HGBASE
A Comprehensive Database of Human Intra-Genic DNA-Polymorphisms

Jana G. Böhm[1] , Marianne Siegfried[2], Jonathan Prince[2], Chandra M. Sarkar[1], Flavio R. Ortigao[1], and Anthony J. Brookes[2].
[1]Interactiva Biotechnologie GmbH, D-89077 Ulm, Germany
[2]Department of Genetics and Pathology, Biomedical Center, Uppsala, Sweden.
email: boehm@interactiva.de, marianne.siegfried@medgen.uu.se, jonathan.price@genpat.uu.se, sarkar@interactiva.de, ortigao@interactiva.de, tony.brookes@medgen.uu.se

SNPs are of immense interest to the Human Genome Project. However, until recently no comprehensive public database on SNPs was available. This changed in August 1998 when we released the HGBASE (Human Genic Bi-Allelic Sequences) database, a compilation of human gene related polymorphisms (primarily SNPs). The URL for HGBASE is http://hgbase.interactiva.de . Release 3.1, available since March 1999, contains ~3,300 human intra-genic polymorphisms, and the rate of growth will likely be exponential for some years with an expected doubling every six months.

HGBASE provides a comprehensive database of human gene polymorphisms located in the span from promoter to end of transcription (including exons and introns). We continually add polymorphism data from the literature, other databases and our own laboratory work. We also welcome contributions from the scientific community, and this is facilitated by our Web page submission forms or by submitting email or spreadsheet files (e.g., Excel, Access). Our data is made available to all other public databases that wish to exchange information with us and we strongly encourage this.

HGBASE data can be retrieved via a web interface with three access modes: i) "simple search" - searches a query word against all data except DNA sequences, ii) the Sequence Retrieval System (SRS) developed by Thure Etzold (EBI) - a common general search interface to many life science databases, and iii) "sequence search" - which uses NCBI's BLAST algorithms to query HGBASE DNA sequences. HGBASE is freely accessible to academia by the Web. A subscription based CD-ROM edition is also available. Polymorphism specific PCR conditions and assay suggestions are being assembled for imminent inclusion in HGBASE. Also, "added value" data such as extended links to other databases, information on gene expression, gene location data, and disease associations, will soon be provided free of charge to the scientific community.