

## Case Study #1: SNP Analysis

In this benchmarking exercise Vmatch, running on a Pentium IV 1.7 GHz PC, outperforms the TimeLogic Decypher BLAST Solution by almost a full order of magnitude! (Under 3 hours versus 28 hours)

[TimeLogic](#) is a leading company in high speed bioinformatics computing. One of its products is the DeCypher BLAST Solution, which is an innovative combination of hardware engineering and software design to speed up BLAST searches. The DeCypher BLAST Solution was shown to map 1.7 million 25-mer nucleic sequences of proprietary SNP data to the human genome in **28 hours** on a Sun Ultra Sparc III server, which is **24X** faster than a regular BLASTN search on a Linux cluster of 30 Pentium III 1 GHz CPUs (see [SNP Analysis graph midway down the DeCypher BLAST Solution page](#)).

For comparison, we generated [1.7 million simulated SNP data](#) (file size about 85M) based on the latest human genome assembly (Release 19.34b, downloaded from Ensembl at [ftp://ftp.ensembl.org/pub/human-19.34b/data/golden\\_path](ftp://ftp.ensembl.org/pub/human-19.34b/data/golden_path)). It took less than **three hours** for Vmatch to complete the entire mapping of these SNPs onto the human genome on a Pentium IV 1.7 GHz CPU. The steps involved in this study are shown below. The first step creates the indices (mkvtree application) necessary for the subsequent mapping step (application of vmatch as specified in the search\_genome.csh script). VmatchNL will provide a GUI-driven alternative to this scripting, with management of indices and output modeled after GeneSeqerNL.