# LITERATURE REVIEW: COMPARISON OF SHORT-READ MAPPING METHODS

AMANDA SHEN

# BACKGROUND AND SIGNIFICANCE

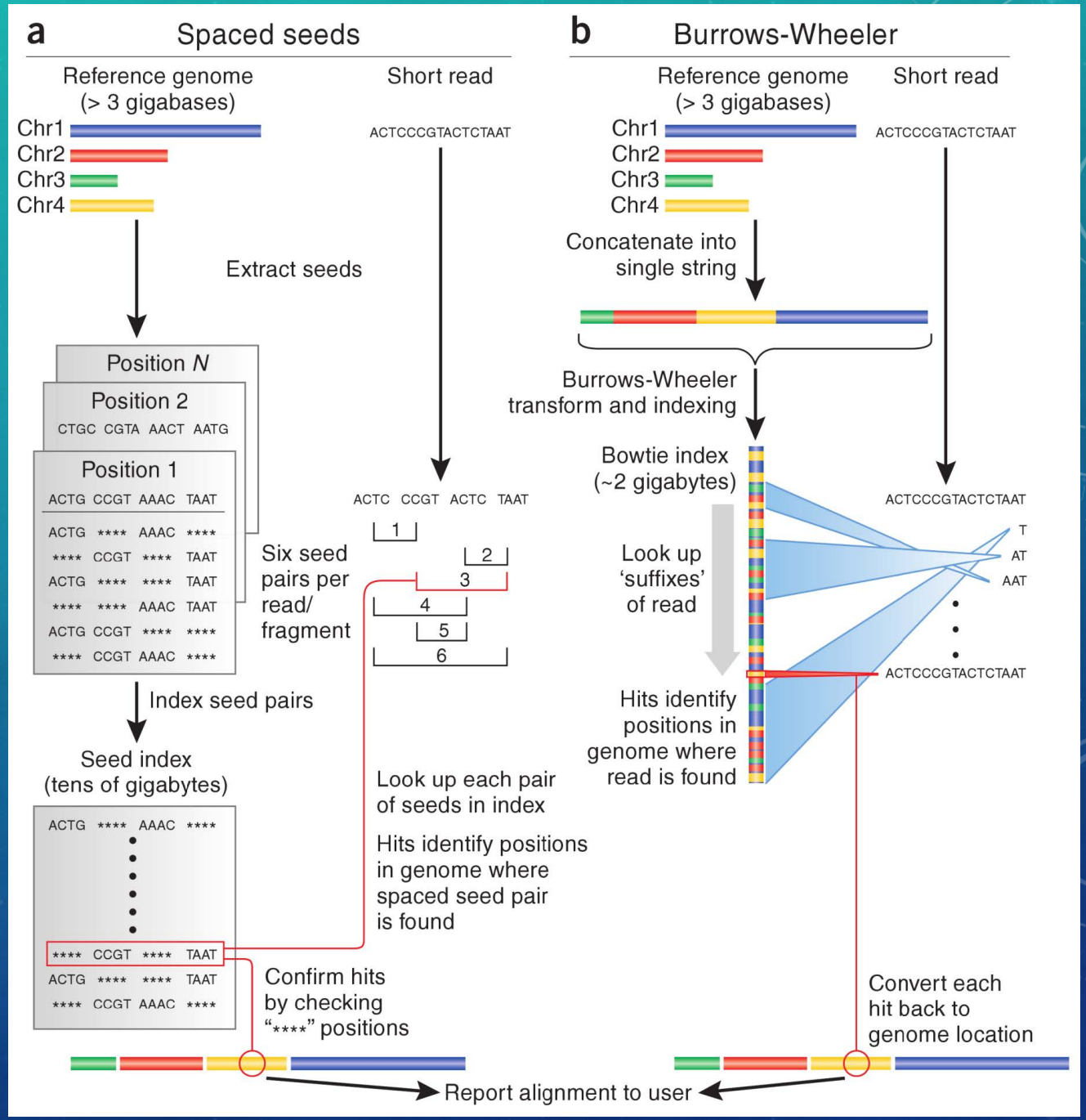
- Personal genomics company: 23andMe

# SHORT-READ MAPPING

- Reads from obtained via shotgun sequencing of the genome of one organism were “aligned” to a reference genome from another closely related organism using a program (Lior Pachter, 2015)

# METHODS

- (Trapnell and Salzberg, 2009)



# METHODS

- Burrows-Wheeler Aligner (BWA) – BWT based tool
- Bowtie – BWT based tool
- SOAP – BWT based tool
- Maq – hashing
- Novoalign – hashing
- …

# STANDARD CRITERIA

- Runtime
- Space efficiency
- Mapping percentage
- Throughput (calculated by dividing the number of reads mapped over the running time.
- SNP (Single Nucleotide Polymorphism) awareness
- False negatives (the intervals missed by the mapper)
- False positives (the intervals returned by the mapper and not included in the gold standard)
- (Hatem et al, 2013)

# SOURCES


- Ayat Hatem, Doruk Bozdag et al (2013). “Benchmarking short sequence mapping tools”. In: *BMC Bioinformatics* 14(184).
- Ben Langmead Cole Trapnell, Mihai Pop and Steven L Salzberg (2009). “Ultrafast and memory-efficient alignment of short DNA sequences to the human genome”. In: *Genome Biology* 10(3).
- E.D. Harrington, C.Y. Mclean et al (2012). “Considerations for the Processing and Direct-to-Consumer Return of Exome Sequences”.
- Matthew Rualo1, Thomas LaFramboise and Mehmet Koyutürk (2011). “Comparative analysis of algorithms for next-generation sequencing read alignment”. In: *Bioinformatics* 27(20), pp. 2790-2796.
- Sophie Schbath, Veronique Martin et al (2012). “Mapping Reads on a Genomic Sequence: An Algorithmic Overview and a Practical Comparative Analysis”. In: *Journal of Computational Biology* 19(6), pp. 796-813.