SENCEL PARALIGN™

High-sensitivity genome searches at blasting speed

"GOLD STANDARD" SENSITIVITY

Finding all homologues of a gene in a sequence database requires a search algorithm that is sensitive enough to identify the most remote matches but also fast enough to deliver results in time. Previously, one could run Smith-Waterman searches to get "gold standard" results, but then searches would take ages to finish. Alternatively, one could run BLAST and get results in reasonable time, but accept missing some matches.

Now you can have the best of both worlds, because SENCEL PARALIGN™ performs searches at blasting speed but with sensitivity at the same high level as the Smith-Waterman algorithm. PARALIGN brings the most sensitive search algorithms to incredible speeds by unlocking the potential of modern microprocessors. It exploits parallel processing technology rarely used by other software, like the Velocity Engine™ (AltiVec™) of the PowerPC™ processors in the Power Macintosh™ computers and the MMX™ technology found in Pentium™ processors in standard PCs.

Actually, one can choose between two search methods when running PARALIGN. The first search method is based on the well-known Smith-Waterman algorithm and gives exactly the same high quality results as a conventional implementation; the speed is just accelerated about eight-fold. The second method is based on the new heuristic ParAlign algorithm and runs at the speed of BLAST, but nevertheless delivers results at the same high level of sensitivity as the Smith-Waterman algorithm.

PARALIGN's sensitive nucleotide sequence comparisons enables the detection of unwanted cross-hybridization loci of siRNAs and DNA microarray probes or PCR primers that could pass undetected by other software.
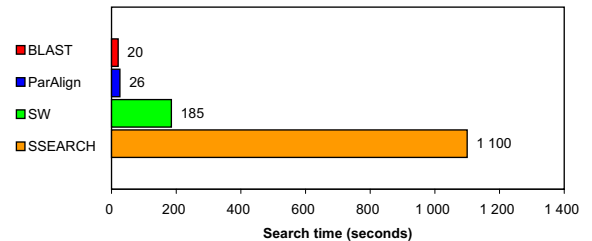
PARALIGN is based on patent pending and published methods (Rognes and Seeberg, 2000; Rognes, 2001) developed by bioinformatics experts at Sencel Bioinformatics™, Rikshospitalet University Hospital and the University of Oslo, Norway.

Visit Sencel on the Internet at **www.sencel.com** to learn more about PARALIGN, or send a message to **sencel@sencel.com** if you have any further questions.

Chart legend:
- BLAST: 20
- ParAlign: 26
- SW: 185
- SSEARCH: 1 100

Search time (seconds)

INCREDIBLE SPEED

The graph above shows the total time used to search the SWISSPROT database with 11 different protein query sequences using BLAST (red; Altschul *et al.*, 1997), Sencel's ParAlign (blue) and Smith-Waterman (SW) (green) methods, and SSEARCH (orange; Pearson, 1991); all on a computer with a Xeon 2.8GHz processor. SSEARCH is the traditional non-parallel Smith-Waterman implementation found in the FASTA package. For further details about this test, see Sencel's website.

PACKED WITH USEFUL FEATURES

- PARALIGN runs on both single- and multiprocessor computers, and clusters using the MPI standard.
- Reads standard sequence databases prepared by the NCBI formatdb tool.
- Writes results in plain text, in html with sequence links, or in an easily parsed tab-delimited format.
- Carries out searches with amino acid or nucleotide sequences as queries or databases sequences.
- Translates query or database nucleotide sequences on-the-fly during searches if necessary.
- Is **free of charge** for academic, non-profit use.
- May be evaluated using free evaluation licenses or online demo searches.
- Can be downloaded directly from Sencel's web site.

LITERATURE

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, 25, 3389-3402.

Pearson WR (1991) Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. **Genomics**, 11, 635-650.

Rognes T and Seeberg E (2000) Six-fold speed-up of Smith-Waterman sequence database searches using parallel processing on common microprocessors. **Bioinformatics**, 16, 699-706.

Rognes T (2001) ParAlign: a parallel sequence alignment algorithm for rapid and sensitive sequence database searches. **Nucleic acids research**, 29, 1647-1652.

SYSTEM REQUIREMENTS
Supported computer systems:
- **Apple® Macintosh® with PowerPC® G4 or G5 processor. Requires Mac OS® X v 10.2 (Jaguar) or later.**
- **PC with Intel® Pentium® MMX™ or compatible processor running Linux® version 2.4 or later.**
- **Intel® Itanium® 2-based computer running Linux® version 2.4 or later.**
- **Intel® Itanium® 2-based computer running HP-UX® version 11i or later.**
- **HP® Alpha®-based computer running Tru64® Unix version 5.1a or later.**