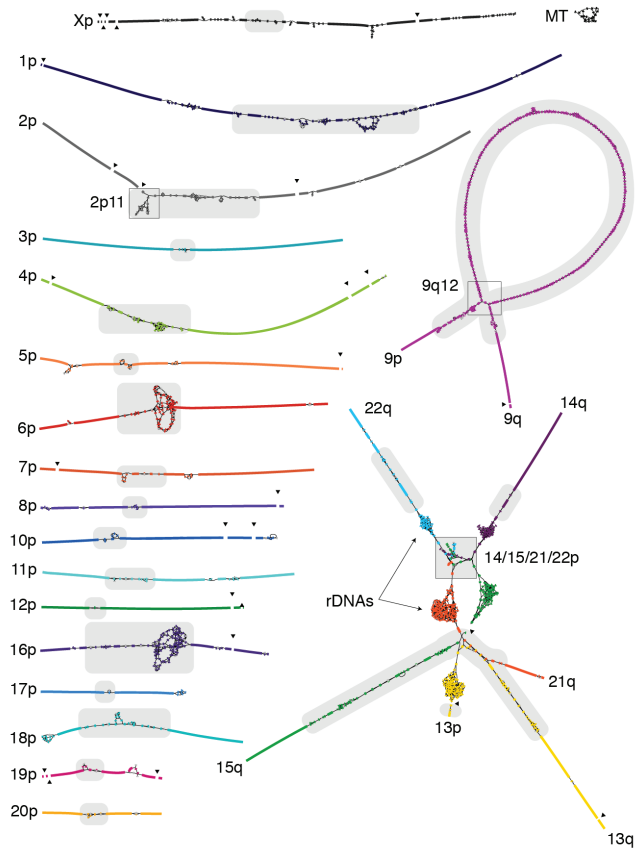
PACBIO®

PUBLICATION AT A GLANCE

THE COMPLETE SEQUENCE OF A HUMAN GENOME

WHY IS THIS PAPER IMPORTANT?

This landmark study by members of the *Telomere-to-Telomere Consortium* is the first fully **complete** assembly to be produced 20 years after the initial drafts of the human genome.



Summary	GRCh38p13	CHM13v1.1	±%
Assembled bases (Gbp)	2.92	3.05	+4.5%
Unplaced bases (Mbp)	11.42	0	-100.0%
Gap bases (Mbp)	120.31	0	-100.0%
# Contigs	949	24	-97.5%
Ctg NG50 (Mbp)	56.41	154.26	+173.5%
# Issues	230	46	-80.0%
Issues (Mbp)	230.43	8.18	-96.5%

"High accuracy long-read sequencing has finally removed this technological barrier, enabling comprehensive studies of genomic variation across the entire human genome. Such studies will necessarily require a complete and accurate human reference genome, ultimately driving adoption of the T2T-CHM13 assembly presented here."

Nurk, S. et al. (2021) The complete sequence of a human genome. ***bioRxiv***

KEY FINDINGS

Pictured as a high-resolution string graph built directly from PacBio® HiFi reads, the new T2T-CHM13 assembly provides:

- A more complete, representative, and accurate reference than GRCh38
- 200 million bp of novel sequence
- Gapless assemblies of 22 autosomes, X chromosome, and mitochondrial genome
- 2,226 paralogous gene copies, including 115 predicted to be protein coding

WHAT ROLE DID PACBIO TECHNOLOGY PLAY?

The authors utilized multiple sequencing techniques and chose PacBio HiFi reads for the genome assembly for their combination of long read lengths (up to 25 kb) and high accuracy (>99.9%) which enabled the assembly of highly repetitive centromeric satellite arrays and closely related segmental duplications.

GET STARTED

Using the PacBio whole genome sequencing for *de novo* assembly application, a 3 Gb genome can be sequenced using three SMRT® Cells targeting 30-fold coverage and can be assembled in less than one day.

Learn more at pacb.com/wgs-human