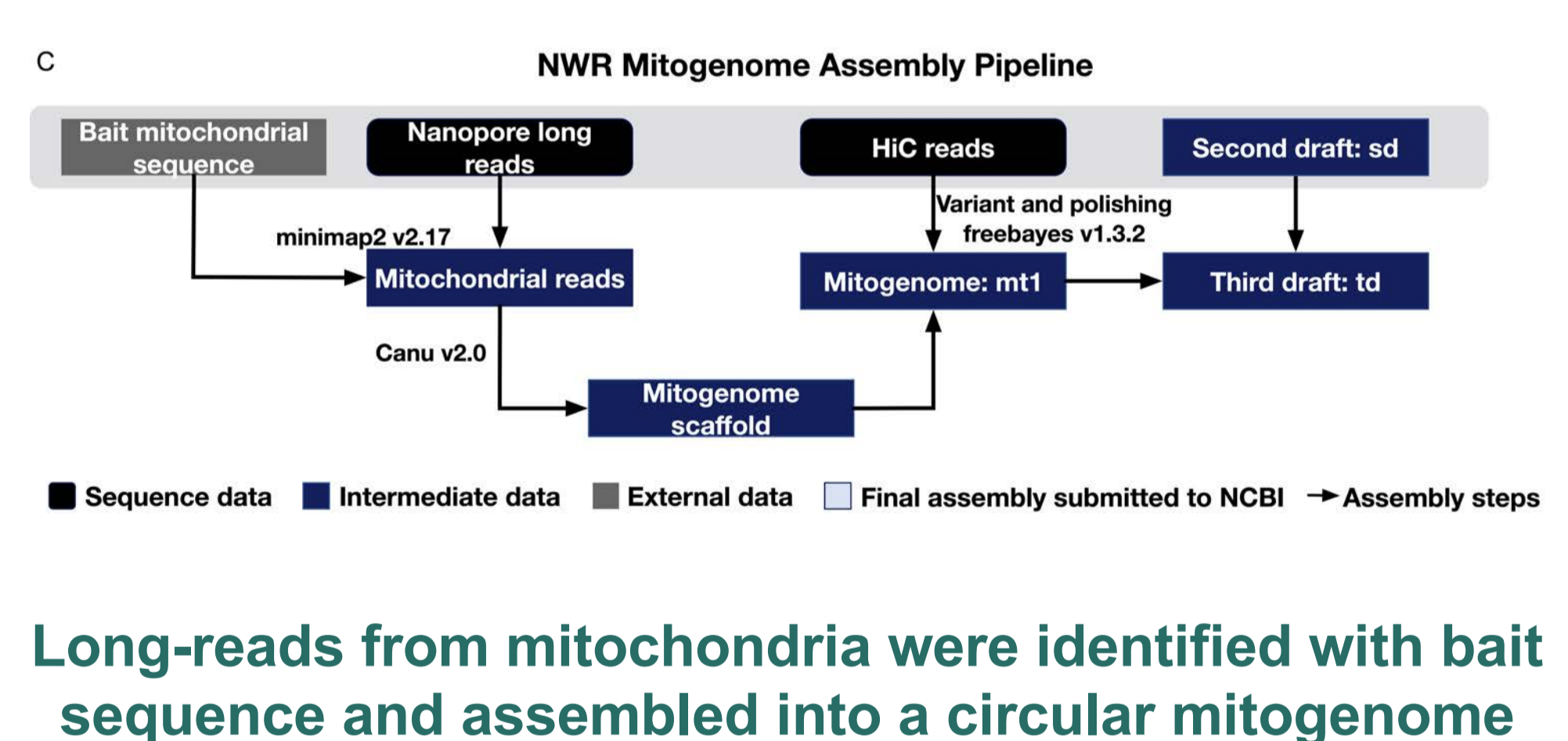
## V. ONT long-read coverage profile



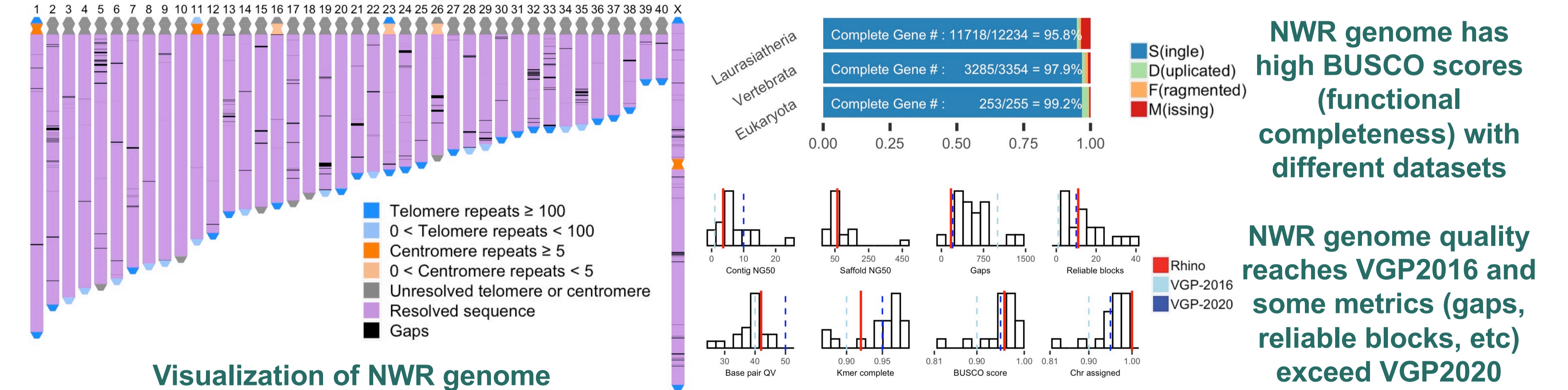
## VI. Allosome assembly pipeline



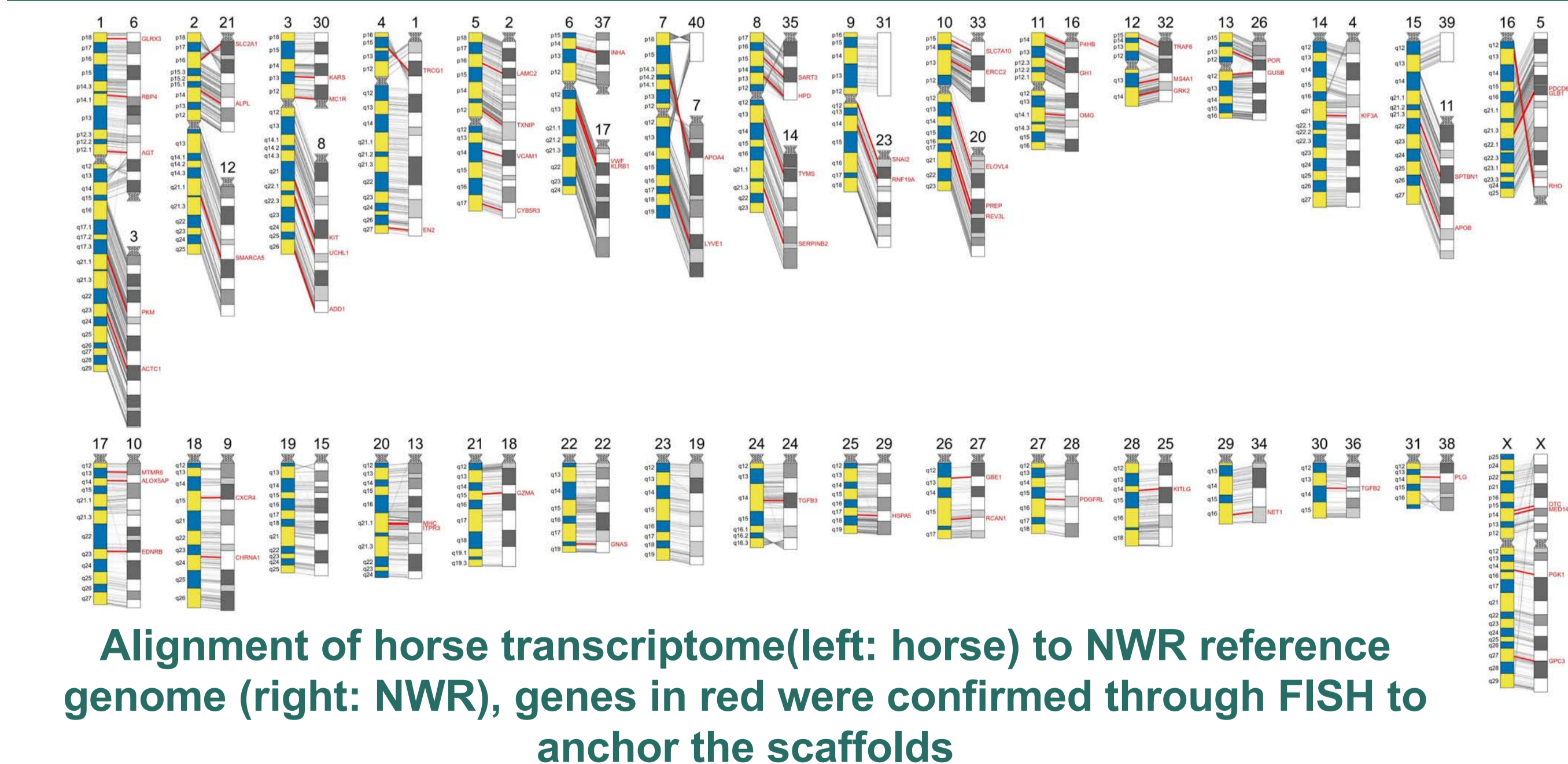
## VII. Mitogenome assembly pipeline



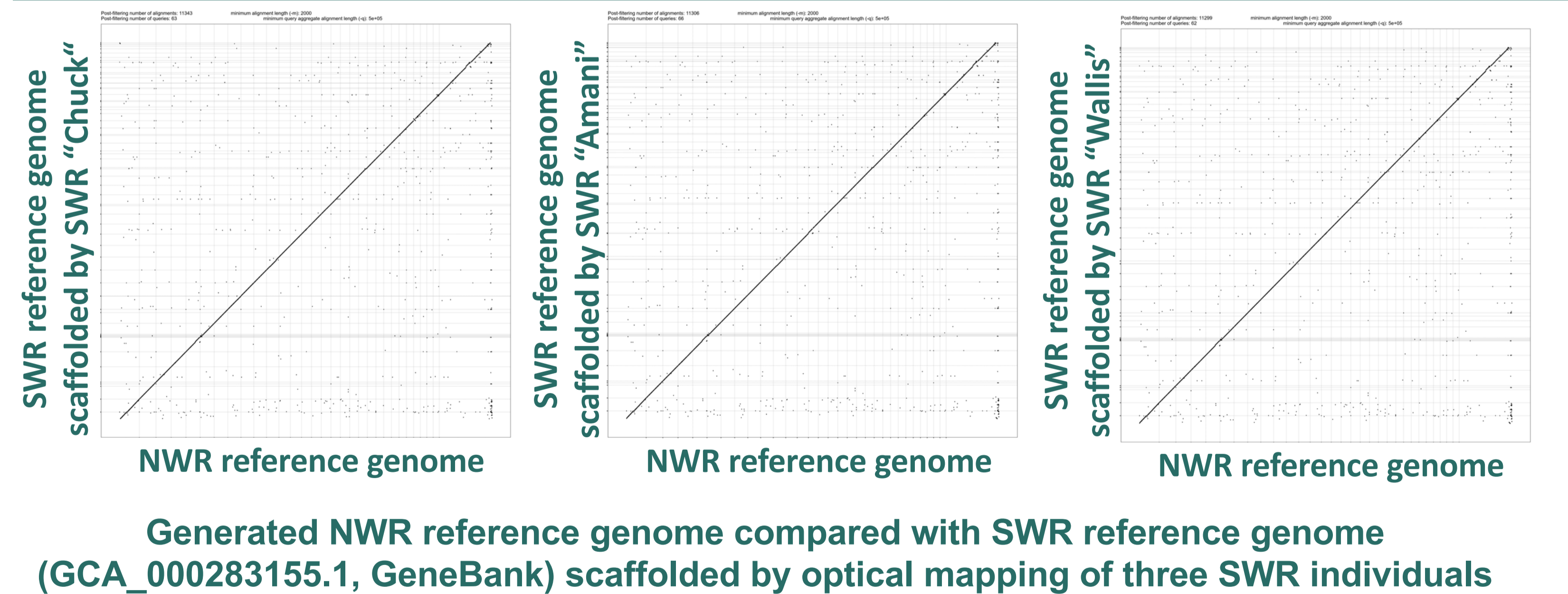
## VIII. NWR reference genome and its quality



## IX. Comparison with horse



## X. Comparison with SWR



## XI. Conclusions

- The NWR genome reaches chromosomal-level assembly for autosomes and X chromosome, contig-level assembly for Y chromosome, and complete assembly of mitochondrial genome.
- The quality of the genome was evaluated through metrics proposed in the VGP. All metrics have reached the VGP2016 standard and some have exceeded VGP2020 standard.
- The high-quality NWR reference genomes confirms that gene order and chromosomes are conserved between horse and NWR. In addition, no chromosomal-level structure variants were identified between NWR and SWR.
- We demonstrate the possibility to generate a high quality, even gapless, reference genome for NWR and possibly for other endangered species with nanopore sequencing

## XII. References

- Korody, M.L., et al., Rewinding Extinction in the Northern White Rhinoceros: Genetically Diverse Induced Pluripotent Stem Cell Bank for Genetic Rescue. *Stem Cells and Development*, 2021. 30(4): p. 177-189.
- Rhie, A., et al., Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 2021. 592(7856): p. 737-746.
- Tunstall, T., et al., Evaluating recovery potential of the northern white rhinoceros from cryopreserved somatic cells. *Genome Research*, 2018.
- Simão, F.A., et al., BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 2015. 31(19): p. 3210-3212.
- Shafiq, K., et al., Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nature Biotechnology*, 2020. 38(9): p. 1044-1053.

**Acknowledgements:** We want to thank the IT department for their support in this research. I would like to extend my appreciation to Karl Hong, Teri Lear, Marlys Houck, Thomas Kreitter, Helene Kretzmer, Romualdas Smicius, Remy Schwab, and everyone who has provided invaluable advice and support to this research