AN EXPEDITION TO COMPLETION

THE WORLD OF THE HUMAN GENOME IS HUGE

31 YEARS TO NAVIGATE

3.2 BILLION BASE PAIRS

92% COMPLETE... UNTIL NOW

WE'VE EXPLORED MANY DESTINATIONS ALONG THE WAY

The *Human Genome Project* started the journey, providing profound insights into biology. It revealed 92% of the human genome.
At the time, this was considered *finished*, but in reality...

8% OF THE TOTAL BASES CONTAINED IN THE HUMAN BODY WERE MISSING

- **1990** — *Human Genome Project* launched in the US
- **1996** — First human gene map established
- **2000** — Draft version of human genome sequence completed
- **2003** — *"Finished"* version of human genome sequence published (92% of the complete version)

STRUCTURAL VARIATIONS

AS WE TRAVELLED, WE KNEW THERE WAS MORE TO DISCOVER

After the *Human Genome Project*, we thought the journey was over. But, we realized that repeats and structural variants had thrown us off course. We couldn't traverse the last 8% of the genome without seeing the full length of these repeats and total expanse of variants.

We needed a vehicle that could read long stretches of DNA, and put the repetitive sections in the right place.

REPEATS

SO, WE SET OFF TO SEE MORE...

The *T2T Consortium* worked extensively to sequence CHM13, using multiple different technologies.

The first T2T assembly of chromosome X (11) relied on ONT sequencing to create a backbone that was then polished with other technologies.

But we weren't quite there...

LONG-READ SEQUENCING PAVED THE ROAD TO THE END

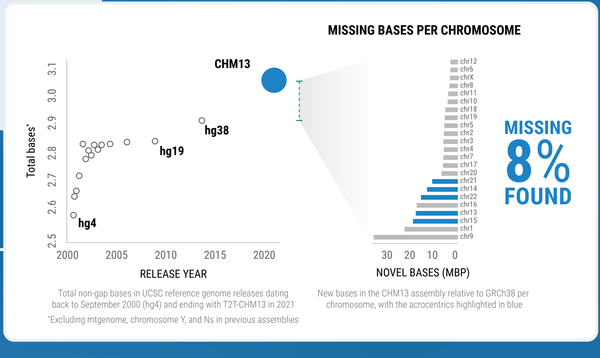
Researchers used HiFi reads to enable assembly of highly repetitive centromeric satellite arrays and closely related segmental duplications (12, 22, 29).

THE RESULT?

In 2021, the *T2T Consortium* assembled the first gapless human reference genome, including all centromeric regions and the entire short arms of five human chromosomes.

The T2T-CHM13 reference assembly impressively removed a barrier that had eluded sequence-based analysis for 20 years and **revealed the missing 8% of the human genome.**

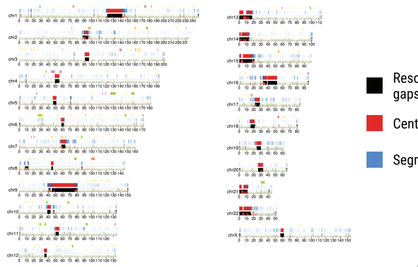
ALL WITH THE HELP OF PACBIO LONG-READ SEQUENCING



THE ASSEMBLY WAS POLISHED, RESULTING IN AN ESTIMATED **ACCURACY OF 99.999996%** (PHRED SCORE OF Q74)

HIFI-BASED ASSEMBLY OF THE CHM13 GENOME

The T2T-CHM13v1.1 assembly includes gapless telomere-to-telomere assemblies for all 22 human autosomes and chromosome X, comprising 3,054,815,472 bp of nuclear DNA, plus a 16,569 bp mitochondrial genome (CHM13 does not have a Y chromosome).



WHERE WILL WE GO IN THE FUTURE?

The complete, telomere-to-telomere assembly of a human genome marks a new era of genomics. New discoveries will be enabled via access to all regions of the genome and improved mapping for all sequencing technologies.

HiFi sequencing is the essential technology for delivering a gold standard reference genome, opening a world of possibilities.

102-193-640 REV02 12MAR2022