

## Internet resources for the human geneticist\*

Sanjoy Kumar Pal<sup>a</sup>, Mahendra K Singh<sup>a</sup>, G S Pandey<sup>b</sup> & Balraj Mittal<sup>b,\*\*</sup>

<sup>a</sup>Department of Gastroenterology, <sup>b</sup>Department of Medical Genetics

Sanjay Gandhi Post Graduate Institute of Medical Sciences, Raebareilly Road, Lucknow 226 014, India

The Internet is a massive expanding body of information, which is likely to play a significant role for clinicians and researchers across the world. Since its inception in December 1969 the Internet has grown rapidly and is anticipated to expand 1000% in the coming next few years. Various useful databases on human genetics are already in 'the Net' and many more are being added constantly. The future of human geneticist is in handling of information. In this review of Internet and compilation of important web site addresses we expect to stimulate and instruct human geneticists in navigating the Net. The list of web sites provided in this article is expected to facilitate their search.

The Internet, more popularly known as the "information superhighway", has become a very pervasive influence in everyday life. Its power is most strongly seen among scientists, as more and more information is made available through "the Net," whether gene sequences, experimental data, chromosome maps, or whole journal articles<sup>1</sup>. With an estimated 320 million web pages available<sup>2</sup>, the Internet is a massive expanding body of information. The last few years have seen an exponential growth of the number of people with access to the Internet. Despite its growing commercial and entertaining contents, the Internet remains, as it was originally perceived, a powerful instrument for the exchange of scientific knowledge<sup>3</sup>.

The amount of scientific information and the number of electronic journals on the Internet continue to increase. The Internet and the World Wide Web (the Web) represent significant advancements for the retrieval and dissemination of scientific and other literature for the advancement of education. With the introduction of full-text search engines such as Alta Vista ([www.altavista.digital.com](http://www.altavista.digital.com)); Excite ([www.excite.com](http://www.excite.com)); Hot Bot ([www.hotbot.com](http://www.hotbot.com)); Infoseek ([www.infoseek.com](http://www.infoseek.com)); Lycos ([www.lycos.com](http://www.lycos.com)); and Northern Light ([www.nlight.com](http://www.nlight.com)), the web can be viewed as a searchable 15 billion word encyclopaedia. Immediate access to all scientific literature has long been a dream of scientists and the web search engine have made a large and growing body of scientific literature and other information resources

accessible within seconds. Scientific information retrieval and literature search, previously dominated by librarians, is now directly available to a widespread group of scientists<sup>4</sup>.

The Internet offers a number of resources to geneticists<sup>5,6</sup>. The MEROPS database<sup>7</sup> provides a catalogue and structure-based classification of peptidases. The review of Scriver and Nowacki<sup>8</sup> highlights the mutation databases on the Internet. Recently, Baxemanis<sup>9</sup> has reviewed various molecular biology databases. In this article, we aim to provide an overview of few databases and preliminary idea of navigating the Net.

### Human Genome Centers

Began in 1990 the human genome project is a 13-year effort coordinated by the US Department of Energy and the National Institute of Health. The project originally was planned to last 15 years, but rapid technology advancement has accelerated the expected completion date to 2003. The main aim of the project is to identify all, more than 100,000 genes in human DNA and determine the sequence of the 3 billion chemical base pairs that make up the human DNA. This task is being carried out in many laboratories throughout the world. A list of different Centers engaged in the human genome analysis is given in the Table 1.

### The Genome Database

The Genome Databases (GDB) is the official central repository for mapping data from the Human Genome Project. It is an encyclopaedia of human genomic information used primarily by gene mapping re-

\*This work was presented in the 25th Annual Conference of Human Genetics held in Nagpur on 9-10th January 2000.

\*\*Correspondent author: Email: [balraj@sgpgi.ac.in](mailto:balraj@sgpgi.ac.in)

searches to find markers for particular regions of the genome and, conversely, the locations of known markers and genes. Site addresses of some genomic databases are given in Table 2.

#### *Online Mendelian Inheritance in Man (OMIM)*

Mendelian Inheritance in Man (MIM) is an encyclopaedia of human genes and their associated phenotypes or genetic traits<sup>10</sup>. First published in 1966 by Victor A. McKusick, MIM is the most widely used compendium of information on genes and genetic disorders, and is of particular interest to investigators with clinical interests. MIM accession numbers are the standard by which genetic disorders are referenced. OMIM, is the continuously updated online version of MIM, maintained at the National Centre for Biotechnology Information (NCBI) and freely available on the Internet at <http://www.ncbi.nlm.nih.gov/omim><sup>11</sup>.

#### **BLAST**

Databases "homology" searching is carried out thousand of times each day by scientists worldwide; it is a critical tool for all molecular biologists. The basic operation is to take a query sequence (consisting of either nucleotides or amino acids), compare it se-

quentially with all known sequences in a database, and rank the database sequence in a "hit list" with respect to how closely they resemble the query. The most important factor is to make sure that the database being searched is comprehensive and up-to-date. With thousands of new sequences entering the public databases every week, it is essential to have access to daily updates and to repeat searches periodically, particularly for those queries that yield no significant match upon initial analysis.

BLAST, an acronym for "Basic Local Alignment Search Tool" represents a family of programs for databases similarly searching. BLAST programs embody a fast algorithm for sequence matching, rigorous statistical methods for judging the significance of the matches and various options for tailoring the program for special situations. A notable feature is the ability to 'mask' portions of the query sequence to reduce or eliminate spurious and misleading output. The most convenient way to access BLAST documentation is by using either e-mail or the www interface. For e-mail, send a message consisting of just the word 'help' to the normal BLAST server address ([blast@ncbi.nlm.nih.gov](mailto:blast@ncbi.nlm.nih.gov)). You will receive three e-mails in return: the help file for the BLAST e-mail

Table 1 — Site addresses of human genome centers

Baylor College of Medicine Genome Centre	<a href="http://gc.bmc.tmc.edu:8088/home.html">http://gc.bmc.tmc.edu:8088/home.html</a>
Columbia University Genome Centre	<a href="http://genome1.ccc.columbia.edu/~genome/">http://genome1.ccc.columbia.edu/~genome/</a>
Co-operative Human Linkage Centre	<a href="http://www.chcl.org">http://www.chcl.org</a>
Eleanor Roosevelt Institute	<a href="http://pg.nci.nih.gov/CHLC">http://pg.nci.nih.gov/CHLC</a>
Fondation Jean Dausset	<a href="http://www.eri.uchsc.edu/welcome.html">http://www.eri.uchsc.edu/welcome.html</a>
GENATLAS	<a href="http://www.cephb.fr">http://www.cephb.fr</a>
Genethon	<a href="http://bisance.cit2.fr/GENATLAS">http://bisance.cit2.fr/GENATLAS</a>
The Genetic Location Database	<a href="http://www.genethon.fr/genethon_en.html">http://www.genethon.fr/genethon_en.html</a>
Lawrence Berkeley Laboratory	<a href="http://cedar.genetics.soton.ac.uk/public.html">http://cedar.genetics.soton.ac.uk/public.html</a>
Human Genome Centre	<a href="http://www-hgc.ibl.gov/GenomeHome.html">http://www-hgc.ibl.gov/GenomeHome.html</a>
Lawrence Livermore National Laboratory	<a href="http://www-bio.llnl.gov/bbrp/genome/genome.html">http://www-bio.llnl.gov/bbrp/genome/genome.html</a>
Biology & Biotechnology Research program	<a href="http://www-lslam1.gov/Lswelcome.html">http://www-lslam1.gov/Lswelcome.html</a>
Los Alamos National Laboratory	<a href="http://www.marshmed.org/genetics">http://www.marshmed.org/genetics</a>
Marshfield Medical Research Foundation	<a href="http://rmc-www.ibl.gov">http://rmc-www.ibl.gov</a>
Resource for Molecular Cytogenetics	<a href="http://www.sanger.ac.uk">http://www.sanger.ac.uk</a>
Sanger Centre	<a href="http://genome-www.stanford.edu/SDSATC">http://genome-www.stanford.edu/SDSATC</a>
Stanford DNA sequence & technology Centre	<a href="http://shgc-www.stanford.edu">http://shgc-www.stanford.edu</a>
Stanford Human genetic Centre	<a href="http://www.tigr.org">http://www.tigr.org</a>
The Institute of Genomic Research	<a href="http://seqcore.brcf.med.umich.edu">http://seqcore.brcf.med.umich.edu</a>
University of Michigan Human Genome Centre	<a href="http://www.genome.washington.edu/uwgc">http://www.genome.washington.edu/uwgc</a>
University of Washington Genome Centre	<a href="http://www.ibt.wustl.edu/egm/egm.html">http://www.ibt.wustl.edu/egm/egm.html</a>
Washington University Centre for Genetics in Medicine	<a href="http://genome.wustl.edu/gsc/gschmpg.html">http://genome.wustl.edu/gsc/gschmpg.html</a>
Washington Univ. Genome Sequence Centre	<a href="http://www.well.ox.ac.uk">http://www.well.ox.ac.uk</a>
Wellcome Trust Centre for Human Genetics	<a href="http://quercus.medgen.uu.se">http://quercus.medgen.uu.se</a>
Uppsala Genome Sequencing Laboratory	<a href="http://www.jgi.doe.gov">http://www.jgi.doe.gov</a>
Centre for Human Genome Studies	

server (e-mail help), the manual for the BLAST programs (manual) and the BLAST Frequently Asked questions file (FAQ). In addition to the above documentation files, the www BLAST interface includes on-line information about the BLAST theory and algorithm, bibliography, FAQ, usage statistics and a "what's new" features. User support for any difficulties or problems encountered can be obtained by sending the questions to [blasthelp@ncbi.nlm.nih.gov](mailto:blasthelp@ncbi.nlm.nih.gov)<sup>12</sup>.

#### *Radiation Hybrid Mapping*

The use of radiation hybrid panels is well established as a tool for construction by high-resolution

maps of genomic regions. It is now possible, to purchase DNA that has been isolated from panels of whole-genome radiation hybrid clones. It is also now possible to have the data obtained from the assay of such commercially available DNA analysed by sending it over the Internet to a server by e-mail or to a world wide web site using Netscape. The European Consortium on radiation hybrid (RH) mapping have constructed a radiation hybrid database with ~10,000 RH entries and the site address is <http://www.ebiac.uk/RHdb/><sup>13</sup>.

#### *Mapping Single Marker of Unknown Location*

The availability of public database with the typing

Table 2—Site addresses of some genome databases

#### *Whole Genome Data*

MIT Genome Centre/Whitehead Institute

<http://www-genome.wi.mit.edu>

[www-genome.wi.mit.edu/cgi-bin/contig/phys\\_map](http://www-genome.wi.mit.edu/cgi-bin/contig/phys_map)

email: [webmaster@genome.wi.mit.edu](mailto:webmaster@genome.wi.mit.edu)

Stanford Human Genome Centre

<http://shgc-www.stanford.edu>

email: [rmyers@camis.stanford.edu](mailto:rmyers@camis.stanford.edu)

CEPH - Genethon (YAC - based fingerprinting & STS maps)

<http://cartagene.cephb.fr/bio/ceph-genethon-map.html>

email: [ceph-genethon-map@cephb.fr](mailto:ceph-genethon-map@cephb.fr)

[ceph-genethon-quickmap@cephb.fr](mailto:ceph-genethon-quickmap@cephb.fr)

email (Help): [denis@ceph.cephb.fr](mailto:denis@ceph.cephb.fr)

#### *Other Related Databases*

GenBank

Genome Sequence Database (GSDB)

EMBL Nucleotide Sequence Database

Nucleic Acid Database (NDB)

Protein Data Bank (PDB)

Protein Information Resources

MEROPS: (peptide database)

Human Gene Mutation Databases (HMGD)

Mutation database

RBI Gene Mutation Database

Human PAX2 Allelic Variant Database

Single nucleotide polymorphism database

DNA Data Bank of Japan

BEAUTY Annotated Protein

ExPASy (Including the SWISS-

PROT databases)

MitoDat (Mitochondrial proteins)

MITOMAP (Human mitochondrial genome)

Allelic frequencies and DNA polymorphism

(ALFRED)

Haemophilia B Mutation Database

5S Ribosomal RNA Databank

Guide RNA Database

Small RNA Database

Molecular Probe Databases

<http://www.ncbi.nlm.nih.gov/Web/Genbank/>

<http://www.ncgr.org/gsdh>

<http://www.ebi.ac.uk/embl.html>

<http://ndbserver.rutgers.edu:80>

<http://www.pdb.bnl.gov>

<http://www-nbrf.georgetown.edu/pir/index.html>

<http://www.bi.bbcr.ac.uk/Merops/Merops.htm>

<http://uwcw.web.cf.ac.uk/uwcw/mg/hgmd0.html>

<http://www2.ebi.ac.uk/mutations/>

<http://ariel.ucs.unimelb.edu.au:80/~cotton/dblist.htm>

<http://home.kamp.net/home/dr.lohmann/>

<http://www.hgu.mrc.ac.uk/Softdata/PAX2/>

<http://www.ncbi.nlm.nih.gov/SNP>

<http://www.ddbj.nig.ac.jp>

<http://dot.imgen.bcm.tmc.edu:9331/seq-search/help/beauty.html>

<http://expasy.heuge.ct>

<http://www-lecb.ncicrf.gov/mitoDat/>

<http://www.gen.emory.edu/mitomap.html>

<http://fondue.med.yale.edu/db2/>

<http://www.umds.ac.uk/molgen/haemBdatabase.htm>

<http://www.rose.man.poznam.pl/5SSdata/5SRNA.html>

<http://www.biochem.mpg.de/~goeringe/>

<http://mbr.bcm.tmc.edu/smallRNA/smallrna.html>

<http://srs.ebi.ac.uk/>

of 5994 markers (G3 panel) and 14,308 markers (G4 panel) greatly facilitates the process of finding the genomic location of an anonymous marker. In contrast to other mapping methods e.g. fluorescence *in situ* hybridization radiation hybrid mapping with these databases allows any small fragments to be localized in a day with great ease. It is worthwhile to check the databases prior to assaying and scoring a marker because the names of markers that have already been typed are available on the databases together with the raw scoring data for those markers. Marker scoring data that has been obtained either from the databases or from PCR assay performed on the Stanford e-mail server or to the Whitehead Institute/MIT Centre for Genome Research via Netscape. Where upon two-point linkage tests will be performed automatically, in case of the Stanford e-mail server, the identity of any marker that is linked to the submitted marker with a LOD score  $\geq 6$  is returned to the sender within minutes of submission. A random marker has a 75% chance of being linked to a mapped marker with a LOD  $\geq 6$ . If the unmapped marker has already been assigned to a chromosome by some other method, it has a 90% chance of being linked to a mapped marker with a LOD  $\geq 3$ . In addition, it is possible to do one's own two-point analysis to identify linked

markers that are in the databases. The uniform resource locator (URL) for the database is [ftp://shgc.stanford.edu/pub/hgmc/RH\\_data/rh\\_scores](http://shgc.stanford.edu/pub/hgmc/RH_data/rh_scores).

Data derived from screening G4 panel can be submitted to the Whitehead/MIT Centre for Genome Research using their www site (the URL is <http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>). A two-point analysis is performed, and framework marker linked with a LOD score of 15 assign the submitted marker to a chromosome. A multipoint analysis with the framework marker for that chromosome is then carried out to give the map position of the submitted marker. Results are returned to the sender by e-mail within 24 hr of data submission. Because of the number of framework marker scored and the panel resolution, an unknown marker can always be localized. The software used at this center is available for in-house analysis if required (the URL is <http://www-genome.wi.mit.edu/ftp/pub/software/rhmapper/>). In addition, it is possible to download the raw data and run the analysis program.

Radiation hybrid scoring data can be sent directly to the rhserver via e-mail (to [rhserver@shgc.stanford.edu](mailto:rhserver@shgc.stanford.edu)). Instruction on how to send data can be obtained by sending an empty e-mail message with the word "info" as the subject line<sup>13</sup>.

Table 3—Chromosome specific physical maps

Chromosome	Laboratory	Map Type	URL
3	University of Texas Health Science Centre Eleanor Roosevelt Institute Stanford Human Genome Centre Sanger Centre The Hospital for sick children, Toronto University of Texas Health Science Centre	STS/YAC+RH STS/YAC STS/YAC RH YAC RH	<a href="http://mars.uthscsa.edu/">http://mars.uthscsa.edu/</a> <a href="http://www-eri.uchsc.edu/">http://www-eri.uchsc.edu/</a> <a href="http://shgc.stanford.edu/">http://shgc.stanford.edu/</a> <a href="http://www.sanger.ac.uk/chr6/chr6.html">http://www.sanger.ac.uk/chr6/chr6.html</a> <a href="http://www.genet.sickkids.on.ca/">http://www.genet.sickkids.on.ca/</a> <a href="http://www-eri.uchsc.edu/">http://www-eri.uchsc.edu/</a>
10	Genome Therapeutics Corporation	STS/YAC	<a href="http://www.cric.com/">http://www.cric.com/</a>
11	University of Texas South Medical Centre Yale/Albert Einstein Columbia University Los Alamos National Labs Lawrence Livermore National Labs Sanger Centre University of Pennsylvania	STS/YAC, RH Cosmid STS/YAC STS/YAC RH+STS/cosmid Cosmid YAC,BAC,PAC STS/YAC	<a href="http://mcdermott.swmed.edu/datapage/">http://mcdermott.swmed.edu/datapage/</a> <a href="http://paella.med.yale.edu/chr12/home.html">http://paella.med.yale.edu/chr12/home.html</a> <a href="http://genome1.ccc.columbia.edu/~genome/">http://genome1.ccc.columbia.edu/~genome/</a> <a href="http://www-ls.lanl.gov/masterhg.html">http://www-ls.lanl.gov/masterhg.html</a> <a href="http://wwwbio.llnl.gov/fbrmp/genome/genome.html">http://wwwbio.llnl.gov/fbrmp/genome/genome.html</a> <a href="http://www.sanger.ac.uk/hum22">http://www.sanger.ac.uk/hum22</a> <a href="http://www.ebil.upenn.edu/HGC22.html">http://www.ebil.upenn.edu/HGC22.html</a>
X	Baylor College of Medicine	STS/YAC	<a href="http://gc.bcm.tmc.edu:8088/chrx/home.html">http://gc.bcm.tmc.edu:8088/chrx/home.html</a>
Mito	Emory University	Sequences	<a href="http://infinitiy.gen.emory.edu/mitomap.html">http://infinitiy.gen.emory.edu/mitomap.html</a>

Mito - Mitochondria genome; PAC - PI artificial chromosome; RH - Radiation hybrid; STS - Sequence tagged sites; YAC - Yeast artificial chromosome

### Navigating Public Physical Mapping

Only a few years ago, just a handful of ready-made maps of the human genome existed, and these were low-resolution maps of small areas. The situation has changed dramatically in recent years. Ironically, the researcher's burden has now shifted from mapping the genome to navigating a vast *terra incognita* of world wide web sites, FTP servers, and databases. There are large databases such as NCBI, Entrez and GDB, smaller databases serving the primary maps published by genome centers, sites sponsored by individual chromosome committees, and site used by smaller laboratories to publish highly detailed maps of specific regions. Each type of resource contains information that overlaps with the information found in other, but valuable in its own right. Finding one's own way around this web of information is not easy. A recent search for the word 'genome' in the AltaVista web search engine turned up to over 80,000 potentially relevant documents.

Among the most accessible of the resources for physical mapping information is the genomes division of Entrez a service provided by the National Centre for Biotechnology Information. Entrez attempt to integrate, in a browsable manner, several genetic and physical maps with DNA and protein-sequencing information, a bibliographic citation database and three-dimensional crystallographic structure information. Although Entrez does not incorporate the most recent physical mapping information (it is updated quarterly intervals), it is recommended as a start point for map searches because of its rich interconnectedness and simple interface. The URL of this site is <http://www.ncbi.nlm.nih.gov/Entrez/>.

Another general source of physical mapping data is the Genome Data Base. Although its original organization was geared towards the preponderance of genetic maps available at the time of its creation, GDB has undergone extensive revision in recent years and now a repository of physical mapping data as well. Like NCBI, GDB is available on the world wide web at <http://gdbwww.gdb.org/>.

Two genome-wide genetic maps are also available in the web. The Genethon map<sup>14</sup> consist of 5264 polymorphic microsatellite repeats that have been placed in 1.6-cM intervals. The complete data files, along with graphical representations of the map in PostScript Formate, are available at <ftp://ftp.genethon.fr/pub/Gmap/Nature-1995/>. These maps are also available in browsable form through GDB.

The second large genetic map is produced by the Cooperative Human Linkage Centre. It consists of 10,775 markers, mostly microsatellite repeats, placed in 3.7-cM intervals. The primary data site for this map is <http://www.chlc.org/>.

In October 1996, a whole-genome transcript map of Homo sapiens was published in Science. This map consists of ~15,000 unique expressed sequences (ESTs) localized by radiation hybrid mapping relative to a framework derived from the Genethon genetic map. The transcript map is available to the public at two web sites. The 'parent' site for the data is at NCBI. Here the full text of the Science article containing the whole-genome transcript map of Homo sapiens will be found, along with image of the colorful, but essentially decorative, wall chart that accom-

Table 4—Site addresses of individual chromosome data

Chromosome	URL
1	<a href="http://linkage.rockefeller.edu/chr1/index.html">http://linkage.rockefeller.edu/chr1/index.html</a> <a href="http://www.sanger.ac.uk/HGP/chr1">http://www.sanger.ac.uk/HGP/chr1</a>
3	<a href="http://mars.uthscsa.edu">http://mars.uthscsa.edu</a> email: <a href="mailto:naylor@uthscsa.edu">naylor@uthscsa.edu</a> <a href="http://www-eri.uchsc.edu/chrom3.html">http://www-eri.uchsc.edu/chrom3.html</a> email: <a href="mailto:gemmill@eri.uchsc.edu">gemmill@eri.uchsc.edu</a>
4	<a href="http://apollo.uthscsa.edu/DB">http://apollo.uthscsa.edu/DB</a> <a href="http://shgc-www.stanford.edu">http://shgc-www.stanford.edu</a> email: <a href="mailto:rmyers@camis.stanford.edu">rmyers@camis.stanford.edu</a>
5	<a href="http://chrom5.hsis.uci.edu/">http://chrom5.hsis.uci.edu/</a>
6	<a href="http://www.sanger.ac.uk/HGP/chr6">http://www.sanger.ac.uk/HGP/chr6</a>
7	<a href="http://www.genet.stickkids.on.ca">http://www.genet.stickkids.on.ca</a> <a href="http://www.nhgri.nih.gov/DIR/GTB/CHR7/">http://www.nhgri.nih.gov/DIR/GTB/CHR7/</a> email: <a href="mailto:egreen@nchgr.nih.gov">egreen@nchgr.nih.gov</a>
8	<a href="http://apollo.uthscsa.edu/">http://apollo.uthscsa.edu/</a>
9	<a href="http://www.gene.ucl.ac.uk/chr9home.html">http://www.gene.ucl.ac.uk/chr9home.html</a>
10	<a href="http://www.genomecorp.com/genesequence/index.html">http://www.genomecorp.com/genesequence/index.html</a> e mail: <a href="mailto:randall.little@genomecorp.com">randall.little@genomecorp.com</a>
11	<a href="http://mcdermott.swmed.edu/human.htm">http://mcdermott.swmed.edu/human.htm</a> email: <a href="mailto:gevans@swmed.edu">gevans@swmed.edu</a>
12	<a href="http://paella.med.yale.edu/home.html">http://paella.med.yale.edu/home.html</a> email: <a href="mailto:kucherla@aecom.yu.edu">kucherla@aecom.yu.edu</a>
13	<a href="http://genomeL.ccc.columbia.edu/~genome/">http://genomeL.ccc.columbia.edu/~genome/</a> email: <a href="mailto:arg@cuca.ccc.columbia.edu">arg@cuca.ccc.columbia.edu</a>
15	<a href="http://mendel.medgen.ubc.ca/medgen/scw15.html">http://mendel.medgen.ubc.ca/medgen/scw15.html</a>
16	<a href="http://www.lanl.gov/masterhgp.html">http://www.lanl.gov/masterhgp.html</a>
17	<a href="http://bioinformatics.weizmann.ac.il/db17/">http://bioinformatics.weizmann.ac.il/db17/</a>
18	<a href="http://www.childrenhospital.org/chromosome18">http://www.childrenhospital.org/chromosome18</a>
19	<a href="http://www-bio.lnl.gov/bbrp/genome.html">http://www-bio.lnl.gov/bbrp/genome.html</a> email: <a href="mailto:ashworth@lnl.gov">ashworth@lnl.gov</a>
20	<a href="http://www.sanger.ac.uk/HGP/chr20">http://www.sanger.ac.uk/HGP/chr20</a>
21	<a href="http://www-eri.uchsc.edu/chr12/c21index.html">http://www-eri.uchsc.edu/chr12/c21index.html</a>
22	<a href="http://www.cbil.upenn.edu/HGC22/index.html">http://www.cbil.upenn.edu/HGC22/index.html</a> <a href="http://www.sanger.ac.uk/hum22">http://www.sanger.ac.uk/hum22</a>
X	<a href="http://gc.bcm.tmc.edu:8088/chrX/home.html">http://gc.bcm.tmc.edu:8088/chrX/home.html</a> email: <a href="mailto:nelson@bcm.tmc.edu">nelson@bcm.tmc.edu</a>

panied it. A limitation of the NCBI web site is that it does not provide a graphical view of the transcript map beyond resolution histograms of marker distributions. The URL of this site is <http://www.ncbi.nlm.nih.gov/Science96/>.

#### *Human Physical Maps*

The Whitehead Institute/MIT Centre for Genome Research is the primary source of two genome wide physical maps. The maps maintained here are an STS content map of more than 10,000 markers assigned to YACs and a radiation hybrid map of some 12,000

markers. These maps have been combined with the Genethon genetic map to create an integrated map of 20,000 STS's at an average spacing of 150 kb. Approximately half the markers on the Whitehead maps are expressed sequences that appear on the human transcript map as well. The Whitehead Institute maps are available via the web at <http://www.genome.wi.mit.edu/>. Follow the link to Human Physical Mapping Project and from there to the map of interest. The map is browsable in a number of ways<sup>15</sup>. A list of chromosome specific physical maps is given in Table 3.

Table 5—Recourses list of few interesting web sites

GenCard	<a href="http://bioinfo.weizmann.ac.il/card/">http://bioinfo.weizmann.ac.il/card/</a>
Gene Clinic	<a href="http://www.geneclinic.org/">http://www.geneclinic.org/</a>
Gene and Disease	<a href="http://www.ncbi.nlm.nih.gov/disease/">http://www.ncbi.nlm.nih.gov/disease/</a>
LocusLink	<a href="http://www.ncbi.nlm.nih.gov/LocusLink/">http://www.ncbi.nlm.nih.gov/LocusLink/</a>
National organisation for rare disorders	<a href="http://www.rarediseases.org">http://www.rarediseases.org</a>
Human Genetic Organisation	<a href="http://www.faseb.org/genetics/ashg/geneintr.htm">http://www.faseb.org/genetics/ashg/geneintr.htm</a>
Genetics - Women Health Net Links	<a href="http://womenhealth.about.com/msub112.htm">http://womenhealth.about.com/msub112.htm</a>
Cloning: Non-Human - Genetics Net Links	<a href="http://genetics.about.com/msub8.htm">http://genetics.about.com/msub8.htm</a>
Journals - Genetics Net Link	<a href="http://genetics.about.com/msub3.htm">http://genetics.about.com/msub3.htm</a>
Genetics - Pediatrics - Net Link	<a href="http://genetics.about.com/msub23.htm">http://genetics.about.com/msub23.htm</a>
Genetics - Home Page	<a href="http://genetics.about.com/mbody.htm">http://genetics.about.com/mbody.htm</a>
Clinical Genetic Site	<a href="http://www.kumc.edu">http://www.kumc.edu</a>
Rare Genetic Disease in Children: (An Internet Resource)	<a href="http://maerctr4.med.nyu.edu/~murrp01/homenew.htm">http://maerctr4.med.nyu.edu/~murrp01/homenew.htm</a>
Dysmorphology Discussion Board	<a href="http://genetics.ich.bprf.ac.uk/DDB/ddb.html">http://genetics.ich.bprf.ac.uk/DDB/ddb.html</a>
Human/Mouse Homology Relationship	<a href="http://www.information.jax.org">http://www.information.jax.org</a>
PubMed (MEDLINE)	<a href="http://www.ncbi.nlm.nih.gov/PubMed/">http://www.ncbi.nlm.nih.gov/PubMed/</a>
FlyBase	<a href="http://www.fruitfly.org">http://www.fruitfly.org</a>
Mendel Database	<a href="http://jjiio6.jic.bbsrc.ac.uk/">http://jjiio6.jic.bbsrc.ac.uk/</a>
GenAtlas	<a href="http://www.citi2.fr/GENATLAS/">http://www.citi2.fr/GENATLAS/</a>

Table 6—Web addresses of few online journals

American Journal of Human Genetics	<a href="http://www.faseb.org/genetics/ashg/jou-sahg.htm">http://www.faseb.org/genetics/ashg/jou-sahg.htm</a>
American Journal of Human Genetics (Electronic Edition)	<a href="http://www.ajhg.org">http://www.ajhg.org</a>
American Journal of Medical Genetics	<a href="http://wiley.edoc.com/">http://wiley.edoc.com/</a>
Annals of Human Genetics	<a href="http://www.cup.org/journals/JNLSCAT/hge/hge.htm">http://www.cup.org/journals/JNLSCAT/hge/hge.htm</a>
British Society of Human Genetics	<a href="http://www.bham.ac.uk/bshg/oldhome.htm">http://www.bham.ac.uk/bshg/oldhome.htm</a>
Cell	<a href="http://www.cell.com/">http://www.cell.com/</a>
Cytogenetics and Cell Genetics	<a href="http://http://karger.ch/journals/ccg/ccgdes.htm">http://http://karger.ch/journals/ccg/ccgdes.htm</a>
Clinical Dysmorphology	<a href="http://www.chapmanhall.com/cd.html">http://www.chapmanhall.com/cd.html</a>
European Journal of Human Genetics	<a href="http://www.stockton-press.co.uk">http://www.stockton-press.co.uk</a>
Gene	<a href="http://www.elservier.com/locate/gene">http://www.elservier.com/locate/gene</a>
Genome	<a href="http://www.idealibrary.com">http://www.idealibrary.com</a>
Human Genetics	<a href="http://link.springer.delink.springer-ny.com">http://link.springer.delink.springer-ny.com</a>
Human Heredity	<a href="http://www.hum-molgen.de/journals/HH/index.html">http://www.hum-molgen.de/journals/HH/index.html</a>
Human Molecular Genetics	<a href="http://karger.ch/journals/hhe/hhedes.htm">http://karger.ch/journals/hhe/hhedes.htm</a>
Journal of Human Genetics	<a href="http://www3.oup.co.uk/html">http://www3.oup.co.uk/html</a>
Nature Genetics	<a href="http://www.link.springer-ny.com">http://www.link.springer-ny.com</a>
The Lancet	<a href="http://genetics.nature.com">http://genetics.nature.com</a>
Nature	<a href="http://www.thelancet.com">http://www.thelancet.com</a>
Science	<a href="http://www.nature.com">http://www.nature.com</a>
Indian Journal of Experimental Biology	<a href="http://www.sciencemag.org">http://www.sciencemag.org</a>
	<a href="http://www.bioline.bdt.org.br/ic">http://www.bioline.bdt.org.br/ic</a>

### *Chromosome Specific Maps*

In addition to the genome-wide maps, research laboratories and genome centers have built physical maps of many individual chromosomes. In many cases, these maps provide greater details than the corresponding whole genome map does. Table 4 lists the maps and where to find them on the web. In addition to the URL shown, an up-to-date list can always be found at GDB's web page at [http://gdbwww.gdb\\_org/gdb/hgp-resources.html](http://gdbwww.gdb_org/gdb/hgp-resources.html).

### *Genes and diseases*

This section presents a series of web pages that are dedicated to specific diseases. They provide the lay person and the medical professionals with short synopses or essays of what is known about a specific genetic disease. The site provides the list of about 60 genetic disorders. It is anticipated that many more will be added in the future. They represent an excellent resource for patients and their families as they explain intelligently, avoiding specialist jargon, a number of genetic diseases. Special emphasis is placed on disorders where the mutation involved is known<sup>6</sup>. The web address of this site is given in Table 5.

### *LocusLink*

LocusLink presents a central platform for Reference mRNA sequences (RefGenes). These are curated human records taken from GenBank that are 'owned' by the NCBI, giving them the liberty to edit, as best they see fit, these records and present the user with a simplified, correct (albeit, currently incomplete) view of the human genome. Currently, NCBI presents curated mRNAs and other proteins. This also allows you to link-out to OMIM, GenBank, Unigene and dbSNP, which presents variations in the databases for single nucleotide polymorphisms. Unigene sequences are the result of clustering (assembling) the ESTs from human, mouse and rat. They provide an excellent resource for gene discovery and identification<sup>6</sup>.

### *Online Journals*

Until recently many feared that increasing use of the Internet would widen yet further knowledge gap between industrialized and developing countries. But such pessimism is giving way to optimism that the new technology may instead help spur a renaissance of science and technology in poor countries. Use of the Internet in developing countries is growing faster than anywhere else. A plethora of initiative is underway to boost Internet networks, for universities

in developing countries. Widespread access to the Internet allows scientists to use a vast range of tools, literature and databases. It promises to offset the drop in journal acquisition by libraries in developing countries that has occurred over the past decade because of inflation<sup>16</sup>. Online versions of many journals are now available through the Internet. A list of few journals related to human genetics is given in the Table 6.

### **Conclusion**

Internet is a rich source of valuable information that can be used to great advantage, for keeping up with the recent developments in the fast-moving disciplines of human genetics. The human genome project is almost completed, soon we will have the entire information on human genome in the Internet. The future of geneticists is in handling of information and therefore we cannot afford to ignore the Internet any more. Computer literacy and an Internet proficiency must be a part of postgraduate medical education, as Internet will be playing an important role in human health care<sup>17</sup>.

### **Acknowledgement**

PDF to SKP from 'Lady Tata Memorial Trust' and Institutional fellowship to MKS and GSP is duly acknowledged.

### **Notes added in proofs**

The sequence analysis of the 'human genome' is now available online and the URL of these sites are <http://www.nature.com/genomics> and <http://www.sciencemag.org/content/vol291/issue5507/>

### **References**

- 1 Baxevasis A D & Ouellette B F F. Internet basics for biologist. In : Dracopoli N C, Haines J L, Korf B R, Moir D T, Morton C C, Seidman C E, Seidman J G, Smith D R, eds. *Current protocols in human genetics*. (John Wiley & Sons, Inc., New York), 1998, 11.1.1.
- 2 Biermann J S, Golladay G J, Greenfield M L V H & Baker L H. Evaluation of cancer information on the Internet. *Cancer*, 86 (1999) 381.
- 3 Ceelen W P & Waele J D. Surgeons and the Internet. *Eur J Surg*, 165 (1999) 83.
- 4 Lawrence S & Giles C L. Searching the world wide web. *Science*, 280 (1998) 98.
- 5 van Steensel M A & Winter R M. Internet databases for clinical geneticists: an overview. *Clin Genet*, 53 (1998) 323.
- 6 Ouellette F. Internet resources for the clinical geneticist. *Clin Genet*, 56 (1999) 179.
- 7 Rawling N D & Barrett A J. MEROPS: the peptidase database. *Nucleic Acids Res*, 27 (1999) 325.

- 8 Sriver C R & Nowacki P M. Genomics, mutation and the Internet: the naming and the use of parts. *J Inherit Metab Dis*, 22 (1999) 519.
- 9 Baxevanis A D. The molecular biology databases collection: An online compilation of relevant databases resources. *Nucleic Acids Res*, 28 (2000) 1.
- 10 McKusick V A. *Mendelian inheritance in man*. (The Johns Hopkins University Press, Baltimore, Md.), 1998.
- 11 Cuticchia A J. Human mapping databases. In : Dracopoli N C, Haines J L, Korf B R, Moir D T, Morton C C, Seidman C E, Seidman J G, Smith D R (eds). *Current protocols in human genetics*. (John Wiley & Sons, Inc., New York), 1998, 11.4.
- 12 Schuler G D, Boguski M S & Recipon H E. Sequence similarity searching using the BLAST family of programs. In : Dracopoli N C, Haines J L, Korf B R, Moir D T, Morton C C, Seidman C E, Seidman J G, Smith D R (eds). *Current protocols in human genetics*. (John Wiley & Sons, Inc., New York), 1998, 11.3.1.
- 13 Carey A H . Use of commercially available radiation hybrid panels. In : Dracopoli N C, Haines J L, Korf B R, Moir D T, Morton C C, Seidman C E, Seidman J G, Smith D R (eds). *Current protocols in human genetics*. (John Wiley & Sons, Inc., New York), 1998, 3.5.1.
- 14 Dip C, Faure C, Fizames D, Samson N *et al*. A comprehensive genetic map of the human genome base on 3264 micro-satellites. *Nature*, 380 (1996) 152.
- 15 Stein L D. Navigating public physical mapping databases. In: Dracopoli N C, Haines J L, Korf B R, Moir D T, Morton C C, Seidman C E, Seidman J G, Smith D R, eds. *Current protocols in human genetics*. (John Wiley & Sons, Inc., New York), 1998, 5.16.1.
- 16 Butler D. Internet may help bridge the gap. *Nature*, 379 (1999) 10.
- 17 Singh M K, Pal S K, Kapoor V K & Choudhuri G. Gastroenterology and the Internet. *Tropical Gastroenterol*, 21 (2000) 165.