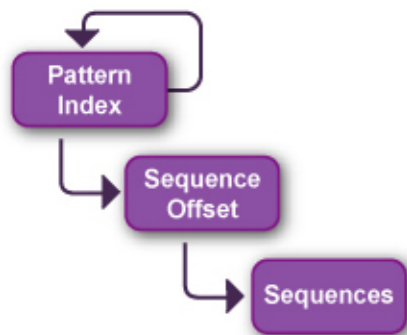


SynaBASE a structured-network pattern database platform



Components of a SynaBASE

The sequences used to build a SynaBASE™ are stored as variable length k-mers structured on the basis of their inter-relationships. This forms the basis of the benefits and advantages of SynaBASE.

SynaBASE is a novel database platform which finds and stores "patterns" rather than fixed length indices.

Computationally demanding analyses such as handling data from 2nd generation sequencers or large scale comparative genomics can be completed 100s of times faster.

SynaBASE key advantages

▶ High-speed analysis

Applications built upon the SynaBASE platform are 100s to 1000s of times faster than conventional software.

▶ Increased sensitivity

Overlapping patterns for each sequence in a genome are indexed within a SynaBASE. As a result, more correct hits are identified.

▶ Flexible platform

Users can utilise existing applications or develop completely new tools specific to their needs using the API.

▶ Long-term ROI

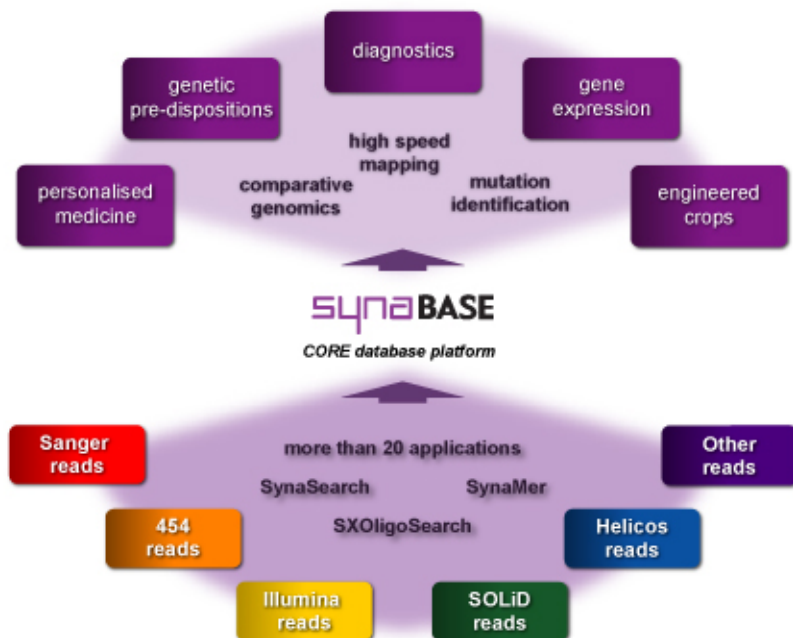
The performance benefits of SynaBASE result from how the data is structured and searched rather than relying on the ever-increasing number of computing clusters or specialised accelerators.



data mining
sequence analysis
comparative genomics
high-speed mapping
phylogenetics



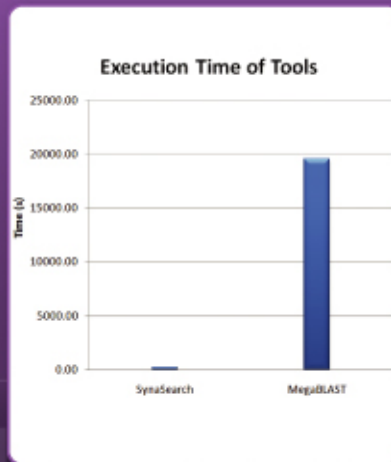
Synamatix approach for next-generation sequence data



Case studies: high-throughput read analysis

Short reads – 454 / Roche

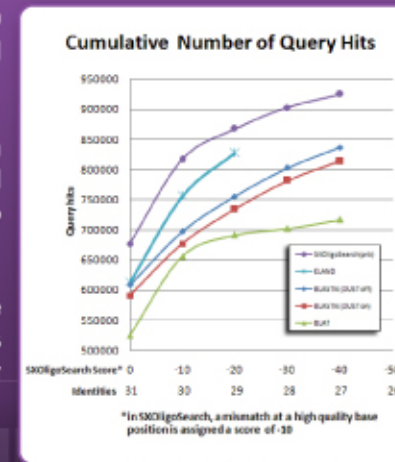
- ▶ By using SynaSearch™ which is built on SynaBASE, 1.68 million 454 reads were mapped back to the Human Genome in 5 hours.
- ▶ In a benchmark study 20,000 sequences with an average length of 100bp were mapped to the Human Genome using MegaBLAST and SynaSearch. SynaSearch performed nearly 80 times faster than MegaBLAST and was more sensitive, with a query coverage of 90.2%, as compared to 77.98% with MegaBLAST.



Microreads – Illumina

A benchmark study was done by mapping Illumina reads with an average length of 31bp against the Human Genome using BLASTN, BLAT, ELAND and SXOligoSearch™ respectively.

- SXOligoSearch performed 1,600 times faster than BLASTN and 100 times faster than ELAND.
- SXOligoSearch placed more than 93% of the reads while BLAT and BLASTN managed only 75.6% and 83.1% respectively.
- SXOligoSearch is more adaptable as it is capable of handling reads longer than 32 bases with any number of mismatches or indels.



Sanger reads - ABI

Detailed performance benchmarking was conducted to compare the analysis time taken by SynaSearch, a sequence alignment application built on SynaBASE, to BLASTz.

In comparison to BLASTz, SynaSearch:

- was 219 times faster
- found 11% more true positives
- had 17% more unique hits to queries
- had a higher specificity
- reported 113% fewer false positives
- had fewer multiple placements per read

