CONF-8605274--Abstr.

Abstracts of papers presented at the
51st Cold Spring Harbor Symposium
on Quantitative Biology

MOLECULAR BIOLOGY OF HOMO SAPIENS

May 28–June 4, 1986

Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Abstracts of papers presented at the
51st Cold Spring Harbor Symposium
on Quantitative Biology

MOLECULAR BIOLOGY OF HOMO SAPIENS

May 28–June 4, 1986

CONF-8605274--Absts.

DE87 011013

Arranged by
James D. Watson, *Cold Spring Harbor Laboratory*
Marcello Siniscalco, *Memorial Sloan-Kettering Cancer Center*

Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Front cover: Jones Laboratory, constructed in 1893, renovated in 1972.

Back cover: Carnegie Library, originally designed as a laboratory in 1905, refitted as a library in 1953.

Contributions from the following Corporate Sponsors provide core support for the Cold Spring Harbor meetings program:

Abbott Laboratories
American Cyanamid Company
Amersham International plc
Becton Dickinson and Company
Cetus Corporation
Ciba-Geigy Corporation
CPC International Inc.
E.I. du Pont de Nemours & Company
Eli Lilly and Company
Genentech, Inc.
Genetics Institute
Hoffmann-La Roche Inc.
Monsanto Company
Pall Corporation
Pfizer Inc.
Schering-Plough Corporation
Smith Kline & French Laboratories
The Upjohn Company
Wyeth Laboratories

This meeting was supported in part by the **Banbury Fund**; the **Department of Energy; Lucille P. Markey Foundation; National Cancer Institute/National Institutes of Health; The National Science Foundation**; and the **Rita Allen Foundation.**

PROGRAM

WEDNESDAY, May 28—7:30 PM

Welcoming Remarks: J.D. Watson

Introduction: W.F. Bodmer, Imperial Cancer Research Fund: Human genetics—The molecular challange.

SESSION 1 HUMAN GENE MAP (1)

Chairman: V. McKusick, Johns Hopkins Hospital

McKusick, V.A., Johns Hopkins University, Baltimore, Maryland: The gene map of *Homo sapiens*—Status and prospectus. 1

Gray, J.W., van den Engh, G., Trask, B., Lucas, J., Pinkel, D., Van Dilla, M., Fuscoe, J., Yu, L., Lawrence Livermore National Laboratory, Livermore, California: Chromosome classification and purification using flow cytometry and sorting. 2

Pinkel, D.,[1] Gray, J.W.,[1] Trask, B.,[1] van den Engh, G.,[1] Fuscoe, J.,[1] van Dekken, H.,[1,2] [1]Lawrence Livermore National Laboratory, Livermore, California; [2]Radiobiological Institute, TNO, The Netherlands: Cytogenetic analysis by in situ hybridization with fluorescently labeled nucleic acid probes. 3

Deaven, L.,[1] Van Dilla, M.,[2] [1]Los Alamos National Laboratory, New Mexico; [2]Lawrence Livermore National Laboratory, Livermore, California: Construction of human chromosome-specific DNA libraries from flow-sorted chromosomes. 4

THURSDAY, May 29—9:00 AM

SESSION 2 HUMAN GENE MAP (2)

Chairman: D. Botstein, Massachusetts Institute of Technology

White, R., Julier, C., Leppert, M., Nakamura, Y., O'Connell, P., Woodward, S., Lathrop, G.M., Lalouel, J.-M., Howard Hughes Medical Institute, Salt Lake City, Utah: Construction of human genetic linkage maps. I. Progress and perspectives. 5

Lalouel, J.-M., Lathrop, G.M., White, R., Howard Hughes Medical Institute, University of Utah Medical Center, Salt Lake City: Construction of human genetic maps. II. Methodological issues. 6

Lander, E.,[1,2,3] Botstein, D.,[2] [1]Whitehead Institute for Biomedical Research, [2]Dept. of Biology, Massachusetts Institute of Technology, [3]Harvard Univer-

THURSDAY, May 29—2:00 PM

SESSION 3 HUMAN CANCER GENES (1)

Chairman: G.F. Vande Woude, NCI-Frederick Cancer Research Facility

THURSDAY, May 29—4:30 PM

Wine and Cheese Party

THURSDAY, May 29—7:30 PM

SESSION 4 HUMAN GENE MAP (3)

Chairman: L. Hood, California Institute of Technology

FRIDAY, May 30—9:00 AM

SESSION 5 GENETIC DIAGNOSIS (1)

Chairman: P.L. Pearson, University of Leiden

FRIDAY, May 30—2:30 PM

SESSION 6 HUMAN EVOLUTION (1)

Chairman: L.L. Cavalli-Sforza, Stanford University

FRIDAY, May 30—1:30 PM

SESSION 7 DRUGS MADE OFF HUMAN GENES 1: CLOTTING ANTI-CLOTTING FACTORS

Chairman: E. Davie, University of Washington

SATURDAY, May 31 – 9:00 AM

SESSION 8 DRUGS MADE OFF HUMAN GENES 2: ANTI-CANCER AGENTS

Chairman: **R.A. Flavell**, Biogen Research Corp.

SATURDAY, May 31—3:00 PM

Picnic

SATURDAY, May 31—7:30 PM

SESSION 9 GENETIC DIAGNOSIS (2)

Chairman: T. Maniatis, Harvard University

SUNDAY, June 1—9:00 PM

SESSION 10 DRUGS MADE OFF HUMAN GENES 3: GROWTH FACTORS

Chairman: G.J. Todaro, Oncogen, Seattle

SUNDAY, June 1—2:00 PM

SESSION 11 HUMAN EVOLUTION (2)

Chairman: M. Singer, National Cancer Institute

SUNDAY, June 1—6:00 PM

DEDICATION OF GRACE AUDITORIUM

MONDAY, June 2—9:00 AM

SESSION 12 RECEPTORS (1)

Chairman: R. Axel, Columbia University College of Physicians & Surgeons

MONDAY, June 2—2:30 PM

SESSION 13 HUMAN GENE MAP (4)

Chairman: M. Siniscalco, Memorial Sloan-Kettering Cancer Center

MONDAY, June 2 - 7:30 PM

SESSION 14 HUMAN CANCER GENES (2)

Chairman: C.M. Croce, Wistar Institute

TUESDAY, June 3—9:00 AM

SESSION 15 GENE THERAPY (1)

Chairman: P. Berg, Stanford University

TUESDAY, June 3—2:30 PM

SESSION 16 RECEPTORS (2)

Chairman: M.J. Goldstein, University of Texas

TUESDAY, June 3—5:00 PM

SESSION 17 SEQUENCING OF THE HUMAN GENOME

Discussion led by **Walter Gilbert**, Harvard University and **Paul Berg**, Stanford University

TUESDAY, JUNE 3

BANQUET

Cocktails 6:00 PM Dinner 7:00 PM

WEDNESDAY, June 4—9:00 AM

SESSION 18 HUMAN CANCER GENES (3)

Chairman: J. Sambrook, University of Texas Health Science Center

WEDNESDAY, June 4—2:00 PM

SESSION 19 GENE THERAPY (2)

Chairman: D.W. Martin, Jr., Genetech, Inc.

SUMMARY: **T. Caskey,** Baylor College of Medicine

AUTHOR INDEX

HUMAN GENETICS : THE MOLECULAR MAP

W.F. Bodmer, Imperial Cancer Research Fund, Lincoln's Inn Fields, London.

The complexity of the human genome is no longer beyond the grasp of molecular biology. The number of different gene products may be no more than 100,000 to 150,000, divided into a maximum of perhaps 10,000 to 15,000 families of related products, many of which may themselves be part of related super-families such as the immunoglobulins, major histocompatibility system, T receptor and other lymphocyte surface molecules. This means that in the foreseeable future we may be able to identify the complete set of genetic functions that contributes to any phenotype, be it a behaviour or a cancer. Knowing where genes are on the genetic map can make a major contribution to sorting out gene-function relationships, as in the elucidation of the significance of the specific translocations in leukaemias and lymphomas, the identification of genetic changes in retinoblastoma, the hunt for the cystic fibrosis gene, or simply using markers for screening or pre-natal diagnosis.

Polymorphic DNA probes now provide an essentially unlimited range of genetic markers for the search for linkage with well-defined Mendelian traits such as familial polyposis coli, or for the identification of genetic factors in less well-defined susceptibilities such as heart disease or cancer. Simply searching for a distortion in the expected Mendelian segregation pattern amongst family members sharing a trait can identify the location of a gene involved in determining the trait.

The HLA system, which is associated with a variety of chronic diseases with immune aetiology, is a good model for such analyses. It also illustrates the complex evolution of a gene cluster, and the importance of the distinction between genetic marker associations in families and in populations.

The ultimate challenge is to unravel the genetic contributions at the molecular level to the determination of common traits, including the major chronic diseases, heart disease, autoimmune disease and cancer, as well as to normal variations, for example in physical features and behaviour.

THE GENE MAP OF HOMO SAPIENS: STATUS AND PROSPECTUS:
Victor A. McKusick, Johns Hopkins University, Baltimore, Maryland

The chromosome that carries each of about 900 structural genes is known and many of these genes have been fairly precisely regionalized. This number represents 47% of the well established loci catalogued in Mendelian Inheritance in Man and 23% of all loci catalogued there. It is less than 2% of the total 50,000 genes estimated for man. In addition, about 500 anonymous DNA segments have been mapped; many of these have sequence variation useful as linkage markers. In order, the most contributory mapping methods have been somatic cell hybridization in all its variations (445 autosomal loci), family linkage study (147 autosomal loci), and in situ hybridization (131 autosomal loci). The centrality of gene mapping in human biology and medicine is reflected in the great diversity of gene products represented by the genes mapped.

The information has revealed frequent clustering of genes of related function; extensive homology of synteny even with species as remote as the mouse; and general rules such as the following: the subunits of heteromeric proteins are coded by different chromosomes; mitochondrial and cytosolic isozymes are coded by different chromosomes; genes encoding enzymes in the same metabolic pathway are usually not syntemic. The gene map has played an important role in validation of the chromosome theory of cancer; viz., specific morphologic changes in the chromosomes consistently associated with particular neoplasms and the correlation of these changes with the location of oncogenes and of specific changes in DNA. Most cancers are somatic cell genetic diseases. Many congenital malformations and autoimmune diseases may be somatic cell genetic diseases. Mapping may help in the understanding of these also.

Mapping is important to diagnosis by the linkage principle. It is also important to delineation of the basic defect by "reverse genetics" (going to the DNA pinpointed by the mapping) and by the candidate gene approach: Is the disease phenotype linked to a RFLP of a given cloned gene? Does the disease phenotype map to the same region as a candidate gene?

Complete sequencing of the human genome will build on complete mapping of the structural genes of man. The dimensions of the task and the information it will engender are useful to consider.

CHROMOSOME CLASSIFICATION AND PURIFICATION USING FLOW CYTOMETRY AND SORTING. J.W. Gray, G. van den Engh, B. Trask, J. Lucas, D. Pinkel, M. Van Dilla, J. Fuscoe and L. Yu. Biomedical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA 94550

Flow cytometry has proved to be a powerful tool for classification of chromosomes from most mammalian cells according to their DNA content, DNA base composition, shape and/or protein composition. In this approach, called flow karyotyping, chromosomes are isolated, stained with one or more fluorescent dyes (eg. DNA specific dyes, monoclonal antibodies against chromatin components and possibly fluorescently labeled nucleic acid probes) and processed flow cytometrically. In conventional flow cytometry, chromosomes are classified according to the intensity of fluorescence emitted as they pass one by one through laser beams whose wavelengths are adjusted to excite the dyes. Scanning flow cytometry allows measurement of the distribution of dyes along the chromosomes.

Applications of flow karyotyping include identification of normal chromosomes and detection of aberrant chromosomes in human lymphocytes, fibroblasts, amniocytes, chorionic villus cells and numerous malignant cell types. Scanning flow cytometry has also proved useful in quantification of the frequency of dicentric chromosomes from irradiated cells as a measure of induced chromosome damage.

Flow sorting allows purification of chromosomes that can be resolved in flow karyotypes. The purity of these sorts can be greater than 95%. Chromosomes can be processed for sorting at rates up to a few thousand per second in conventional sorters and at rates up to 30,000 per sec using high speed sorting. The sorting process seems to cause little damage to the chromosomal DNA beyond that produced during chromosome isolation. Sorted chromosomes have proved useful for gene mapping and as sources of DNA for production of chromosome-specific recombinant DNA libraries. Sources for the sorted chromosomes include human fibroblasts, lymphoblastoid cells and inter-species hybrid cells.

Work performed under the auspices of the US DOE contract number W-7405-ENG-48 with support from USPHS grants HD 17665 and GM 25076.

CYTOGENETIC ANALYSIS BY IN SITU HYBRIDIZATION WITH FLUORESCENTLY LABELED NUCLEIC ACID PROBES. D. Pinkel[1], J. W. Gray[1], B. Trask[1], G. van den Engh[1], J. Fuscoe[1], and H. van Dekken[1,2]; [1]Biological Sciences Division, Lawrence Livermore National Laboratory, Livermore, Ca, 94550; [2]Radiobiological Institute, TNO, The Netherlands.

The rapidly increasing availability of nucleic acid sequences homologous to specific chromosomal location, make in situ hybridization an exciting approach for identifying chromosomes in metaphase spreads and nuclei. Fluorescence detection of hybridized probes is advantageous because of the high spatial localization of the signal, ease of quantitative analysis, and the opportunity for simultaneous use of multiple probes, each with a different fluorochrome.

Genomic DNA can be used as a species-specific probe for the chromosomes of that species. Using human-hamster hybrids we have demonstrated that the probe fluorescence intensities are proportional to the amount of target sequence in chromosomes and nuclei, that interspecies translocations can be detected with high sensitivity, and that chromosomes of both species can be simultaneously labeled with different fluorochromes using genomic probes modified with biotin and AAF respectively. Hybridization has been accomplished in suspension, with probe binding quantitated flow cytometrically.

Human chromosome-specific repetitive probes for a growing number of chromosomes have been discovered. Binding of these probes is usually concentrated to limited regions of the chromosome, such as the centromeres, making them ideal for analysis of interphase cells. We have demonstrated detection of XYY amniocytes using a Y specific probe, and have shown the ability to determine the three dimensional localization of both homologues of chromosome 18 in nuclei using optical sectioning. Repetitive sequences with target sizes below 50 kb, such as the ribosomal RNA genes can be reliably detected.

We are now developing techniques for use of collections of unique sequence probes to label individual chromosomes along their length. This will permit identification of translocations and aneuploidy where banding analysis is inconvenient or impossible. These composite probes will also be useful for analysis of the positions of chromosomes in interphase nuclei. Work performed under auspices of U. S. DOE under contract no.W-7405-ENG-48 with support from USPHS grant HD 17665.

CONSTRUCTION OF HUMAN CHROMOSOME-SPECIFIC DNA LIBRARIES FROM FLOW SORTED CHROMOSOMES: L. Deaven[1] and M. Van Dilla[2], [1]Los Alamos National Laboratory, Los Alamos, New Mexico 87545; [2]Lawrence Livermore National Laboratory, Livermore, CA 94550

The Los Alamos and Lawrence Livermore National Laboratories' biomedical research groups are currently constructing chromosome-specific DNA libraries from each of the human chromosomes. These libraries are available to the international scientific community to aid investigators in studies of gene mapping, genetic disease diagnosis, linkage and pedigree analysis and related areas of research. Chromosomes are isolated from normal human fibroblasts and human lymphoblastoid cells (#13 and smaller) or from Chinese hamster-human hybrids (#1-12) for flow sorting. DNA from flow sorted chromosomes is digested to completion with Eco R1 (Los Alamos) and Hind III (Livermore) and cloned into Charon 21A (acceptance range: 0-9.1 kb). Thus, the first libraries will consist of two complete sets, each in the same vector, but from cuts with different restriction enzymes. Library purity is determined by flow histogram analysis, cytogenetic analysis of sorted chromosomes, and analysis of cloned DNA. Each of the 24 human chromosomal types has been sorted, and libraries have been constructed for at least one restriction cut for each chromosome except number 1. These libraries are currently available from the American Type Culture Collection, Rockville, MD. We expect to have two complete sets of small insert libraries completed by 6/86.

We are currently doing preliminary studies on the construction of a set of large insert libraries to be cloned into bacteriophage vectors (about 20 kb) or cosmids (about 40 kb) with characteristics that are better suited to basic studies of gene structure and function. The first large insert libraries should be available in the later part of 1986. This work is supported by the U.S. Dept. of Energy.

CONSTRUCTION OF HUMAN GENETIC LINAKGE MAPS. I. PROGRESS AND PERSPECTIVES; R. White, C. Julier, M. Leppert, Y. Nakamura, P. O'Connell, S. Woodward, G.M. Lathrop, and J.-M. Lalouel, Howard Hughes Medical Institute, Salt Lake City, Utah.

The last several years have seen considerable progress in development of methodologies for the construction of genetic linkage maps of human chromosomes. Families whose structure is optimal for revealing linkage relationships have been sampled and archived. The associated genotypic database and cell lines are available through the CEPH in Paris, and form the basis for an international collaboration that will be described. New protocols are allowing characterization of a number of new loci with hypervariable regions; these highly polymorphic markers for human linkage studies will be described also.

The characteristics of primary maps spanning chromosomes 12, 13, and X will be reported. Differences in recombination frequencies between male and female meioses force the conclusion that recombination is not directly proportional to physical distance, and a model for these differences will be discussed.

Recent progress in developing family resources for the study of several genetic diseases, including retinitis pigmentosa, familial polyposis, ataxia telangiectasia, and FSH dystrophy will be presented. Finally, evidence supporting very tight linkage between the met locus and cystic fibrosis will be reviewed.

CONSTRUCTION OF HUMAN GENETIC LINKAGE MAPS. II. METHODOLOGICAL ISSUES; J.-M. Lalouel, G.M. Lathrop, and R. White, Howard Hughes Medical Institute, University of Utah Medical Center, Salt Lake City, Utah.

As the number of genetic markers which can be defined at the molecular level increases, extensive linkage groups that cover entire chromosomes can be developed. Efficient construction of linkage maps requires that genotypes be characterized on a common panel of families of optimal structure. By following Mendelian segregations at several loci jointly, new and powerful analytical strategies can be brought to bear for the detection of linkage, the estimation of recombination, and the determination of gene order. As human linkage maps expand, new multilocus strategies become necessary to add a new locus to a genetic map and to meet the computing challenge resulting from the large number of possible genetic outcomes which have to be accounted for. Current experience will serve to illustrate these points and to highlight areas where further methodological advances are needed.

MAPPING COMPLEX GENETIC TRAITS IN HUMANS: NEW METHODS USING A LINKAGE MAP OF RFLPS: E. Lander[1,2,3] and D. Botstein[2], [1]Whitehead Institute for Biomedical Research, Cambridge MA; [2]Department of Biology, Massacussetts Institute of Technology, Cambridge, MA; [3]Harvard University, Cambridge MA.

The continuing accumulation of hundreds of restriction fragment length polymorphisms (RFLPs) in the human genome has made it increasingly straightforward to map simple Mendelian traits given adequate numbers of human pedigrees.

Many human trait of medical or biological interest, however, show complex patterns of inheritance, for reasons of genetic heterogeneity, variable penetrance, synthetic interactions, and/or environmental interactions. Traditional methods of linkage analysis (which rely on mapping a trait relative to a single marker in families with multiple affected members) perform poorly when faced with such complications, even when highly informative RFLP markers are used.

We have devised new linkage methods that in theory allow more efficient mapping of such complex genetic traits. These methods exploit the additional information contained in a complete, pre-existing linkage map of the RFLPs; completion of such a map can be anticipated in the near future. The new techniques include "simultaneous search" for mapping each of several loci capable singly of causing a trait (i.e. the case of genetic heterogeneity) and "homozygosity mapping" for exploiting additional information that can be gleaned from consanguineous affected individuals (useful when diseases are very rare).

Considering several human diseases with different modes of inheritance, we have computed how large a pool of families will be required to map the genes involved, using the new methods based on complete RFLP linkage maps. In this context, we shall discuss what aspects of human heredity should eventually be within the power of linkage mapping.

APPROACHES TO PHYSICAL MAPPING OF THE HUMAN GENOME

C. R. Cantor and C.L. Smith, Department of Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, NY 10032

Four techniques have been developed that together should make it possible to begin to construct a coarse restriction map of the human genome. Each technique was developed or tested on simple organisms like yeast or bacteria but each has now been shown to be feasible for comparable studies on human samples.

High molecular weight DNA samples can be prepared by cell lysis in situ in agarose and extensive protease, detergent and salt treatments. The resulting DNA samples show a negligible level of double strand breaks and are stable indefinitely. These samples can be cut into discrete large DNA fragments in situ by treatment with various restriction nucleases or nuclease-methylase combinations. Appropriate choice of enzymes can yield fragments averaging anywhere from a few hundred kB in size to more than a thousand kB. Pulsed field gel electrophoresis affords excellent separations of these large fragments of human DNA. The electrophoretic separations can be visualized by ordinary Southern blotting and hybridization with cloned single copy probes. In this way, large genomic DNA fragments encompassing any desired known small DNA sequence can be identified.

For construction of physical maps, several different enzyme digests can be overlapped just as in conventional restriction enzyme analysis. However, much more efficient strategies have been designed that take unique advantage of the availability of large DNA fragments. For example, it is possible to select from a library just those junction clones that contain a rare restriction enzyme cutting site. Probing a genomic digest of DNA generated by the same enzyme will reveal two large fragments. These must be adjacent, and thus a physical map of the distances between all rare cutting sites can be generated systematically, by combining his approach with the use of other specially selected libraries. Examples of all of these techniques will be demonstrated and estimates of the feasibility of applying them to generate complete physical maps of individual human chromosomes will be described.

THE HUMAN MHC-APPROACHES TO CHARACTERIZATION OF LARGE REGIONS OF DNA:

S.M. Weissman[1], S.K. Lawrance[1], R. Srivastava[1], M.J. Cherney[1], B. Kigas[1], H. Vasavada[1], G.A. Gillespie[1], C. Smith[2], C. Cantor[2], and F.S. Collins[3], [1]Department of Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510. [2]Department of Human Genetics, Columbia University, 701 W. 68th Street, New York, NY 10032. [3]Department of Internal Medicine, University of Michigan Medical School, 4708 Med. Sci. II - Box 015, Ann Arbor, MI 48103.

The human major histocompatibility complex is an extended genetic region covering several centimorgans of DNA containing genes for cell surface and secreted molecules important in various aspects of immune function. We have directly cloned about 1.2 million base pairs of this complex in cosmids and in addition are developing methods that would expedite cloning and characterization of other large gene complexes and are potentially applicable to global characterization and cloning of the human genome. The approaches that we have used include: 1) Application of chromosome hopping libraries to obtain probes scattered at a distance from the initial probe; 2) PFGE (Pulse Field Gel Electrophoresis) fractionation and DNA blotting to obtain a course scale restriction map of the human MHC; 3) Development of convenient restriction mapping methods that are generally applicable; these including methods that are in principal indeperdent of gel electrophoretigrams, as well as approaches for determining restriction map from single partial digests with single enzymes without direct or indirect labeling; 4) With Drs. D. Ward and C. Padding, improved methods for physical isolation of DNA fragments or clones prior to or in place of the use of colony or plaque hybridization techniques.

We have found a number of genes for the human MHC including a family that are homologous to the murine TL genes and additional genes or pseudogenes of the class II type beyond those described in literature. So far, we have obtained a coarse restriction map for about 3.1 million base pairs of the human MHC and are currently involved in mapping the short arm of the human chromosome 6. In addition, we have developed a rapid procedure for identifying tissue-specific enhancers in cloned genes.

MOLECULAR APPROACHES TO MAMMALIAN GENETICS:

Hans Lehrach, Annemarie Poustka, Thomas Pohl, Denise Barlow, Alister Craig, Guenther Zehetner and Anna-Maria Frischauf, EMBL, Heidelberg, Germany

To approach a molecular analysis of genes defined by mammalian (mouse and human) mutations and to derive combined physical and genetic maps of mammalian chromosomes, cloning and DNA analysis techniques capable of analysing genetic distances in mammals have been developed and applied.

The main new developments in this direction are the construction and use of "chromosome jumping" libraries derived by selective cloning of the ends of large DNA fragments created by different combinations of rare or commonly cutting restriction enzymes and the derivation and use of "junction fragment" libraries containing fragments bridging rare restriction sites.

The application of these techniques to specific problems, as well as the progress in developing techniques potentially capable of deriving ordered cosmid libraries from mammalian chromosomes or mammalian genomes will be discussed.

REDUCED RECOMBINATION RATE IN CHROMOSOMES 21 WHICH UNDERGO NONDISJUNCTION IN TRISOMY 21. S.E. Antonarakis[1], A.C. Warren[1], C. Wong[1], C. Metaxotou[2], A. Chakravarti[3]. [1]Dept. Peds, Johns Hopkins Univ. Sch. of Med., Baltimore, MD, [2]Dept. Peds., Athens Univ. Sch. of Med., Athens, Greece, [3]Dept. Biostatistics, Univ. of Pitts., Pitts., PA.

We have used DNA polymorphisms adjacent to single copy DNA fragments which map on human chromosome 21 to create a linkage map of this chromosome. The DNA fragments used were: D21513 (B) cloned by K. Davies et al; D2151 (C), D21511 (C), D2153 (D), D21523 (D) cloned by P. Watkins et al; SOD-1 cloned by Y. Groner et al; and cW21pc (A), a 700 nt fragment which maps very close to the centromere and is the short arm junction fragment from a ring chromosome 21. The DNA polymorphisms associated with these fragments were: Taq I-D21513, BamHI-D2151, Msp I-D2151, EcoRI-D21411, Taq I-D21511, Taq I-D2153, Hind III-D2153, EcoRI-D21523, Msp I-SOD-1, Sst I-cW21pc, Hinc II-cW21pc. Linkage analysis using these markers in appropriate informative families revealed the following linkage map:

21pter--A,cen—14 cM—B--17 cM—C—50 cM—D—21qter.

We have then used chromosome 21 DNA polymorphism markers in families with trisomy 21 to investigate the recombination rate in chromosomes which undergo nondisjunction (NDJ). DNA from members of 35 Greek families with a trisomy 21 and at least one normal offspring has been examined. When the map distances (y) of these markers from the centromere were computed using the nuclear trisomy 21 families, the following distances were found: y_A=0, lod 1.14; y_C=0, lod 1.13; y_D=0, lod 0.54. We then examined the recombination between these markers in chromosomes 21 which participated in NDJ (CHR-NDJ) vs. chromosomes 21 which did not participate in NDJ (CHR-NO-NDJ) in the informative trisomic individuals as shown below. (DNA markers were informative and the origin of NDJ was known in 22/35 families.) These results suggest that crossing-over is strikingly suppressed in chromosomes 21 which participate in NDJ and may reflect defective pairing of these chromosomes in first meiotic division.

A,cen——B——C————D

	recomb/inf. meioses	
	CHR.NDJ	CHR.NO.NDJ
———	0/4	0/2
———	0/2	0/1
———————	0/5	7/11
———————	0/5	3/9
———————	1*/3	0/1
———————	1/2	1/1
Total	2/21	11/25 χ^2:6.7 p<0.01

*the recombination occurred in chromosomes of a maternal meiosis 2 error.

THE HUMAN *MET* ONCOGENE IS A MEMBER OF THE TYROSINE KINASE FAMILY: M. Park[1], M. Dean[1], K. Kaul[1], M. Gonzatti-Haces[1], A. Iyer[1], T. Robins[1], D. Blair[2], and G.F. Vande Woude, [1]LBI-Basic Research Program, NCI-Frederick Cancer Research Facility, Frederick, Maryland; [2]Laboratory of Molecular Oncology, National Cancer Institute, Frederick, Maryland.

Prolonged exposure of a nontumorigenic human osteogenic sarcoma cell line (HOS) with the direct acting carcinogen N-methyl-N'-nitronitrosoguanidine (MNNG) gave rise to morphologically distinct cells which were tumorigenic in nude mice (MNNG-HOS). We have shown that DNA from MNNG-HOS cells will transform NIH/3T3 cells and have isolated >35 kb of human DNA containing an oncogene, termed *met*. The activated *met* oncogene in MNNG-HOS cells and in *met* transformed mouse cells expresses a novel 5.0 kb RNA transcript which is a fused hybrid RNA derived from a DNA rearrangement involving two distinct genetic loci termed *met* and *tpr* (translocated promoter region). The *met* proto-oncogene has been localized to 7q21-31 by *in situ* hybridization. This locus expresses a 9.0 kb RNA in fibroblast and epithelial cell lines, but is not commonly expressed in cell lines derived from the hematopoietic cell lineage. In contrast, the *tpr* locus, which has been mapped to chromosome 1, expresses a 10.0 kb RNA in all human cell lines tested. Although the novel 5.0 kb *met* oncogene RNA is largely derived from (>80%), and is 3' co-terminal with, the 9.0 kb *met* proto-oncogene RNA, the 5' portion of this RNA uses some exons from the 10.0 kb *tpr* RNA. These exons are small and are presumably in the promoter region of both transcripts. Nucleotide sequence analysis of the 3' end of *met* shows that it is a member of the tyrosine kinase family of genes. This coding region is most homologous to the human insulin receptor and murine v-*abl* genes. Peptide antibody to the c-terminal coding region of *met* in immunoprecipitation analyses detects a 65 kilodalton polypeptide in both MNNG-HOS cells and *met* transformed NIH/3T3 cells which can be phosphorylated *in vitro*. The *met* oncogene activation in MNNG-HOS cells therefore appears to result from a DNA rearrangement possibly mediated *in vitro* by MNNG. The mode of activation of *met* may therefore be similar to EGF^R/*erbB* or *bcr*/c-*abl* in the Philadelphia chromosome translocation in chronic myelogenous leukemias.

C-erbB-2 GENE ENCODES A RECEPTOR LIKE PROTEIN WITH TYROSINE KINASE ACTIVITY

K.Toyoshima[1], K.Semba[1], T.Akiyama[2], S.Ikawa[1], and T. Yamamoto[1]. [1]Institute of Medical Science, University of Tokyo, Tokyo 108, [2]Meiji College of Pharmacy, Tokyo 154.

The c-erbB-2 gene was first found as a v-erbB related sequence in human genomic library. It was mapped to human chromosome 17q21 and produces a 4.6 kb mRNA. Nucleotide sequnce analysis of c-erbB-2 cDNA clones showed that the gene encodes a possible receptor protein with calculated molecular weight of 137,895. The antibody raised against a synthetic peptide corresponding to 14 amino acid residues of the carboxyl terminus could immuno-precipitate a 185k dalton glycoprotein. The immuno-precipitate showed protein kinase activity specific to tyrosine residue. In spite of the extensive structural similarity to EGF receptor, tyrosine kinase activity of this protein was not stimulated by addition of EGF either *in vitro* or *in vivo*. PDGF, FGF, TGF-α, TGF-β, NGF and erythropoietin also failed to activate the kinase activity.

This gene was found to be conserved at least in mammalian and avian. Comparison of amino acid sequence of c-erbB-2 to that of *neu* gene published by Bargmann et al. shows about 88 % homology indicating that *neu* gene is the rat counterpart of c-erbB-2 gene.

Involvement of this gene to human cancer was observed in 6 out of 66 adenocarcinomas as amplification of the gene. However, we did not observe amplification in squamous cell carcinoma in which amplification of EGF receptor gene was often observed. Shubert et al. also reported amplification of this gene in a human mammary carcinoma. These observations suggest that the increased expression of c-erbB-2 may provide growth advantage to adenocarcinomas.

This study was supported by a GRANT-in-Aid for Special Project Research, Cancer-Bioscience from the Ministry of Education, Science and Culture of Japan.

ras and trk: TWO TRANSFORMING GENES OF HUMAN TUMORS: M. Barbacid, S. Sukumar, D. Martin-Zanca, S. Katzav and G. Mitra. Basic Research Program, NCI-Frederick Cancer Research Facility, Frederick, Maryland 21701

Genetic alterations in at least fifteen human loci have been shown to contribute to the development of human neoplasia. To date, members of the ras gene family are the oncogenes most frequently identified in human tumors. In order to elucidate the role of ras oncogenes in the complex multistep process of carcinogenesis, we have turned our research efforts to animal model systems. Induction of mammary carcinomas in rats by injection of a single dose of nitroso-methyl-urea (NMU) during sexual development involves the reproducible activation of the H-ras oncogene. Each of these oncogenes acquired their transforming properties by a G → A transition, the type of mutation most commonly induced by NMU. Based on these observations, we have postulated that in this model system, H-ras oncogenes are activated during initiation of carcinogenesis as a direct consequence of the mutagenic activity of NMU. I will present recent evidence suggesting that (a) ras oncogenes can be activated at various times during the life span of the animal. However, the phenotypic expression of their transforming properties must await cellular proliferation (e.g., hormone-induced proliferation of the mammary gland during female sexual development); (b) activation of ras oncogenes does not predetermine further genetic changes that occur during tumor development (e.g., loss of hormone dependency).

Approximately half of the known human oncogenes have been identified by their ability to transform rodent cell lines in gene transfer assays. One such transforming gene, trk, has been recently characterized in our laboratory. This oncogene was generated by a somatic rearrangement during the development of a colon carcinoma. Nucleotide sequence analysis suggests that the trk oncogene was created by substitution of the external domain of a putative tyrosine-kinase growth factor receptor by the 221 amino terminal residues of a non-muscle tropomyosin molecule. We are currently investigating the biochemical properties of the trk gene product and the nature of the human proto-oncogene that participated in the generation of trk.

TUMORIGENICITY ASSAYS FOR HUMAN PROTO-ONCOGENES: M. Wigler, [1]O. Fasano, [2]D. Birnbaum, C. Birchmeier and D. Young. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724. [1]EMBY Meyerhof Strasse 1, D 6900 Heidelberg, W. Germany. [2]U.119 I.N.S.E.R.M., 27 Bd. Lei Roure, 13009 Marseille, France.

We have attempted to develop a new assay for human oncogenes based on the tumorigenicity of cotransfected NIH3T3 cells in nude mice, using human tumor DNA as donor. While we sometimes can detect genetic abnormalities in donor DNA, we have also observed that de novo activation of proto-oncogenes due to rearrangement or gene amplification can occur following gene transfer. One such event involved the activation of ros1, a human homolog of the avian v-ros oncogene. This gene appears to encode a transmembrane receptor of the tryosine kinase family. A second event involved the activation of a previously unknown proto-oncogene which we call mas1. This gene encodes a protein with multiple potential transmembrane domains, and resembles the visual rhodopsins in overall structure.

STRUCTURE OF THE ACTIVATED C-RAF-1 GENE FROM HUMAN STOMACH CANCER: K. Shimizu[1], Y. Nakatsu[1], M. Oh-uchida[1], S. Nomoto[1], and M. Sekiguchi[2], [1]Department of Biology, Faculty of Science, [2] Department of Biochemistry, Faculty of Medicine, Kyushu University, Fukuoka 812, Japan

We have isolated a novel human transforming sequence from a primary stomach cancer. Of 57 Kirobase-pairs (Kbp) of the human sequence isolated, a region of 39 Kbp was found to be a minimum functional unit for the transforming activity, because a cosmid clone harboring this region was capable of inducing foci upon transfection.

Direct tests for homology with known retroviral and cellular oncogenes revealed that the transforming sequence contains the c-raf-1 gene, a human homologue of v-raf, first found in 3611-murine sarcoma virus. The v-raf gene is known to encode a serine/threonine-specific protein kinase. The size of the transcript of the transforming c-raf-1 gene was estimated about 2.8 kb and we have isolated a cDNA clone which covers about 85% of the message.

Analyses of cDNA and genomic clones of this gene revealed that the gene was generated by substitution of 5'-sequence (exons 0-4) of normal c-raf-1 gene with an unrelated human sequence. We identified a region in the genomic clone, where the rearrangement has occurred. Moreover, the substitued cDNA sequence is composed of an open reading frame, which joins to the exon 5 of c-raf-1 gene in an in-phase manner. The substitued open reading frame encodes an extremely hydrophobic polypeptide. The main body of the cDNA, composed of exons 5-16 of c-raf-1 gene, had no base-substitution compared with normal c-raf-1 sequence.

These results suggest that the mechanism(s) of activation of the transforming c-raf-1 gene may involve either altered regulation of expression or altered structure/function of the c-raf-1 gene-product itself. Reproducible occurrence of foci-induction in two independent experiments suggests that the activated c-raf-1 gene has been generated in the primary stomach cancer. Experiments with plasmids designed to express the cDNA sequence as well as searches for structurally altered c-raf-1 gene in other human cancers are in progress.

P450 GENES: STRUCTURE, EVOLUTION, REGULATION, AND RELATIONSHIP TO CANCER: Frank J. Gonzalez[1], Anil K. Jaiswal[2] and Daniel W. Nebert[2], [1]NCI and [2]NICHD, NIH, Bethesda, Maryland 20892.

P450 proteins are monooxygenases important in the biosynthesis of steroids, fatty acids, prostaglandins, leukotrienes, biogenic amines, pheromones and phytoalexins. In addition to endogenous substrates, these hemoprotein enzymes metabolize innumerable drugs, chemical carcinogens and other environmental pollutants. At least six P450 gene classes have been documented so far by sequencing and chromosomal mapping; the protein encoded by a gene in one class is 15-35% similar in amino acid sequence to the protein encoded by a gene in any of the other five classes. P450 class I genes include P_1450 and P_3450; positive and negative trans-acting regulatory control regions have been identified in the mouse and human P_1450 upstream sequences. In laboratory animal studies, certain procarcinogens are known to be converted by specific P450s to ultimate carcinogenic intermediates: combustion products including benzpyrene for P_1450; two tryptophan-derived pyrolysis products and acetylaminofluorene and aminobiphenyls for P_3450. In man it appears that enhanced P_1450 inducibility is correlated with increased risk of cigarette smoking-induced bronchogenic carcinoma. Human cultured lymphocytes treated with mitogens plus 3-methylcholanthrene exhibit inducible P_1450 activity [aryl hydrocarbon (benzo[a]pyrene) hydroxylase (AHH)]; genetic differences in AHH inducibility exist in the human population and are highly correlated with P_1450 mRNA concentrations that have been quantitated by slot blot analysis. Near-full-length cDNA and/or genomic clones for human P_1450 and P_3450 and more than ten other P450s have been isolated, sequenced, and used as probes to localize their chromosomal sites. It might be possible to correlate restriction fragment length polymorphism (RFLP) patterns of one or more of these genes with increased risk of cancer caused by a certain class of environmental chemicals.

T-CELL RECEPTOR GENES: ORGANIZATION, DIVERSIFICATION AND REARRANGEMENTS. L. Hood[1], P. Concannon[1], M. Klein[3], E. Lai[1], G. Siu[1], E. Strauss[1], and L. Pickering[2], [1]Division of Biology, California Institute of Technology, Pasadena, California 91125. [2]T-Cell Sciences, Inc., Cambridge, Massachusetts 02139. [3]Toronto Cancer Institute, Toronto, Canada.

Human T-cell receptor genes present a fascinating opportunity for study in several regards. First, they constitute a two multigene families, which encode the α and β chains of the T-cell receptor. Second, they are divided into gene segments (variable and joining or variable, diversity and joining) which rearrange and join together to generate a V coding region. Third, there are a multiplicity of V, D and J gene segments in each of these gene families. Fourth, T cells play an important role in immunologically mediated disease, and accordingly, an understanding of the diversity in these gene families will probably eventually provide important insights into the pathophysiology of diseases. With these issues in mind we undertook to characterize in detail the nature of the α and β T-cell receptor gene families.

We have collaborated with Dr. Tak Mak and his colleagues characterizing the duplicated D-J-C region clusters of the β gene family. We then analyzed 27 β cDNA clones derived from peripheral blood cells and found that the V_β gene segments of the cDNA clones fell into 14 distinct subfamilies. This analysis revealed a variety of mechanisms for diversification including the D_β-D_β joining. We also analyzed in detail one subfamily containing five members and made the interesting evolutionary observation that the T-cell receptor V gene segments appear to be evolving at about the same rate other genes. We have used the new technique of pulsed field gradient gel electrophoresis to analyze the linkage relationships among members of the V_β gene segments. We have also studied in some detail their polymorphism among different individual humans. These observations suggest that humans have about 60 V_β gene segments. Similar studies are now being carried out with the α gene family. These molecular probes should provide some unusual opportunities for analyzing T-cell associated diseases in the future.

GENES OF THE T CELL ANTIGEN RECEPTORS IN NORMAL AND MALIGNANT T CELLS: T.W. Mak, Y. Yoshikai, I. Iwamoto, M. Minden, P. Ohashi, N. Caccia and B. Toyonaga, The Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada.

Antigen recognitions for the cell mediated immune responses are known to be mediated by T-lymphocytes; until recently, the molecular mechanisms underlying T-cell recognition had eluded investigation. In the past two years, we and others have isolated genetic information for the α- and β-chain of the cell receptor on T-lymphocytes (T-cell receptor, TcR). Molecular analysis indicate that α- and β-chains of the TcR are genes distinct from immunoglobulin being only partially homologous to Igs. Furthermore, the genes coding these proteins are situated at different locations of the human and mouse chromosomes. The organization of these genes, however, are similar to those of Ig genes being composed of non-contiguous segments of variable, joining, and constant region segments. Using gene transfer technology, we have also been able to reconstitute a biologically active T cell receptor by DNA transfer of T cell receptor genes. This finding provides direct support for the hypothesis that the α-β heterodimer encodes the T cell antigen receptor. With the use of the T cell receptor genes, we have also examined cells from hemopoietic malignancies and allied disorders from somatic rearrangements of these genes. These studies should help in the diagnosis and maintenance of patients with T cell malignancies and lymphocytosis. In addition, these genes are now known to be involved in activating known (e.g. c-myc) and "new" "oncogenes" in malignant cells with chromosomal translocations.

HUMAN HISTOCOMPATIBILITY ANTIGENS: STRUCTURE, GENETICS AND REGULATION OF EXPRESSION: Jack L. Strominger, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138

The human Major Histocompatibility Complex (MHC) encoded on chromosome 6 can be subdivided into three regions, termed Class I, Class II and Class III. Class I genes encode the HLA-A, -B and -C antigens. Class II genes encode the HLA-DR, -DQ, and -DP antigens and Class III genes encode some components of the complement system. Class I and Class II antigens function as restricting elements for the elimination of virus infected cells and for presentation of foreign antigens for antibody formation, respectively. Structurally they belong to a large family of proteins composed of 'immunoglobulin-like' domains. The following topics will be discussed:

1. More than 400 kb of DNA from the human Class II region has been cloned. Fourteen α and β chain genes have been described, twice as many as in the mouse.
2. The regulation of expression of these Class II genes in vivo (for example, γ-interferon induction) and the cis- and trans-acting factors involved in this regulation will be discussed and is related to a newly discovered immunodeficiency disease in children called the Bare Lymphocyte Syndrome.
3. The expression of Class I genes is specifically up-regulated by tumor necrosis factor and may be related to immunosurveillance of tumors.
4. Insight into the functioning of Class I antigens has been obtained by the study of naturally occurring CTL variants and by site-directed mutagenesis.
5. Two genes encoding 21-steroid hydroxylase are linked to the C4A and C4B genes in the Class III (complement) cluster. A deletion of one of these (21-B) is associated with congenital adrenal hyperplasia, and the deletion of the other (21-A), associated with the -A1,-B8,-DR3 haplotype, is phenotypically normal.

HUMAN Ia ANTIGENS: COMPLEXITY, POLYMORPHISM AND REGULATION

B. Mach, P. Rollini, J. Berdoz, C. Berte, W. Reith and J. Gorski, Department of Microbiology, University of Geneva Medical School, Geneva, Switzerland

HLA-class II genes code for highly polymorphic heterodimers (α and β chains) and, as target for T cell recognition, are essential for the immune response and for autoimmunity. Genetic complexity: A large number of structurally related α and β chain genes have been identified and mapped within the HLA-D region on chromosome 6. Polymorphism: The extensive allelic polymorphism can be accounted for by gene duplication divergence and gene conversion, creating new specificities, while maintaining preselected epitopes. The human HLA-class II polymorphism can be analyzed in population studies by hybridization with selected oligonucleotides. This provides an essential tool in studies of HLA-disease association. Regulation: There is a global regulation of the entire HLA-D region. Studies with congenital immunodeficiency patients (SCID) have demonstrated the existence of a transacting class II regulatory gene, defective in the patients. HLA and disease association: An autoimmune reaction can be initiated by the aberrant expression of certain Ia alleles, possibly in the form of hybrid α/β dimers, on cells that normally do not express Ia antigens.

POLYMORPHISM IN THE CLASS II REGION OF THE HUMAN MAJOR HISTOCOMPATIBILITY COMPLEX. Hugh O. McDevitt and John I. Bell, Departments of Medicine and Medical Microbiology, Stanford University, Stanford, California, 94305

The Class II region of the human Major Histocompatibility Complex (MHC) encodes a set of highly polymorphic cell surface glycoproteins involved in regulation of the immune response. The products of the genes in this region have been previously defined serologically and certain alleles have been demonstrated to be strongly associated with susceptibility to a wide variety of autoimmune diseases. Using Restriction Fragment Length Polymorphisms (RFLP) and allelic nucleic acid sequencing, we have studied the polymorphism present in the DR and DQ subregions of the MHC, and have begun to define the basis for disease association at a molecular level. We have used a collection of consanguineous lymphoblastoid lines homozygous at the MHC to study the standard RFLP patterns associated with DR beta, DQ beta and DQ alpha probes after digestion with a variety of enzymes. These studies have established that extensive RFLPs exist adjacent to all of these loci, and in part they reflect serological specificities. In addition, RFLPs define a large number of subtypes of serological haplotypes, many of them previously unrecognized. Based on our studies of multiple representatives of each DR haplotype, it appears likely that the amount of polymorphism in this region far exceeds that predicted serologically. We have extensively studied the DR haplotype where cellularly defined subtypes are known to exist with RFLPs. DR2 lines that represent Dw2, Dw12, tb24, AZH and Db9 types can be distinguished easily by RFLPs using both the DR beta and DQ beta probes. Allelic sequencing of DR beta chains from cDNA libraries made from consanguineous HTCs representing DR types 1-5 and 7, indicate that two DR beta chains are expressed in all but the DR1 haplotype. Of these, one beta chain in the DR3 haplotype is invariant compared to sequences from the DR5 and DR6 haplotypes. This indicates that the MT2 determinant may be encoded by this invariant chain. The variability between these DR beta chains is largely confined to three regions within the first domain between residues 9-13, 26-38 and 70-77. We have used this information on the extent and nature of Class II polymorphism to study a range of haplotypes associated with human diseases in an attempt to define the established disease associations. We have studied the frequency of different subtypes by serological specificities as defined by RFLPs in patients and controls. Insulin-dependent diabetes mellitus, celiac disease, myasthenia gravis and multiple sclerosis have been studied. We have observed a strong association between a rare subtype of DR3 identified with the enzyme Hinc II and the DQ beta probe and the disease myasthenia gravis. We have also demonstrated that the restriction fragment pattern seen in MS patients is that associated with the Dw2 subtype of DR2. We are currently obtaining sequence data from disease associated haplotypes to compare with our normal sequences to establish if disease associated alleles exist. The presence of such sequences would greatly improve our ability to study molecular interactions involved in the pathogenesis of disease and potentially to screen for susceptible individuals.

HLA CLASS II RFLPs ARE HAPLOTYPE SPECIFIC: M.R.J. Kohonen-Corish, H. Dunckley, M.A. Reid, S.W. Serjeantson, John Curtin School of Medical Research, Australian National University, Canberra. A.C.T.

The marked linkage disequilibrium between the HLA-B and -DR loci in the major histocompatibility complex of man is considered indirect evidence for the importance of natural selection in the evolution of this region of the genome. Our studies show that current linkage disquilibrium values for HLA-B and -DR, based on serological techniques, are gross under-estimates of the strength of the HLA-B and DR linkage relationships.

RFLPs in HLA-DRα, DRβ, DQα and DQβ were generated with five different restriction enzymes in a panel of 250 cells and revealed heterogeneity within DR types that was haplotype specific. For instance, using TaqI and the DRβ probe, four different patterns were seen on DR7-positive haplotypes, each in strong linkage disequilibrium with an HLA-B antigen, namely HLA-B50, B57, B13, or B44. Similarly, haplotype-specific class II RFLPs were seen for HLA-DR4, such that the DQβ 3.7kb BamHI fragment is in significant linkage disequilibrium (Δ=.023, P<.001) with B44.DR4, while the allelic 12kb fragment occurs on B62.DR4, (Δ=.033, P<.001) and B60.DR4 (Δ=.019, P<.001) haplotypes.

Following the method developed by R.A. Fisher to predict the order of the Rhesus blood group genes, the class II RFLPs and their linkage relationships with HLA-B are used to predict the order and orientation of the DQ and DR α and β chain genes.

INSULIN-DEPENDENT DIABETES MELLITUS IS ASSOCIATED WITH POLYMORPHIC FORMS OF THE T CELL RECEPTOR BETA CHAIN GENE

Marie L. Hoover, Ph.D., J. Donald Capra, M.D., and Kay Black

The association of the HLA-D alleles, HLA-DR3 and/or -DR4, with insulin dependent diabetes mellitus (IDDM) is well known. However, there is accumulating evidence that a second locus, unlinked to the HLA complex, may be involved in this disease. We and others have previously shown that the restriction enzyme Bgl II reveals polymorphic forms of the alpha and beta genes of the human T cell receptor. While the distribution of the polymorphic forms of T alpha was not significantly different in IDDM patients (N = 44) versus controls (N = 107), the T beta forms were. Twenty-one of the 22 DR3 or DR4 IDDM patients were positive for the 9.3 kb T beta allele whereas 77 of 107 controls were positive [relative risk (RR) 8.2, $p < 0.05$)]. In the patients that were DR 3,4, a correlation with the 8.6 kb allele was found in that 21 of 22 patients had the 8.6 kb allele as compared with 78 of 107 controls (RR 7.8, $p < 0.05$). This study documents for the first time an association between the allelic forms of a T cell receptor gene and a human disease. The ability to genetically type individuals for two unlinked loci involved in IDDM should significantly enhance our understanding of the genetic basis of this disorder.

THE CYSTIC FIBROSIS LOCUS: G. Bell, J. Bell, G. Bates, K.A. Davies, X. Estavill, M. Farrall, H. Kruyer, H.Y. Law, N. Lench, P. Scambler, P. Stanier, B. Wainwright, E. Watson, R. Williamson, Cystic Fibrosis Genetics Research Group, St. Mary's Hospital Medical School, University of London.

Recombinant DNA sequences are now available which allow the mapping of the entire human genome, but the first linkage to CF came using classical protein polymorphisms. The enzyme paraoxonase was shown to be loosely linked to CF (Copenhagen), to be followed quickly by six cloned DNA sequences: pJ3.11, 7C22, COL1A2 and TCRB (St. Mary's), 917 (Toronto) and MET (Salt Lake City). Both pJ3.11 and MET are very closely linked, and can be used for carrier detection and prenatal diagnosis in many informative families where there is a CF child. There is no evidence for heterogeneity of the CF locus. The collection of informative markers surrounding the CF locus is now sufficient to permit attempts to be made to isolate the defective gene using a combination of chromosome-mediated gene transfer, pulse field gel electrophoresis, cosmid mapping and chromosome walking techniques, although the difficulty of obtaining tissue in which the defect is known to be expressed remains a problem.

HIGHLY POLYMORPHIC RFLP PROBES AS DIAGNOSTIC TOOLS: H. Donis-Keller, D. Barker, R. Knowlton, J. Schumm, and J. Braman, Department of Human Genetics, Collaborative Research, Inc., Lexington, Massachusetts.

We identified more than 500 RFLPs by screening unique-sequence-containing clones from a human genomic library. Several hundred clones reveal more than 2 alleles and virtually all RFLPs tested represent single loci. More than 30 probes reveal RFLPs with PIC (polymorphism information content) values of $\geq$ 0.7 with a single enzyme; about half of these result from clustered site changes while the remainder appear to be DNA rearrangements. Genetic linkage relationships and chromosome distribution of the RFLP collection are being determined. Preliminary findings indicate that highly polymorphic genetic loci are distributed throughout the genome.

One RFLP (Lam4-917) identified for the first time the chromosomal origin of the cystic fibrosis locus. Further linkage studies utilizing the RFLP probe collection and chromosome 7 library probes have led to a high resolution genetic linkage map of chromosome seven and the CF locus. Several RFLP probes that closely flank the locus have been identified and can now be used for antenatal diagnosis of cystic fibrosis.

In addition to their use in genome mapping, highly polymorphic RFLP probes from this collection are diagnostically useful for determining donor or host cell origin of hematopoeitic cells in a recipient following bone marrow transplantation. This determination is otherwise difficult because the donor and recipient are almost invariably HLA matched and frequently siblings. These RFLP probes are especially useful in studies of hematopoietic chimerism (presence of both donor and recipient cells) since donor and recipient-specific patterns are routinely observed with a single probe-enzyme combination.

MAPPING OF THE CYSTIC FIBROSIS LOCUS ON CHROMOSOME 7:

L.-C. Tsui and M. Buchwald, Department of Genetics, The Hospital for Sick Children, Toronto, Ontario, CANADA.

Cystic fibrosis (CF) is the most common, severe autosomal recessive disorder in Caucasian children. Although the clinical symptoms- chronic pulmonary infections, pancreatic enzyme insufficiency and elevated sweat electrolyte levels- are suggestive of a general exocrine malfunction in the pateints, the basic defect of this disease remains unknown.

As a first step in our attempt to identify the genetic lesion, we recently demonstrated a linkage between *CF* and a randomly selected DNA marker in a large number of two-generation CF families [Tsui et al (1985) *Science* 230, 1054]. Subsequent mapping studies using human/rodent cell hybrids placed this DNA marker (*D7S15*) and therefore *CF*, on chromosome 7 [Knowlton et al (1985) *Nature* 318, 380]. Close linkage was also detected between *CF* and two other chromosome 7 markers, *met* [White et al (1985) *Nature* 318, 382] and *D7S8* [Wainwright et al (1985) *Nature* 318, 384].

To obtain a more precise chromosome location of *CF*, we have examined the linkage relationship between *CF* and a number of other chromosome 7 markers:

	CF	*met*	*D7S8*	*PON*	*D7S15*	*COL1A2*	*Epo*	*TCRB*	
CF	-	0.01	0.01	0.12	0.17	0.17	0.17	0.25	
met	19.28	-	0.02	0.22	0.18	0.17	0.12	0.16	
D7S8	19.58	16.75	-	0.25	0.21	0.22	0.15	0.14	
PON	2.88	1.44	0.67	-	0.06	0.13	0.10	0.50	θ
D7S15	6.68	3.55	2.89	6.47	-	0.05	0.07	0.22	
COL1A2	3.87	3.40	1.82	2.33	13.25	-	0.07	0.25	
Epo	0.87	2.87	1.16	2.51	5.45	4.86	-	0.15	
TCRB	0.81	1.82	2.90	0.00	1.14	1.03	0.77	-	

lod (z) scores at recombination fractions (θ)

Based on the above data and the result of multipoint linkage analysis, the most likely order for these markers is:

COL1A2(q21.3-q22.1)-[*Epo*-*D7S15*]-*PON*-[***CF***-*met*-*D7S8*]-*TCRB*(q33)

Our present data also show that both *met* and *D7S8* are closely linked to *CF* but their order with respect to *CF* is unclear from our analysis. In addition, a strong linkage disequilibrium has been detected between *met* and *CF*.

THE MET ONCOGENE LOCUS IS TIGHTLY LINKED TO CYSTIC FIBROSIS: M. Dean[1], M. Park[1], S. Woodward[2], R. White[2], and G.F. Vande Woude[1], [1]LBI-Basic Research Program, NCI-Frederick Cancer Research Facility, Frederick, Maryland; [2]University of Utah, Salt Lake City, Utah.

The met oncogene in MNNG-HOS cells was activated by a DNA rearrangement fusing sequences on chromosome 1 into the met proto-oncogene locus on chromosome 7q21-31. Monosomy in #7, 7q⁻ or interstitial deletions in the vicinity of met occurs in acute non-lymphocytic leukemia (ANLL). Patients with these abnormalities often have a history of either occupational exposure to carcinogens or, in secondary ANLL, to intensive chemotherapy treatment for a primary malignancy. Since met was activated by treating human cells with the direct-acting carcinogen N-methyl-N'-nitronitrosoguanidine (MNNG), we wished to evaluate whether met was involved in ANLL and developed restriction fragment length polymorphism (RFLP) probes in order to distinguish between the two met alleles. Several polymorphisms were detected: met D detects 7.2 and 5.2 kb TaqI fragments in human DNA at frequencies of 0.54 and 0.46, respectively; met H detects 7.5 and 4.0 kb TaqI fragments at frequencies of 0.56 and 0.44, respectively, and 6.5, 2.3 and 1.8 kb MspI fragments at frequencies of 0.07, 0.50 and 0.43, respectively. The TaqI and MspI probe H RFLPs are in disequilibrium. We tested met polymorphisms for linkage to disease in cancer families and families with genetic diseases and discovered that met is tightly linked to the locus for the recessive genetic disease cystic fibrosis (CF). At the very least, this maps the CF locus in man to chromosome 7q21-31. By genetic analysis met is within 1000-2000 kb of the CF gene and is presently the closest gene mapping to this locus. A 9.0 kb RNA transcribed from the met proto-oncogene locus has been detected in all fibroblast and epithelial cell lines studied. This pattern of expression is not inconsistent with that of a putative CF gene, considering the potential involvement of epithelial cells in this exocrine disorder. These observations, and the relationship of met to the tyrosine kinase gene family, of which some members appear to influence salt equilibrium following mitogenic stimulation, requires that met be rigorously included while in search for other candidates for the CF gene upstream and downstream from the met locus.

MOLECULAR CHARACTERISATION OF X-LINKED DISEASES; K.E. Davies[1], S. Forrest[1], S. Kenwrick[1], M. Patterson[1], K. Paulsen[2], T. Smith[1], L. Wilson[1], I. Lavenir[1] and S. Ball[1], [1]Nuffield Department of Clinical Medicine, Oxford, England; [2]Institut fur Humangenetik und Anthropologie, Freiburg, Germany.

Over the last few years a complete genetic map of the human X chromosome has been constructed. This has enabled the localisation of Duchenne and Becker muscular dystrophies to the same region of Xp21, the localisation of X-linked phosphataemic rickets and the development of markers for X-linked mental retardation. However, although this research has resulted in some improvement in the accuracy of diagnosis, a full understanding of the molecular basis of these disorders remains elusive.

Genetic studies of the DMD locus suggest that it lies in a region where recombination occurs frequently. Sequences that lie physically close to the mutation recombine with the DMD and BMD loci in approximately 5% of meioses. This suggests a large gene, several genes, or a hot spot for inversions and rearrangements which alter the gene order. In view of the high complexity of the region, we are studying the molecular arrangement of sequences within Xp21 in different patients using pulsed field gel electrophoresis.

The fragile X mutation also lies in a region demonstrating a high frequency of recombination. This disorder is unusual since transmission can occur through normal males. We have proposed the existence of a premutation to explain the segregation patterns observed. We are testing this hypothesis by the analysis of the region for chromosomal rearrangements using pulsed field electrophoresis.

ANALYSIS OF AN X-AUTOSOME TRANSLOCATION RESPONSIBLE FOR X-LINKED MUSCULAR DYSTROPHY: R.G. Worton, P.N. Ray, S. Bodrug, and M.W. Thompson, Department of Genetics and Research Institute, Hospital for Sick Children, and Dept. of Medical Genetics, University of Toronto, Toronto, Canada.

Duchenne muscular dystrophy (DMD) is rare in females and is almost always associated with a translocation that disrupts band Xp21, the site of the DMD locus. The normal gene on the other X chromosome is silenced in most somatic cells by X inactivation thereby allowing expression of the disease.

One of the translocation cases is a Belgian girl with a mild form of DMD resulting from the translocation t(X;21)(p21;p12). The translocation exchange point in chromosome 21 is within a block of genes encoding 18S and 28S ribosomal RNA.

Using ribosomal RNA gene probes we have recently isolated a phage clone, XJ-1, carrying a DNA segment that spans the junction on the derivative X chromosome. Chromosome walking from this clone has yielded 40 kb of the normal X chromosome spanning the site of the translocation junction. The restriction maps of the two translocation derived chromosomes are co-linear with the maps of the normal X chromosome and the normal 21 suggesting little or no deletion or addition of sequences at the site of the translocation. Unique sequence probes from the 40 kb region detect deletions in a small proportion of male DMD patients and reveal three restriction fragment length polymorphisms that segregate with the DMD gene in families.

The latter analysis has revealed two recombinations between XJ and the DMD gene in 30 meiotic events examined, one of the cross-overs occurring within the 40 kb region. This suggests a high frequency of crossing-over in the region of the DMD gene.

The fact that the XJ region is non-overlapping with the pERT 87 region* suggests that the locus whose disruption may lead to DMD is several hundred kb in size and may include a very large gene or a gene cluster.

* described by Kunkel and his colleagues

MOLECULAR GENETICS OF DUCHENNE MUSCULAR DYSTROPHY

L. Kunkel, A. Monaco, C. Colletti and C. Bertelson.
Division of Genetics, Children's Hospital, Boston, MA

The locus which is disrupted in patients with Duchenne Muscular Dystrophy (DMD) has been localized to the Xp21 region of the human X chromosome. Eight cloned fragments of DNA were isolated which were absent from the DNA isolated from a patient with a rare cytological deletion in Xp21 and three X-linked diseases, including DMD. With the assumption that a certain proportion of patients with solely the phenotype of DMD would be structural alterations at or nearby the locus responsible for the disease, the DNA isolated from patients was tested for deletion or rearrangement of one or more of the eight cloned fragments. One cloned segment, pERT87 (DXS164), was found to be completely absent from the DNA isolated from 6.5% of patients with the phenotype of DMD. Over 230kb of DNA surrounding the original 200bp pERT87 cloned segment was obtained by 9 consecutive bidirectional chromosome "walks". Six unique sequence subclones spaced over 120kb have been determined to identify 14 different restriction fragment length polymorphisms. Three of these subclones have been made available to more than 100 investigators throughout the world for genetic and physical analysis of DNA isolated from DMD patients. The DNA samples isolated from 57 different patients who were shown to be deletions were sent to Boston and the extent of the deletions mapped within the 230kb of the DXS164 locus. Nearly one half (24) of the deletions had at least one break within the DXS164 locus, the remainder were deleted for the entire 230kb of DNA. Three regions were indicated as potentially important by deletion breakpoint analysis and these have failed to identify a clear transcript when utilized as hybridization probes against Northern blots prepared from a variety of human RNA samples. Two of the three regions though are extensively conserved between different mammalian species. Both conserved regions have been narrowed to less than 500bp and the nucleotide sequence of each is currently being determined. The equivalent segments have been isolated from the mouse and each maps to the mouse X chromosome. Sequence comparision between the mouse and human segments should indicate whether these regions are potential portions of the DMD gene.

CARRIER DETECTION AND GENE ANALYSIS OF INHERITED MUSCULAR DYSTROPHIES: P.L. Pearson, State University of Leiden, Genetics Department, Leiden, The Netherlands

Inherited muscular dystrophies represent a wide range of diseases in which the primary genetic defect(s) has not yet been defined, despite intense biochemical investigation over many years. However, recent developments in human gene mapping methods have permitted localisation of the mutant genes concerned to specific chromosome regions for the two most commonly occuring disorders, namely, the X-linked recessive Duchenne muscular dystrophy and the autosomal dominant Dystrophia myotonica. By the use of genetic markers, restriction fragment length polymorphisms (RFLP's), present on the same chromosome region as the disease gene itself, it has been possible to trace the inheritence of the disease within families. It is important to realise that the presence of the disease gene only can be predicted in individuals by its association to the RFLP's as long as sufficient information over the association is present within the particular pedigrees under consideration. Our own group has been particularly interested in applying these methods to Duchenne muscular dystrophy and have used a large number of DNA markers for carrier detection and prenatal diagnosis. In many circumstances it is possible to determine the disease status of an individual with a reliability exceeding 99%. The broad conclusions of the Leiden carrier and prenatal detection programs will be presented. Of particular interest is the observation that carrier determinations based on DNA analysis have eliminated the necessity for carrying out prenatal diagnosis in about 50% of cases. Unfortunately, in some cases the prenatal chorion biopsy was carried out at a stage in which the carrier investigations on the mother were not ready and in retrospect it turned out that the prenatal investigation was unnecessary. This emphasises the necessity for completing family investigations before the 10th week of pregnancy, a fact which will require a better communication and appreciation of the problems involved not only by the family members but also clinicians.

We have noted that a significant proportion of sporadic cases (new mutations) arise in association with meiotic exchange in the same chromosome region and poses the question of whether mutation in Duchenne may be primarily caused by unequal meiotic exchange causing duplication/deletion abormmalities. Evidence will be presented supporting this concept.

Approximately 10% of Duchenne patients have demonstrable deletions with particular DNA markers. The deletions appear to overlap and to vary in size between a few to several hundred kb. Investigations are taking place in our own and other laboratories to make a large scale restriction fragment map of the Duchenne gene region using pulse field gradient techniques. It seems that the chromosome region covered by the gene is very large since meiotic recombinations have been recorded between particular DNA marker sites within the gene and the presumptive mutation site.

Of paramount importance will be the identification of a gene transcript, because only then can researchers begin to unravel the nature of the basic genetic defect itself, a knowledge which conceivably could lead to realistic curative therapies.

MOLECULAR GENETICS OF HUNTINGTON'S DISEASE

J.F. Gusella, T.C. Gilliam, R.E. Tanzi, M.E. MacDonald, S. Cheng, M. Wallace (2), J. Haines (2), P.M. Conneally (2) and N.S. Wexler (3), Neurogenetics Laboratory, Massachusetts General Hospital and Department of Genetics, Harvard University, Boston, Massachusetts; (2) Department of Medical Genetics, Indiana University, Indianapolis, Indiana; (3) Department of Neurology, Columbia Univeristy, New York, New York.

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder with mid-life onset characterized by progressive chorea, cognitive deficits and psychiatric disturbance. Symptoms result from premature loss of neurons that is most marked in the basal ganglia, but the nature of the primary defect is not known. We are attempting to clone and characterize the HD gene based on its on its map location in the human genome.

HD is genetically linked to restriction fragment length polymorphism (RFLP) at the D4S10 locus detected by anonymous DNA probes G8 and R7 from the terminal band of the short arm of chromosome 4 (band 4p16). We have typed a total of 12 large HD families that all showed linkage to the D4S10 marker with no evidence of non-allelic heterogeneity. The two loci are separated by approximately 4-5% recombination with the HD gene being distal to the DNA marker. Detection of a recombination event occurring between individual RFLP sites of the marker has provided the orientation of the D4S10 locus relative to the centromere and the HD gene.

In one family, homozygotes for the HD gene have been identified. Neurological and psychological assessments of these individuals indicate that they do not differ in phenotype from HD heterozygotes, and that HD is therefore a "true" dominant defect. This suggests that the HD mutation results in a "gain of function".

Attempts to identify a second marker flanking the defect have thus far been frustrated by the terminal position of the disease gene on the chromosome. Such a marker is required for more accurate presymptomatic diagnosis, and for efficient application of physical and genetic mapping approaches which should ultimately allow identification of the HD defect.

CONTRIBUTIONS OF DNA STUDIES TO THE PROBLEM OF THE HISTORY OF HUMAN ETHNIC GROUPS

L.L. Cavalli-Sforza, Dept. of Genetics, Stanford University, Stanford, California.

The differentiation of human ethnic groups is sufficiently recent that there has been no time for fixation of different alleles in different groups. Moreover, exchanges between groups have almost always been sufficiently frequent that possible fixations have not occurred. One therefore studies differences in frequency of the various alleles or haplotypes in polymorphic systems.

The introduction of restriction enzyme analysis has enormously increased the number of polymorphisms that are available for study. Mitochondrial DNA has been analyzed by two groups. There are still very few data on polymorphisms for nuclear DNA in different ethnic groups, and available information will be presented. There is some discrepancy between results obtained with classical systems like blood groups, HLA, proteins and enzymes, mtDNA and nuclear DNA. The resolution of these discrepancies will most probably come from DNA segments that are not exposed to natural selection, e.g. pseudogenes, some repetitive DNA, and others. In fact, parts of the genome exposed to natural selection may be more influenced by different environmental factors than by evolutionary history.

Today many unique populations are disappearing at a fast rate. Modern techniques offer, among other things, the possibility of storing their DNA for future analysis. I believe this is an urgent task.

FOSSIL EVIDENCE ON HUMAN ORIGINS AND DISPERSAL:

P. Andrews, Department of Paleontology, British Museum (Natural History), Cromwell Road, London SW7 5BD, U.K.

With increasingly sophisticated biochemical methods, the sequence and timing of evolutionary events are becoming more precisely documented. With a similar increase in the methods of analysis and availability of fossils for study, the timing and place of evolutionary events can be inferred with an increasing degree of confidence. Both approaches support Darwin's suggestion that humans originated in Africa, and a series of events can now be recognized in the fossil record suggesting successive African origins for several hominoid lineages followed by dispersals into Europe and Asia. The order and dates for these branching events will be described here.

The Hominoidea are clearly an African group, being restricted to this continent for the first 8-10 Myr of their known history. They reached their highest levels of species diversity during the first half of the Miocene period, and the subsequent decline in diversity has produced the restricted populations of great apes living today. The emerging human species must therefore be seen in the context of a group well on the way to extinction. The time and place of gibbon divergence from the great apes was earlier than 17 Myr in Africa, and the evidence for the divergence of the orang utan from the African apes and humans suggests a date of earlier than 13 Myr. The order and time of branching within the African ape and human clade is not yet clear, but fossil evidence for the clade places it in Africa 7-9 Myr ago. The human fossil record is restricted to Africa from 1-5 Myr, but since this time, successive radiations gave rise firstly to Homo erectus in Asia at about 1 Myr from an African radiation dating back to more than 1.6 Myr; secondly to archaic Homo (perhaps as many as three species) in Africa and Europe at about 400 Kyr extending to China by 300 Kyr; and finally to modern Homo sapiens in Europe and Asia at about 40 to 50 Kyr from an African radiation dating back to over 100 Kyr.

SEVEN UNIDENTIFIED READING FRAMES OF HUMAN MITOCHONDRIAL DNA ENCODE SUBUNITS OF THE RESPIRATORY CHAIN NADH DEHYDROGENASE

G. Attardi, A. Chomyn and P. Mariottini
Division of Biology, California Institute of Technology, Pasadena, California 91125

More than half of the protein coding capacity of the human mitochondrial genome resides in seven reading frames (originally designated URFs) which have no homology to any of the genes of yeast mitochondrial DNA, although several of them occur in the mitochondrial genome of other lower eukaryotic cells. By the use of antibodies against synthetic peptides predicted from the DNA sequence and by analysis of trypsin fingerprints, it has been shown that all seven URFs are expressed in HeLa cells, and the polypeptide corresponding to each URF has been identified.

Several of the HeLa cell URF products are coprecipitated from a Triton X-100 mitochondrial lysate by individual URF-specific antibodies, an observation which points to an association of these products in some kind of complex. The nature of this complex has been revealed by immunoprecipitation experiments utilizing antibodies against highly purified rotenone-sensitive NADH dehydrogenase from bovine heart mitochondria. In fact, all seven URFs are specifically precipitated by these antibodies from a Triton X-100 or a deoxycholate mitochondrial lysate. Furthermore, antibodies raised against a single subunit of the bovine NADH dehydrogenase, the cytoplasmically synthesized, 49 kilodalton iron-sulfur protein, precipitate a complex containing all seven URF products.

In a different approach to the same problem, enzyme fractionation studies have shown that the URF products copurify with the NADH-Q_1 and NADH-$K_3Fe(CN)_6$ oxidoreductase activities. These observations reinforce the conclusion from the immunochemical experiments that the seven URF products are subunits of the respiratory chain NADH dehydrogenase. Immunoblotting experiments have identified three of these URF products as components of the hydrophobic shell which surrounds the catalytic moiety of the enzyme complex.

MORE ON MITOCHONDRIAL DNA AND HUMAN EVOLUTION: Mark Stoneking[1], Kuldeep Bhatia[2] and Allan C. Wilson[1] [1]Department of Biochemistry, University of California, Berkeley, California USA; [2]Institute of Medical Research, Goroka, Papua, New Guinea.

We have expanded the tree built by R. Cann, M. Stoneking and A. C. Wilson relating mitochondrial DNAs from 147 human beings representing 5 geographic populations: (1) Africa, (2) Australia, (3) New Guinea, (4) East Asia, and (5) Europe including North Africa and the Middle East. The tree now includes mtDNAs from 29 additional New Guineans.

All 176 mtDNAs were mapped at an average of 370 sites per mtDNA by comparing the sizes of fragments produced by restriction enzymes to the sizes expected for one human mtDNA of known sequence. A tree relating the 146 types of mtDNA found confirms that (a) each major race contains many of the main branches in this tree and (b) for the two deepest clusters of lineages specific to New Guinea, the mean sequence divergence through the root of each cluster is 0.13%. This value corresponds, we suggest, to the extent of divergence since the initial colonisation of New Guinea, which is estimated by archaeologists to have begun 30 to 50 thousand years ago. With these two internal points of temporal calibration for the mtDNA tree, we estimate that all known human mtDNA lineages can be traced back to one ancestral woman who lived 140 to 290 thousand years ago, probably in Africa.

MOLECULAR GENETIC INVESTIGATIONS OF ANCIENT HUMAN REMAINS

Svante Pääbo, Department of Cell Research, The Wallenberg Laboratory, University of Uppsala, Box 562, S-751 22 Uppsala, Sweden.

Ancient human remains are found in various archaeological contexts around the world. One area where a particularly rich material of human soft tissues has been preserved is Egypt. A number of Egyptian mummies, representing both naturally preserved bodies from neolithic time (before approx. 3000 BC) and artificially mummified bodies from Pharaonic times, have been investigated. Superficial and periferal parts of the bodies in a majority of cases contain DNA that sometimes can be cloned in bacteria. However, in all cases the DNA is found to be degraded to an average size of a few hundred base pairs. Preliminary investigations on precolumbian and postcolumbian Peruvian mummies indicate that these findings apply also to South American mummies. Peat bogs represent another context in which human soft tissues can be preserved. 8000-year-old brain tissue, recently found in neutral and highly anaerobic peat in Florida, USA, has been analyzed and found to contain high molecular weight nuclear DNA.

Possible applications of paleomolecular biology to the study of the evolution of man and his parasites will be discussed.

THE RELATIONSHIP OF HUMAN PROTEIN SEQUENCES TO THOSE OF OTHER ORGANISMS: R. F. Doolittle, D. F. Feng, M. S. Johnson and M. McClure, Department of Chemistry, University of California, San Diego, La Jolla, California 92093

The amino acid sequences of more than 200 human proteins are known, many having been inferred from nucleic acid sequences. In many cases the sequences of the same or comparable proteins are known from other species. As such, scaled lineages can be derived that in some instances, especially mainstream enzymes, extend back to prokaryotic ancestors. Multiprotein relationships to various organisms, or groups of organisms, have been tabulated and the amount of sequence conservation from various common ancestors to humans estimated. The human protein sequences can be categorized in diverse ways; we have loosely grouped them into "ancient" and "modern" types. The "ancient" group includes virtually all the enzyme sequences and a few non-enzyme proteins. The "modern" group includes numerous extracellular proteins that appear to have been concocted during the last billion years. Some are modular proteins that have been assembled by exon shuffling. Most of the modules in these cases are compact structures containing several intramolecular disulfide bonds. Occasionally these units have been grafted on to enzymes, which are themselves derived from ancient stock. Examples include several serine proteases found in vertebrate blood plasma. The goal of these comparisons and tabulations is to establish better the molecular roots of human proteins and to chronicle the appearance of modern proteins that have allowed the evolution of modern organisms, including homo sapiens.

CLONING OF CDNA AND GENOMIC DNA FOR HUMAN VON WILLEBRAND FACTOR: J.E. Sadler[1], B.B. Shelton-Inloes[1], J.M. Sorace[1], and Ko Titani[2], [1]Howard Hughes Medical Institute and Department of Medicine, Washington University, St. Louis, Missouri; [2]Department of Biochemistry, Univ. of Washington, Seattle, Washington.

Von Willebrand factor (VWF) is a plasma protein that is synthesized by vascular endothelial cells and megakaryocytes. It is essential for normal platelet adhesion to damaged endothelium, and also for normal factor VIII survival in the circulation. VWF is composed of identical subunits with Mr ~250,000 that form a series of multimers, from dimers to species with Mr >10 million. Von Willebrand disease, the most common inherited bleeding disorder of humans, is characterized by deficiency of VWF activity. Using antibody and cDNA probes, overlapping cDNA clones for VWF have been isolated from λgt11 cDNA libraries constructed from cultured human endothelial cell RNA. These span ~6.5 KB of the VWF mRNA and specify a portion of the leader peptide, all 2050 amino acids of the mature VWF subunit, the 3'-noncoding sequence, and a poly(A) tail. By Northern blotting, the VWF mRNA is ~8.5 KB in length. Some structure-function relationships of VWF can be inferred from the translated amino acid sequence. The protein is synthesized as a precursor that is cleaved after a Lys-Arg dipeptide to generate the mature subunit. The VWF polypeptide contains 5 types of repeated domains (A-E), of which 2 are triplicated (A and B), and 3 are duplicated (C,D, and E). Together, repeated domains comprise over 75 percent of the sequence. The A domains each show apparent homology to a 225 amino acid segment of complement factor B. A binding site for platelet glycoprotein Ib (GPIb) appears to include the first A domain, and a binding site for the GPIIb/IIIa complex of activated platelets is in the second C domain. The cDNA inserts have been used to isolate genomic DNA from a Charon 4A λ-phage library. Eight clones spanning over 70 kilobases of the gene have been characterized that correspond to ~3.3 kilobases of the cDNA sequence. At least 14 exons have been identified by Southern blotting, and 7 exons have been sequenced. These studies form a basis for the investigation of the cause of von Willebrand disease at the level of gene structure.

STRUCTURE AND EVOLUTION OF THE HUMAN PROTEIN C GENE: George L. Long, Lilly Research Labs., Eli Lilly and Co., Indianapolis, Indiana.

Protein C is a representative member of a class of vitamin K-dependent plasma proteins involved in hemostasis. Other members in this group include coagulation factors VII, IX, X and prothrombin. Each of these proteins circulates as an inactive zymogen that, upon limited proteolytic cleavage, is converted to an active serine protease. Protein C serves as a feed-back down-regulator of the coagulation cascade by specifically degrading the protein cofactors VIIIa and Va. The protein C precursor consists of the following domains: leader peptide, "gla" region, two epidermal growth factor segments, and the activation peptide/serine protease. The human gene for protein C has been cloned and characterized [Plutzky, et al. PNAS 83, 546-550 (1986)]. The gene is contained within a 12 Kb segment of DNA and consists of coding sequence for a 461 amino acid precursor, interrupted by seven intervening sequences (introns). Intron positions correspond to presumed protein structure/function domain boundaries. An analysis of the protein C gene structure in forty-five protein C deficient individuals using recombinant DNA probes has revealed no observable alterations, except that in two individuals (siblings) it appears that one entire copy of the gene has been deleted. Comparison of amino acid sequences reveals that protein C and factor IX are homologous. A comparison of the genes for protein C and factor IX shows that all of the introns are in identical positions. However, the base compositions of the two genes (coding and non-coding regions) are remarkably different: ~60% G+C for protein C versus ~40% G+C for factor IX. One possible explanation for this phenomenon is that the factor IX gene (located on the X chromosome) has undergone extensive deoxycytosine methylation and subsequent mutation, resulting in a net C to T (and G to A) transition. This would suggest that the protein C gene may represent a more primitive form of the duplication precursor.

ISOLATION AND EXPRESSION OF cDNAs ENCODING HUMAN FACTOR VII. K. Berkner[1], S. Busby[1], E. Davie[2], C. Hart[1], M. Insley[1], W. Kisiel[3], A. Kumar[1], M. Murray[1], P. O'Hara[1], R. Woodbury[1], F. Hagen[1], [1]ZymoGenetics, Seattle, WA, [2]Dept. of Bioc., Univ. of Wash., Seattle, WA, [3]Dept. of Path., Univ. of New Mexico, Albuquerque, N.M.

Factor VII exists in plasma as a precursor to a serine protease. When converted to the active form by cleavage, it activates factor X and/or factor IX in the presence of tissue factor and calcium, in the extrinsic pathway of blood coagulation. Biological activity of factor VII is dependent upon the γ-carboxylation of approximately 10 glutamic acid residues present in the amino-terminus (the "gla" domain), and this post-translational modification requires vitamin K. We report here the isolation of human factor VII cDNAs from cDNA libraries prepared from HepG2 cells and human liver polyA$^+$ RNA. The identity of these cDNAs has been confirmed by comparison to factor VII amino acid sequence data, and to DNA sequence data derived from factor VII-containing genomic clones. The cDNAs encode a mature factor VII protein of 406 amino acids, which includes a gla domain, two potential EGF domains, and a serine protease domain. DNAs encoding differing leader lengths (38 or 60 amino acids) have been identified in the cDNA population. Factor VII exhibits a high degree of homology with the other vitamin K-dependent plasma proteins.

Full-length factor VII cDNAs with the two different leaders have each been inserted into a mammalian expression vector. These plasmids have been transfected into mammalian cells, in some cases along with a plasmid encoding dihydrofolate reductase. Secretion of recombinant factor VII has been observed in a variety of cell lines (e.g. cos cells, BHK cells and HepG2 cells) in transient analysis, and stably transfected cell lines secreting factor VII have been established. The levels of factor VII expression from the two cDNAs with different leader sequences are approximately the same. The biological activity of the recombinant factor VII has been demonstrated in a one step clotting assay: in preliminary studies, the percentage of biologically active protein is at least 60%. We are currently examining the extent to which factor VII is γ-carboxylated in our expression system.

RECOMBINANT HUMAN FACTOR VIII: R. Adamson, G. Amphlett, M. Bologna, A. Dorner, W.B. Foster, R. Hewick, D. Israel, R. Kamen, R. Kaufman, D. Pittman, B. Schmidli, K. Smith, P. Szklut, J. Toole, J. Wang, L. Wasley, Genetics Institute, Inc., Cambridge, Massachusetts, U.S.A.

The use of gene technology to produce pure, safe and inexpensive Factor VIII has been a major goal of the biotechnology industry. The cloning and characterization of the human Factor VIII gene has now been accomplished. Site-directed mutagenesis has been used to explore fundamental domains of the molecule, with the surprising result that a substantial portion of the protein sequence can be deleted with little effect on procoagulant activity. Although Factor VIII is an extremely large and complex glycoprotein, it has been expressed in a scaleable mammalian cell system. The protein purified from this system closely resembles natural human Factor VIII in structure and in functionality in vitro. Results of in vivo experiments using the hemophilic dog model strongly suggest that the recombinant product will be useful for human therapy.

MOLECULAR BIOLOGY OF HEMOPHILIA

R.M. Lawn, D.J. Capon, J. Gitschier, E.G.D. Tuddenham*, G.A. Vehar, K.L. Wion and W.I. Wood, Genentech, Inc., South San Francisco, CA, *Royal Free Hospital School of Medicine, London

Classic hemophilia (hemophilia A) is caused by a deficiency of factor VIII, a component of the intrinsic clotting pathway. The gene for human factor VIII has been cloned and sequenced, and recombinant factor VIII has been produced by transfected tissue culture cells. The gene sequence provides information about the factor VIII protein and its proteolytic processing, and provides a framework for the understanding of the molecular basis of haemophilia.

The factor VIII gene spans 186,000 nucleotides of the human X chromosome and codes for a protein of M_r ~300,000. The protein contains an internally triplicated domain plus another duplicated domain, and a large intermediate region which is eliminated during activation. The triplicated domain, which comprises most of the activated form of factor VIII, shares significant homology with the plasma protein ceruloplasmin.

Cloned DNA probes have been used to detect factor VIII in hepatocytes and in a number of human tissues, including kidney, spleen, pancreas and lymph nodes. DNA probes also detect restriction fragment length polymorphisms in the factor VIII gene which have been used in antenatal diagnosis and carrier detection of haemophilia A. We have also analyzed over 200 haemophilic DNAs by blot hybridization and have detected a number of different deletion, nonsense and missense mutations in the factor VIII genes of haemophiliacs.

Production of human tissue plasminogen activator (t-PA) using recombinant DNA technology. Gordon A. Vehar, Bruce Keyt., Michael Spellman, William Hancock, Stuart Builder. Genentech, Inc., South San Francisco, CA.

The production of large quantities of recombinant t-PA has allowed charactization of the biochemical and pharmacological properties of this important mediator of fibrinolysis. The fibrin-stimulated activation of plasminogen to plasmin by t-PA results in the proteolysis of fibrin and in turn the dissolution of blood clots as part of the normal vascular healing process. The rate at which this cascade of events occurs is directly related to the circulating t-PA levels. Nucleic acid sequence determination of the t-PA gene and cDNA, as well as characterization of the protein, have led to the proposed structure of a multi-domain, 527 amino acid glycoprotein containing 17 disulfide bonds. By analogy with other known proteins, t-PA consists of a serine protease domain attached to fibrin binding structures. The inability of bacteria to glycosylate proteins and to form disulfides has led to the use of mammalian cell expression systems for large scale production. The resulting material is now available in highly pure form. Analysis of the protein has shown the presence of two forms, designated types I and II, which differ in their carbohydrate content. In addition, the glycosylation at position 117 of t-PA has a high mannose structure; the remaining carbohydrate(s) are of the complex type. Recombinant t-PA has been shown to be equivalent to non-recombinant material when analyzed by a variety of in vitro methods. The availability of large quantities of recombinant t-PA (Activase™) has allowed the evaluation of this material for fibrinolytic activity in clinical trials.

BIOCHEMICAL AND BIOLOGICAL PROPERTIES OF SINGLE CHAIN UROKINASE-TYPE PLASMINOGEN ACTIVATOR : D. Stump, H.R. Lijnen and D. Collen, Center for Thrombosis and Vascular Research, University of Leuven, Belgium.

Single chain urokinase-type plasminogen activator (scu-PA) occurs naturally in man in both plasma and urine. When purified in active form from either human urine or conditioned media from cultured CALU-1 human lung adenocarcinoma cells, it is characterized by plasminogen dependent fibrinolytic activity on fibrin films, migration on sodium dodecyl sulfate-polyacrylamide gel electrophoresis as a single polypeptide chain with apparent M_r 54,000 under reducing conditions, and immunological identity with urinary two-chain urokinase. It is virtually inactive toward low molecular weight chromogenic substrates for urokinase but is converted to amidolytically active urokinase by limited digestion with plasmin. scu-PA directly activates plasminogen with uniquely high affinity (K_m 0.5 uM) but is competitively inhibited in plasma. This inhibition is reversed by the presence of fibrin. It induces fibrinolysis in vitro with a high degree of clot-specificity. In rabbits, dogs, and man, it can cause thrombolysis without systemic fibrinolytic activation, although in some patients with acute myocardial and coronary artery occlusion, its use may be associated with variable degrees of fibrinogen degradation.

The cloning and expression of scu-PA in E. coli (rscu-PA) (Holmes et al, Biotechnology 3: 923, 1985) now has provided an alternative source for this natural fibrin-specific plasminogen activator. rscu-PA possesses similar biochemical properties as scu-PA and is also capable of mediating clot-selective fibrinolysis in rabbits and baboons. Its use has also now been extended to patients with acute myocardial infarction where a 75 percent reperfusion rate was obtained in the first 12 subjects treated with a mean reperfusion time of about 35 minutes (Van de Werf et al, in press), but this was associated with extensive fibrinogen degradation in three patients. The potentially more widely available recombinant scu-PA shows promise as a fibrin-specific thrombolytic agent in patients with thromboembolic disease.

INTERFERON PRODUCTION FROM HUMAN CELL CULTURES

N.B.Finter, G.D.Ball, K.H.Fantes, M.D.Johnston and W.G.Lewis
Wellcome Biotechnology, Beckenham, Kent BR3 3BS U.K.

Bulk human interferon(IFN) can be obtained by two routes: by obtaining expression of a human IFN-a gene inserted into some heterologous host cell; or by growing human cells in mass culture and stimulating their many IFN genes. Wellcome has pioneered the development of this latter route.

After screening 150 lines, we chose the Namalwa line of human B lymphoblasts, derived from a Burkitt lymphoma patient, for development. We have built plants in which these cells have been grown in steel tanks of progressively larger size from 1000 litres in 1977 to 10,000 litres today; the latter provide more than $10^{13.5}$ cells per production batch. Cells are conditioned by incubation with butyrate for 48h and induced to form IFN by treatment with Sendai virus.

The final product, Wellferon, has been shown to contain a mixture of at least 17 distinct IFN-a species at a purity of more than 95%, and with no detectable contamination with DNA (less than 10pg/dose). It can be produced abundantly and relatively cheaply. It may prove to have clinical advantages over any single recombinant IFN because of its broader spectrum of biological activity and the fact that it is naturally formed in and released from human cells. Wellferon has been used in extensive clinical trials in patients with various virus infections or forms of cancer, and has been licensed in the U.K. for the treatment of hairy cell leukaemia.

MOLECULAR BIOLOGY OF THE INTERLEUKIN-2 SYSTEM:

T.Taniguchi,T.Fujita,M.Hatakeyama,G.Yamada,S. Minamoto,H.Shibuya,T.Ohashi,K.Yamanishi,H.Mori, R.R.Hardy,*H.Matsui,*J.Hamuro and ^T.Uchiyama, Institute for Molecular and Cellular Biology, Osaka University, Osaka, *Central Research Laboratories, Ajinomoto,Co.Inc., Yokohama, ^Faculty of Medicine, Kyoto University, Kyoto, Japan

Interleukin-2(IL-2), the major lymphokine for T cells, plays a key role in the growth control of T cell clones by interacting with a specific cell surface receptor(IL-2R). As both IL-2 and IL-2R genes are induced mainly by T3/Ti receptor triggering, antigen-specific T cell growth appears to occur via a process of signal transduction wherein the specific antigen/MHC-T3/Ti interaction is converted into a growth pathway mediated by the IL-2 system. IL-2 is also known to function in the activation and differenciation of a variety of lymphoid cells such as NK and LAK(lymphokine activated killer). Hence, there has been increasing interests on this lymphokine for the use of cancer immunotherapy. In order to elucidate the molecular mechanism regulating the IL-2 system and to provide IL-2 in unlimited quantity, we cloned and expressed cDNAs for human and murine IL-2. Human IL-2 has been produced in E.coli and the recombinant IL-2 is being used for cancer therapy in many institutions. We next analysed chromosomal genes for Il-2. Expression studies of the human IL-2 gene revealed the presence of unique DNA element(s) that functions as T cell-specific regulatory enhancer. We also cloned the gene for Tac antigen, a putative IL-2R. We carried out a series of expression studies with the cloned cDNA and demonstrated that the cDNA encodes both high(functional) and low(non-functional) IL-2R in T-lymphoid but not in fibroblast cells. We also found that IL-2 mediates reversed growth signal transduction in the transformed T lymphoid cells by altering the expression of various cellular onc genes. The role of intracytoplasmic region of the IL-2R in the signal transduction by IL-2 has been investigated by generating mutations in the IL-2R cDNA.

LYMPHOKINES AND MONOKINES IN ANTICANCER THERAPY:

W. Fiers, P. Brouckaert, R. Devos, L. Fransen[1], G. Leroux-Roels, E. Remaut, P. Suffys, J. Tavernier[1], J. Van der Heyden[1] and F. Van Roy
Laboratory of Molecular Biology, State University; [1]Biogent (subsidiary of Biogen S.A), Ghent, Belgium

Various interferons have been cloned and expressed at high levels in bacteria or in yeast. Although some of these interferons appear promising for treatment of certain forms of malignancies, they are usually not active on the major types of human cancer. Also interleukin 2 is now available in abundant supply and has been used with some success in anticancer trials. However, what seems to be the most specific anticancer drug is Tumour Necrosis Factor. Native TNF is a trimer composed of 17,000 dalton subunits. It is coded for by a single gene in man and mouse. TNF is almost not species-specific. It is specifically toxic for a fairly large number of malignant cell lines, especially when the latter are also treated with interferon gamma. Also in vivo, long-term curing of mice carrying syngeneic B16BL6 melanoma tumours has been achieved. Furthermore, complete responses were obtained in nude mice carrying human tumours. TNF receptors are present on susceptible cells but also on many types of normal cells. Remarkably, TNF acts mitogenic on several types of normal cells, rather than cytotoxic or cytostatic. The action on a number of cells is closely similar to that of IL1, although the latter interacts with a different receptor. The action on some types of normal cells is undoubtedly responsible for the pronounced toxicity of TNF; a better understanding of the latter phenomenon is obviously required. Therefore, mutants were constructed by site-specific mutagenesis in order to probe the active centre of the molecule. The other important question is why some malignant cells respond to TNF (in combination with interferon gamma) while others do not.

TUMOR NECROSIS FACTOR: GENE STRUCTURE, EXPRESSION AND BIOLOGICAL PROPERTIES, D. Goeddel, B. Aggarwal, M.A. Palladino, D. Pennica, B. Sugarman, H.M. Shepard and G.H.W. Wong; Genentech, Inc., South San Francisco, California.

Tumor necrosis factor (TNF) is a polypeptide of 157 amino acids that is synthesized and secreted by activated macrophages and by mitogen-stimulated promyelocytic cell lines. TNF is cytotoxic toward a variety of tumor cells *in vitro* and *in vivo* while acting as a growth factor for normal fibroblasts. The anti-tumor activity of TNF is enhanced by the presence of IFN-γ, a lymphokine produced by the immune system under similar physiological conditions.

Natural human TNF was purified from mitogen-treated HL-60 promyelocytic leukemia cells. The cDNA for human TNF was isolated, sequenced and expressed in *E. coli*. The TNF gene is closely linked on chromosome 6 to the gene for the structurally-related cytotoxin known as lymphotoxin (LT). Studies on the regulation of these genes indicate they are expressed in a tissue-specific manner but that the expression of both genes is induced similarly. Receptor binding experiments demonstrate that TNF and LT recognize the same cell surface receptor and that this receptor is up-regulated by pre-treating cells with IFN-γ. Additional properties and activities of recombinant TNF will be discussed.

TANDEM ARRANGEMENT OF TUMOR NECROSIS FACTOR AND LYMPHOTOXIN GENES IN HUMAN GENOME:

S.A. Nedospasov[1], A.N. Shakhov[1], R.L. Turetskaya[1], V.A. Mett[1], G.P. Georgiev[1], V.G. Korobko[2], V.N. Dobrynin[2], S.A. Filippov[2], N.S. Bystrov[2], E.F. Boldyreva[2], S.A. Chuvpilo[2], A.M. Chumakov[2] and Yu.A.Ovchinnikov[2], [1]Institute of Molecular Biology, USSR Academy of Sciences, Moscow; [2]Shemyakin Institute of Bioorganic Chemistry, USSR Academy of Sciences, Moscow, USSR.

Tumor necrosis factor (TNF-alpha) and lymphotoxin (TNF-beta) genes have been isolated from human genomic libraries using oligonucleotide probes. We will present data on the structure and expression of these related genes and on their tandem arrangement in the same orientation within a short segment of human genome.

CACHECTIN: A MACROPHAGE PROTEIN THAT INDUCES A CATABOLIC STATE IN INFECTED ANIMALS. A. Cerami, Laboratory of Medical Biochemistry, The Rockefeller University, New York, NY 10021.

Mammals infected with parasitic, bacterial, or viral organisms or bearing tumors characteristically display a catabolic state and weight loss which can if not resolved, advance to cachexia (or wasting), shock and death. Although commonly observed in many parasitic diseases, the mechanism of this phenomenon has not been understood. We have recently identified and isolated a macrophage protein, cachectin, as the molecule that may be responsible for cachexia and shock. Cachectin is produced by macrophages in response to endotoxin or a number of other bacterial or protozoal products. The released cachectin then acts as a hormone where it binds to specific high affinity receptors and elicits a number of biological responses. In the adipocyte, for example, several anabolic enzymes, eg., lipoprotein lipase, are selectively suppressed because of a selective inhibition of mRNA production. One of the more intriguing aspects of cachectin is its pivotal role in the pathogenesis of endotoxin-induced shock. Cachectin causes fever, anorexia and can induce a lethal shock state in experimental animals. During the course of the chemical characterization of cachectin it was shown that cachectin was identical to tumor necrosis factor (TNF), a macrophage protein that kills tumor cells. This finding has served to emphasize the extensive range of effects that are associated with this protein. Cachectin has many properties in common with IL-1, however, it binds to a different receptor and lacks structural homology. Presumably, low levels of cachectin are helpful to the host in its battle to remove invasive pathogens; although, prolonged or extensive production of cachectin can lead to severe wasting and shock. These findings have added a new dimension to the biological properties of cachectin, its production, and its role in cachexia and shock.

INTERLEUKIN-1 MOLECULAR BIOLOGY:

Peter T. Lomedico, Patricia L. Kilian, Ueli Gubler, Alvin S. Stern and Richard Chizzonite; Hoffmann-La Roche Inc., Roche Research Center, Nutley, N.J., 07110

Interleukin-1 (IL-1), a polypeptide hormone (Mr, ≈17,000) synthesized and secreted by activated macrophages, acts on many different cells to modulate a variety of biological activities including lymphocyte activation, fever, liver cell function, connective tissue cell activation and proliferation, etc. Through these diverse activities, IL-1 helps to stimulate immune and inflammatory responses in the body's defense against infection and other forms of trauma. IL-1 may also contribute to the pathology observed in certain chronic inflammatory conditions.

The genes for murine, rabbit and human IL-1 have been cloned and expressed in E. coli. This work, together with protein sequencing and biosynthetic studies, has demonstrated that IL-1 is initially synthesized as a precursor protein of ≈270 amino acids, which is processed to liberate from the carboxy-terminus the biologically-active hormone (≈156 amino acids). The gene cloning experiments have proven that a family of evolutionary related genes code for at least two distinct IL-1 proteins. Mammalian IL-1 proteins (IL-1α=pI5 protein; IL-1 β=pI7 protein) are ≈60% homologous in amino acid sequence within a subfamily, but only ≈25% homologous between subfamilies.

Recent experiments have demonstrated that IL-1 is not a macrophage-specific gene product: we have been able to measure IL-1 mRNAs in other human cells (eg. epithelial cells, keratinocytes, endothelial cells, fibroblasts, B-cells, etc.) and have found that the ratio of IL-1α mRNA vs. IL-1 β mRNA varies in different cells.

Recombinant mouse and human IL-1α have been radiolabeled with ^{125}I in a biologically active form. High affinity binding sites for IL-1 have been identified on murine EL-4 thymoma cells as well as other mouse and human cell lines. Importantly, it appears that both human IL-1α and β recognize the same receptor site. Characterization of the IL-1 receptor is continuing.

DEVELOPMENT OF MULLERIAN INHIBITING SUBSTANCE AS A POTENTIAL ANTI-CANCER DRUG: R. L. Cate[1] and P. K. Donahoe[2]
[1]Biogen Research Corporation, Cambridge, Massachusetts;
[2]Department of Surgery, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts.

Mullerian Inhibiting Substance (MIS) is a testicular glycoprotein expressed early in the sexual development of the male embryo. MIS causes regression of the Mullerian duct which, in the normal female embryo, develops into the uterus, vagina and Fallopian tubes. In addition to this developmental role, MIS also can inhibit the growth of tumors that derive from tissues related to the Mullerian duct. In order to understand the mechanism by which MIS acts as a natural regressor during embryonic development and to further test the efficacy of MIS as an anti-cancer agent, we have isolated the bovine and human genes for MIS. The mRNA sequence of bovine MIS, determined from an analysis of cDNA and genomic clones, encodes a protein of 575 amino acids which contains a 24 amino acid leader peptide. The human gene has five exons that encode a protein of 560 amino acids. A comparison of the bovine and human MIS proteins reveals a highly conserved C-terminal domain that shows marked homology with human transforming growth factor-beta and the beta chain of porcine inhibin. Expression of the human gene in CHO cells produces an active protein that causes the regression of the rat Mullerian duct *in vitro*. The level of expression is being increased in order to produce large quantities of MIS for postimplantation anti-cancer studies.

USE OF OLIGONUCLEOTIDE PROBES FOR DIAGNOSTIC PURPOSES:

R. Bruce Wallace, Department of Molecular Genetics, Beckman Research Institute of the City of Hope, Duarte, CA.

When radiolabeled and used as hybridization probes, synthetic oligonucleotides display essentially absolute hybridization specificity. That is, every nucleotide must form a Watson-Crick base pair in order that the probe forms a stable duplex. All of the non-Watson-Crick base pairs have a destabilizing effect. Thus, it is possible to choose stringent conditions of hybridization such that, while a perfectly matched duplex between an oligonucleotide and complementary DNA will form, duplexes mismatched at one or more position will not.

Since mutations in a single base in the DNA sequence of a gene can and do result in genetic diseases, the hybridization of oligonucleotides to the region of DNA containing these base changes provides a means of detecting these single base differences. An oligonucleotide complementary to a normal gene will not hybridize to the mutant allele while an oligonucletide complementary to the mutant gene will not hybridize to the normal allele. The pattern of hybridization obtained using two such oligonucleotide probes gives an unambiguous genotype of an individual's DNA. It has been possible to use such an assay for the diagnosis of sickle cell anemia, α_1-antitrypsin deficiency, β^0-thalassemia, Hb-C, Hb-E and β^+-thalassemia. This approach appears to be generally applicable to any genetic disease that is due to a known base change.

In order to improve the sensitivity of oligonucleotide hybridization, we have developed a plasmid vector which allows the specific transcription of a radiolabeled oligoribonucleotide with a defined sequence. The vector is based on an *E. coli* concensus sequence promoter, with a cloning site near the initiation of transcription. Oligoribonucleotide probes can be produced by run-off transcription *in vitro* using *E. coli* RNA polymerase.

SPECIFIC ENZYMATIC AMPLIFICATION OF DNA IN VITRO

Kary Mullis, Fred Faloona, Stephen Scharf, Randall Saiki, Glenn Horn and Henry Erlich, Department of Human Genetics, CETUS Corporation, Emeryville, California

We have devised a method whereby a specific nucleic acid sequence can be enzymatically amplified. The method is based on the DNA polymerase catalyzed extension of two oligonucleotide primers which anneal to the ends of the target sequence, and on different strands, such that the extension product from one of the primers will serve as a template for the other. Repeated cycles of denaturation, primer annealing and extension result in an exponential accumulation of the target sequence. It is not necessary that the sequence to be amplified be present initially in a pure form; it can be a minor fraction of a complex mixture, such as a segment of a single copy gene in whole human DNA. The sequence can be present initially as a discrete molecule or it can be part of a larger molecule. In either case the product of the reaction will be a discrete dsDNA molecule with termini corresponding to the 5'-ends of the oligonucleotides employed.

This procedure has been used to amplify a specific betaglobin segment in genomic DNA approximately one million fold, enabling the direct determination of betaglobin alleles from ethidium stained gels. Using specific hybridization probes and amplified DNA in a dot-blot format we have been able to determine betaglobin genotypes with as little as 33 ng of DNA.

With primers designed to amplify specific sequences and also to introduce terminal restriction sites into the amplified fragments, we have been able to rapidly construct clones for sequence analysis derived from specific betaglobin, HLA-DQ-α, HLA-DQ-β and N-ras genomic segments.

The amplification procedure, which we have called the Polymerase Chain Reaction (PCR), requires less than five minutes per cycle and lends itself easily to automation. We have accomplished this with a microprocessor controlled liquid handling device.

This procedure promises to be very useful for molecular analysis of the human genome as well as for DNA based diagnosis of infectious diseases. It is also proving to be a convenient tool in a variety of molecular biological techniques.

DETECTION AND LOCALIZATION OF SINGLE BASE MUTATIONS IN TOTAL GENOMIC DNA: Richard M. Myers[1], Leonard S. Lerman[2], and Tom Maniatis[3]; [1]Department of Physiology, University of California, San Francisco; [2]Genetics Institute, Cambridge, MA; [3]Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA.

We have been developing methods for the detection and localization of single base mutations in total genomic DNA, with the aim of applying these methods towards the diagnosis of human genetic diseases. In the first method, single base mutations are detected by denaturing gradient gel electrophoresis. In the second method, single base mutations are detected and localized by cleavage of mismatches in RNA:DNA hybrids by ribonuclease. Each of these procedures will detect approximately fifty percent of all possible single base substitutions in a cloned or genomic DNA fragment. Although there is some overlap in the mutations detected by the two procedures, we estimate that their combined use will detect seventy five percent of all mutations in a fragment. We will discuss improvements in the use of both procedures and their applications towards human genetic diseases.

HELIX STABILITY AND GENETIC ANALYSIS: L. Lerman[1], R. Myers[2], N. Lumelsky[3], and T. Maniatis[4], [1]Genetics Institute, Cambridge, MA; [2]Dept. of Physiology, Univ. of California, San Francisco; [3]Dept. of Biochem. and Molecular Biology, Harvard University, Cambridge, MA; [4]Dept. of Biology, Yale University, New Haven, CT.

The intimate relation between base sequence of the double helix and the stability of the helix toward heat characterizes the sequence and provides a basis for physical examination of sequence variations. The transition from full helicity to a partly helical, partly disordered structure by heat or solvents severely reduces the electrophoretic mobility of DNA in gels and can establish a sensitive correspondence between gel position, thermal stability, and sequence. Sensitivity is further enhanced if sequence differences are presented as an error in perfect pairing in a heteroduplex composed of strands from each of the two DNA molecules to be compared.

We have surveyed the detection of single base changes, including insertions or deletions, base substitutions, and postsynthetic modifications, by means of denaturing gradient gel electrophoresis. These procedures have shown the presence of polymorphisms and of substitutions identified with genetic disease in human genomic DNA.

The pattern of thermal stability sets limits to the detection of base changes by simple heteroduplexes with a wild-type probe. Theoretical simulation permits analysis of the distribution of sequences accessible or inaccessible to this type of scrutiny. Calculations applied to the extended human -globin region indicate the extent of response to mutations and polymorphisms and show where the response can be expanded by use of a clamp sequence. The patterns are also intrinsically interesting in displaying apparent relations between helix stability and biological function.

MOLECULAR BASIS OF PHENYLKETONURIA AND POTENTIAL SOMATIC GENE THERAPY

S.L.C. Woo, Howard Hughes Medical Institute, Department of Cell Biology and Institute of Molecular Genetics, Baylor College of Medicine, Houston, Texas, 77030.

Phenylketonuria (PKU) is a recessive human genetic disorder caused by a deficiency of the hepatic enzyme phenylalanine hydroxylase (PAH). The disorder has a prevalence of about 1 in 10,000 births among Caucasians, and predisposes affected individuals to development of severe mental retardation. The human chromosomal PAH gene contains 13 exons and is 90 Kb in length to code for a mature mRNA of 2.4 Kb. Ten polymorphic restriction sites have been identified in the human PAH locus, which have been utilized for prenatal diagnosis of the genetic disorder. Detailed RFLP haplotype analysis of the Northern European population has demonstrated that over 90% of PKU genes are represented by four distinct haplotypes. All four mutation genes have been isolated by cosmid cloning from patients who are haplotype homozygotes. Sequencing analysis followed by expression studies will define the molecular basis of the hereditary disorder and determine whether each of the RFLP haplotype is associated with a specific mutation due to linkage disequilibrium. These studies may lead to development of specific probes for detection of the corresponding mutant alleles in the population to allow carrier detection without prior family PKU history.

We have also investigated the prospect of somatic gene therapy for PKU by retroviral mediated gene transfer. The full-length human PAH cDNA has been inserted into a cryptic retroviral vector which has been transfected into cells providing all the viral functions in trans. Defective recombinant retroviruses containing the human PAH cDNA were detected in culture media which were used to infect fibroblasts and a mouse hepatoma cell line that is PAH deficient. Infected cells express PAH immunoreactive protein and enzymatic activity characteristic of human PAH. In addition, the infected mouse hepatoma cell acquired the capability to grow in tyrosine-free medium, indicating the reconstitution of the entire phenylalanine hydroxylation system. Defective recombinant retroviruses provide an efficient means for gene transfer and will enable experiments in which human PAH be transduced into whole laboratory animals as a model for somatic gene therapy of PKU.

ANALYSIS OF NORMAL AND ABNORMAL HUMAN GENOMES WITH FLOW SORTED CHROMOSOMES: R. Lebo, F. Chehab, Y.F. Lau and Y.W. Kan, Howard Hughes Medical Institute and Department of Medicine, University of California, San Francisco, California.

Flow analysis of human chromosomes is a powerful new approach to study the human genome. Human chromosomes can be resolved almost completely as a result of improvements in metaphase chromosome suspensions, new fluorescent dye combinations, and advances in sorter optical design. We have applied this technique to map genes, construct individual recombinant chromosomal DNA libraries, and detect abnormal chromosome constitutions.

The dual laser sorter can separate all the human chromosomes except 10 and 11. Individual chromosomes are sorted directly onto filter paper, and hybridized to gene probe. Chromosomes from cell lines with the appropriate translocations are used to subchromosomally localize genes. With these methods we have located 37 recently cloned genes. When homologous gene sequences such as ferritin-L and Von Willebrand factor are located on different chromosomes, Southern blot analysis of sorted chromosomal DNA directly assigns the genes to specific chromosomes.

A chromosome-specific library facilitates the identification of many chromosome-specific DNA fragments that can be tested to locate a polymorphism close to a mapped disease locus. Once a closely linked polymorphic locus has been identified, the library will also facilitate walking toward a disease locus. Currently we have constructed an EMBL-4 phage library of chromosome 1 to locate a polymorphic DNA fragment close to the Charcot-Marie-Tooth Disease locus. The fragments are being sublocalized by spot-blot analysis and those in the vicinity of the disease locus tested for polymorphism. Subsequent pedigree analysis will establish which fragment can be used to test the disease status of an at-risk fetus.

Flow analysis can detect altered chromosome frequencies and abnormal chromosomes with a 5%-10% change in total DNA content. We have examined 11 clinical samples with deletions, insertions, translocations, and aneuploidy and found excellent agreement with results obtained by Giemsa-banded karyotypes. These results indicate that flow cytogenetics is potentially useful for patient screening.

DNA-BASED DETECTION OF CHROMOSOME DELETION AND AMPLIFICATION: DIAGNOSTIC AND MECHANISTIC SIGNIFICANCE:

S. Latt[1,2], M. Lalande[1,4], T. Donlon[1], A. Wyman[5], E. Rose[1], Y. Shiloh[1], B. Korf[1], K. Sakai[1], J. Kang[1], H. Stroh[1], P. Harris[1], G. Bruns[1], R. Wharton[3], and L. Kaplan[1], [1]Division of Genetics and [3]Developmental Evaluation Clinic, Children's Hospital and Departments of [1]Pediatrics and [2]Genetics, Harvard Medical School; [4]National Research Council, Canada (current address); [5]Department of Biology, M.I.T.

Recombinant DNA libraries constructed from highly purified flow-sorted metaphase chromosomes provide convenient starting points for in-depth analysis of specific human chromosomal regions. In addition to providing molecular linkage markers for specific human diseases, such libraries can yield probes detecting and differentiating barely detectable microscopic deletions or submicroscopic insertions within human chromosomes. Specific examples of the above include detection of structural deletions in human chromosome #13 (in retinoblastoma) or #15 (in the Prader-Willi syndrome, or PWS). Molecular studies with chromosome-specific DNA segments can serve not only to complement cytological analyses of metaphase chromosome abnormalities but also, e.g. in the PWS, to guide studies examining the nature of the molecular processes responsible for the chromosomal abnormalities observed. At the other extreme, molecular probes from chromosomes containing homogeneously staining regions (HSRs) can facilitate detection of DNA amplification in neuroblastomas and provide insight about novel processes associated with gene amplification. These include DNA relocation among chromosomes, long range DNA splicing, and DNA rearrangements of varying complexity. Overall, these studies provide a means to detect and understand the remarkable fluidity of the human genome.

TUMOR GROWTH FACTORS AND VACCINIA VIRUS GROWTH FACTORS: ROLE IN EPITHELIAL WOUND HEALING

G.J. Todaro[1], D. Twardzik[1], T. Purchio[1], R. Eiferman[2], and G. Schultz[2]. [1]Oncogen, Seattle, Washington; [2]University of Louisville Dept. of Biochemistry, Louisville, KY

Cells that have been transformed by various retroviruses, as well as many human tumors, produce biologically-active polypeptides called tumor growth factors. TGF-a, a prototype, is related to epidermal growth factor (EGF). Two major forms are seen; the smaller one contains 50 amino acids, It is produced from the larger form that, in the rat, contains 159 amino acids. by a novel protease that cleaves between alanine and valine at both the N-terminal and C-terminal ends. The 159 amino acid form has an extremely hydrophob c region that may serve to anchor the molecule in the cell membrane. If this hypothesis is correct, TGF-a could also be a cell membrane receptor and a potential tumor-specific cell surface antigen. Another member of the EGF/TGF family is an early protein from vaccinia virus which shows extensive sequence homology with both EGF and TGF-a. We speculate that the vaccinia virus growth factor (VGF) plays a role in the viral replicaton cycle and pernaps in the proliferative responses seen in poxvirus-infected cells.

Treatment of burn wounds on the backs of pigs show that both TGF-a and VGF greatly accelerate the re-epithelialization process when applied topically to the wound site. TGF-a has also been shown to be effective in the healing of penetrating wounds. A hybrid molecule that is partially TGF-a and partially VGF has also been created and has been shown to be biologically active.

TRANSFORMING GROWTH FACTORS - α AND - : PRECURSOR STRUCTURES AND BIOLOGICAL ACTIVITIES, R. Derynck, P. Lindquist, A. Rosenthal, T. Bringman, M. Winkler and D.V. Goeddel; Genentech, Inc., South San Francisco, California.

TGF-α is a 50 amino acid peptide which is structurally related to EGF and which mediates its action through the EGF receptor. TGF-β is a homodimeric protein consisting of two 112 amino acid subunits, is structurally unrelated to TGF-α and binds to a distinct receptor.

The TGF-α monomer is encoded as the C-terminal segment of a 390 amino acid long precursor which has a putative N-terminal signal peptide. The 160 amino acid TGF-α precursor also has an apparent N-terminal signal sequence but incorporates in addition, a putative membrane spanning domain downstream from the 50 amino acid TGF-α sequence. We overexpressed the human TGF-α in CHO-cells and determined that the precursor is a transmembrane protein which has a cytoplasmic domain which remains attached to the membrane. Processing at the periplasmic side releases the 50 amino acid TGF-α and two larger glycosylated TGF-α species.

The 50 amino acid TGF-α expressed in E. coli and EGF were compared for their biological activities in several systems. TGF-α and EGF are equally effective in inducing a mitogenic response in various cells and soft agar colony formation in the presence of TGF-β. However, TGF-α is much more potent than EGF in inducing neovascularization and bone resorption. These latter activities may be physiologically relevant in tumor formation in vivo.

Examination of a large variety of tumors and tumor cell lines has indicated that TGF-α expression is higher in [illegible] than in the surrounding tissues. TGF-α is not expressed in human pituitary tumor cell lines but is often synthesized in [illegible] tumors and carcinomas. The TGF-α expression is highest and most consistent in squamous and renal carcinomas which have elevated EGF receptor mRNA levels, in mammary carcinomas and in tumors of neuroectodermal origin. Because of the importance of TGF-α expression in solid tumors, we [illegible] TGF-α expression can [illegible]. Expression of [illegible] TGF-α mRNA in [illegible] results in [illegible] by the [illegible] independence and by efficient [illegible]. Anti-TGF-α antibodies prevent the expression of the transformed phenotype.

ISOLATION AND CHARACTERIZATION OF CLONES ENCODING BASIC FIBROBLAST GROWTH FACTOR: J.C. Fiddes, J.L. Whang, A. Mergia, J. Friedman, A. Tumolo, K.A. Hjerrild, D. Gospodarowicz[+], and J.A. Abraham, California Biotechnology, Inc., Mountain View, California; [+]Cancer Research Institute, University of California, San Francisco, California.

Basic fibroblast growth factor (FGF) is a potent mitogen for vascular endothelial cells *in vitro*, and stimulates new capillary growth (angiogenesis) *in vivo*; because of these two activities, this factor has been proposed as playing a role in such processes as tissue repair and the vascularization of solid tumors. Support for the latter proposed role comes from the observation that, based on amino acid composition, heparin-binding affinity, and migration on reverse-phase HPLC, basic FGF is very similar if not identical to tumor-derived mitogens isolated from a rat chondrosarcoma and from the SK-HEP1 human hepatoma cell line (Lobb *et al.*, *J. Biol. Chem.* 261:1924 [1986]).

We have isolated a clone encoding basic FGF from a bovine pituitary cDNA library. Sequence analysis of this clone indicates that basic FGF is initially synthesized with a 9 amino-acid N-terminal extension not found on the purified forms whose amino acid sequences have been determined. No classic signal sequence is evident. The identical 9 amino-acid extension is also encoded in the human basic FGF nucleotide sequence, which was derived from a combination of genomic clones and clones from kidney, fetal liver, and placenta cDNA libraries. Consistent with the proposal that basic FGF plays a role in tumor growth, we have also isolated cDNA clones from a human breast carcinoma cDNA library; in addition, two RNA species (5.0 and 2.2kb) were detected in Northern blots of SK-HEP1 polyA RNA when the bovine cDNA was used as a probe. Southern blot analysis and genomic cloning results indicate that basic FGF is encoded by a single gene split by at least two large (>15kb) introns.

HUMAN GROWTH HORMONE: FROM CLONE TO CLINIC P.H. Seeburg, Genentech, Inc., So. San Francisco, CA 94080.

A variety of genetic and acquired conditions (familial isolated growth hormone deficiency, Turner's syndrome, pituitary traumas) result in inadequate pituitary growth hormone levels leading to growth disorders. Human growth hormone (HGH), isolated from human cadavers, became available for the treatment of such disorders in the early 1960's. A decade later the novel field of recombinant DNA technology opened the possibility of using genetically engineered bacteria as a new source for HGH. Complementary DNA (cDNA) sequences encoding human chorionic somatomammotropin and rat growth hormone were characterized and cloned to serve as probes for the isolation of highly homologous HGH cDNA sequences. An historical account of the cloning of these sequences will be presented as well as the attempts to achieve their expression in Escherichia coli. Only a few years after it was shown that this bacterium could be engineered to synthesize rat GH sequences fused to the bacterial protein β-lactamase, HGH was produced at high yield. The expression vector used for this production directs the biosynthesis of HGH as a cytoplasmic protein carrying a methionine residue at the N terminus. This protein is deposited in the bacterial cells in the form of "refractile bodies" facilitating in purification of HGH which is achieved by a multiple step procedure yielding >95% pure hormone. This material proves indistinguishable from natural HGH in all biological assays and biological tests. It has been distributed through hospital pharmacies since late 1985 and will be succeeded by a recombinant product identical to natural HGH.

CLONING OF THE cDNA AND BIOLOGICAL PROPERTIES OF HUMAN MACROPHAGE GROWTH FACTOR, CSF-1: P. Ralph, K. Warren, J. Weaver, E. Kawasaki, M. Ladner, L. McConlogue, D. Mark, and T. White, Cetus Corporation, Emeryville, California.

CSF-1 is a glycoprotein which belongs to the family of colony-stimulating factors (CSF) that regulate the production of the blood cells. CSF-1 specifically promotes the growth of monocytes and macrophages from undifferentiated bone marrow progenitor cells. It also has a variety of stimulatory effects on mature cells. CSF-1 stimulates the production by monocytes of interferon, tumor necrosis factor, and a myeloid CSF which produces mainly mixed neutrophil-macrophage colonies in bone marrow culture. Pretreatment with CSF-1 also promotes resistance to viral infection and tumor cytotoxicity in murine peritoneal macrophages. Amino acid sequence data of purified CSF-1 from human urine and the pancreatic carcinoma line MIA PaCa show that the molecule appears to be a disulfide-linked homodimer with about 80% homology with murine CSF-1 produced by L929 cells. The human CSF-1 gene and cDNA clones were identified based on the protein sequence data. The gene is 18 kilobases (kb) in length and contains at least 9 exons. Although CSF-1 seems to be encoded by a single-copy gene, cells expressing the factor synthesize several CSF-1 messenger RNA species, ranging in size from 1.5 to 4.5 kb. One cDNA clone specifies the synthesis of biologically active protein in transfected, primate COS-7 cells as determined by CSF-1 specific radioreceptor and radioimmunoassay, macrophage bone marrow colony formation, and antibody neutralization. Nine of ten cDNA clones that were inactive in COS cells contain an aberrantly spliced intron sequence which interrupts the normal reading frame. The active cDNA specifies a 32 amino acid signal peptide followed by a protein of 224 amino acids. Several facts suggest, however, that one-third of the molecule at the C-terminal end is processed off intracellularly to derive the secreted growth factor. A genetically engineered polynucleotide with this region missing produced active CSF-1 in COS cells. CSF-1 has no obvious sequence homology with other hematopoietic factors or other proteins. In vitro and in vivo studies of CSF-1 suggest that it will have useful pharmacological properties. High level expression of the recombinant protein will allow careful preclinical testing in infectious disease and cancer models for therapeutic efficacy.

IN VIVO STUDIES WITH RECOMBINANT HUMAN GM-CSF:
R. Donahue and S. Clark, Genetics Institute, Cambridge, Mass.

Granulocyte-macrophage colony stimulating factor is a glycoprotein made by activated T-cells that displays a variety of biological activities. The protein was identified and its corresponding cDNA cloned based on its ability to stimulate the proliferation of granulocyte-macrophage progenitors. The recombinant protein has subsequently been shown to affect the biological activities of mature blood cells. In addition, we have found that the recombinant protein is a growth factor for erythroid progenitors indicating that GM-CSF interacts with earlier precursor cells than originally believed.

Initial experiments with ^{35}S-labeled CSF in rats demonstrated that the protein is cleared from circulation with multiphasic kinetics typical of many glycoproteins. The rate of clearance was not affected by desialylation of the CSF in vitro with neuraminidase. However, desialylated CSF was preferentially cleared in the liver while the untreated protein was cleared in the kidney.

Experiments in primates have demonstrated that continuous intravenous infusion of GM-CSF into healthy animals elicits a dramatic leukocytosis (WBC greater than 50,000) with an absolute increace in neutrophil, eosinophil, monocyte and lymphocyte levels. This leukocytosis can be observed 24 to 48 hours after the start of the infusion. The elevated WBC is absolutely dependent upon continued administration of the hormone. The leukocytosis has always resolved within 10 days of termination of the infusion. Initial experiments with pancytopenic rhesus monkeys (virally induced) have shown that some cytopenias may respond to GM-CSF treatment. One pancytopenic animal also exhibited a reticulocytosis upon CSF treatment consistent with the in vitro observations that GM- CSF stimulates early erythroid progenitors. These results suggest that GM-CSF may be clinically useful in treating some cases of cytopenias.

ERYTHROPOIETIN: GENE CLONING, PROTEIN STRUCTURE AND BIOLOGICAL PROPERTIES. J.K.Browne, A.M. Cohen, J.C. Egrie, P.H. Lai, F.K. Lin, T. Strickland, and N. Stebbing, Amgen, Thousand Oaks, California 91320.

Limited amino acid sequence analysis data, obtained from urinary erythropoietin (EPO) allowed the construction of oligonucleotide probes that were used to identify EPO specific mRNA isolated from the kidney of an anemic monkey. Human EPO genomic clones were similarly isolated from a human fetal liver library using mixed oligonucleotide probes. Comparison of the sequence of the genomic gene clones with the monkey cDNA sequence allowed determination of the intron-exon structure and of the amino acid sequence of human EPO. Mature EPO is 166 residues in length as determined by direct sequencing of the entire protein. Purified recombinant human erythropoietin (rHuEPO), expressed in CHO cells, and urinary EPO each contain three N-linked and one O-linked oligosaccharide chains with similar carbohydrate compositions. rHuEPO is indistinguishable from urinary EPO in three different assays: the in vivo exhypoxic polycythemic mouse bioassay; an in vitro rat bone marrow bioassay, and an RIA. Additionally, mammalian cell derived rHuEPO has equivalent activities in each of the bioassays and in the RIA indicating that CHO cell derived material is fully biologically active. Activity is dependent on glycosylation; EPO produced in E. coli or in yeast shows greatly reduced activity in the in vivo bioassay although it has equivalent activity in the in vitro bioassay and the RIA. The pharmacokinetics of rHuEPO is similar to the urinary EPO in mice and has been measured in nephrectomized animals and also in humans. CHO cell derived rHuEPO causes a dose dependent increase in the red blood cell mass in normal rodents and dogs. In rat models, rHuEPO is also effective in treating the anemia of chronic renal failure (CRF) and of inflammatory disease. In pre-clinical toxicity studies, rHuEPO has proved to be well tolerated. Anemia is a major complication of CRF due to a deficiency in the production of the hormone and is a primary clinical target for replacement therapy with rHuEPO. Purified, cell culture derived EPO is currently being evaluated in human Phase I clinical trials.

THE ABUNDANT LINE-1 FAMILY OF REPEATED DNA SEQUENCES IN MAMMALS: GENES AND PSEUDOGENES?

Maxine Singer and Jacek Skowronski, National Cancer Institute, NIH, Bethesda, MD 20892

Members of the LINE-1 family of long interspersed repeated sequences that is found in all mammalian DNA are structurally analogous to processed pseudogenes. The longest known elements are 6 to 7 kbp, occur 10^3 to 10^4 times per genome, have interrupted open reading frames (ORF) totalling as much as 5 kbp, terminate in 3'- A-rich stretches and can be surrounded by target site duplications of varying lengths. As many as 10^5 additional family members that are truncated at the 5'- end also occur; some of these are internally rearranged. LINE-1 units in different mammalian orders (e.g., primates and rodents) are homologous in the ORF segment but not in the regions that flank the ORF at the 5'- and 3'- sides.

In principle, functional gene(s) could exist among the thousands of LINE-1 sequences in most mammalian genomes. However, no family member(s) has been identified as a functional gene nor has a gene product derived from such a putative gene been identified. Homologous polyA$^+$ cytoplasmic RNA containing the full 6 kbp of the sense strand of a human LINE-1 is detected in human teratocarcinoma cells (NTera2D1) having the EC morphology, but not in any of a variety of other cell types. The 5'-end of the RNA coincides with the 5'-end of the longest LINE-1 genomic units. Twenty independently selected cDNA's, some of them nearly full length, were cloned from a library prepared with NTera2D1 cytoplasmic, polyA$^+$-RNA; no two are identical although most vary less than 3% in nucleotide sequence. Not all of them have uninterrupted ORF's. The consensus sequence of the 3'-noncoding region of the cDNA's is distinct from that determined for random genomic family members, suggesting that a particular subset of LINE-1 genomic sequences are transcribed in NTera2D1 cells and that this subset is independently homogenized. These data are relevant to various models proposed to explain the dispersal of this remarkably large family of repeats and its concerted evolution within individual mammalian species.

I. PROPERTIES OF A HUMAN TRANSPOSON-LIKE SEQUENCE AND II. EVOLUTION OF THE PRIMATE ALPHA GLOBIN GENE CLUSTER AND ITS INTERSPERSED ALU REPEATS: N. Deka, K. Eric Paulson, I. Sawada, C. Willard and C. W. Schmid, Department of Chemistry, University of California, Davis.

I: A repetitive family of human sequences, called the THE-1 family, is structurally similar to transposons such as yeast Ty, Drosophilla copia and mammalian proretrovirus. The THE-family consists of 10,000 member sequences sharing a 2 Kb consensus sequence. Members of this family are flanked by short direct repeats, which in some cases are similar sequences, suggesting THE-1 elements insert into preferred target sites. THE-1 elements are present in extrachromosomal closed circular DNAs. Both transcriptional senses are represented in poly A plus RNAs as well as polysome bound RNAs. In one case, a THE-1 element processes the 3' end of a transcript resulting from an upstream promoter. A 4 Kb transcript, which is specific to Hela cells, results from insertion of a THE-1 element into this transcription unit.

Monkey, chimpanzee and human share repetitive THE-1 families having a similar consensus sequence including orthologous member sequences. In contrast parts of the THE-1 human sequence are present as single copy sequence elements in a lower primate, galago, Conceivably these single copy sequence elements result from ancestral sequences which rearranged prior to their successful amplification in the human genome.

II: The structure of the galago alpha globin gene cluster more nearly resembles the structure of this region in such mammals as goat and horse rather than the structure in higher primates. The Alu family repeats in human and galago result from entirely independent insertion events. In contrast higher primates, human, chimpanzee and monkey, share orthologous Alu repeats. The divergence of these orthologous Alu repeats approximates the value expected for non selected DNA sequences.

THE HUMAN GENOME AND ITS EVOLUTIONARY CONTEXT :
Giorgio Bernardi and Giacomo Bernardi,
Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 75005 Paris, France.

The nuclear genome of man, like those of all warm-blooded vertebrates, exhibits a compositional compartmentalization, in that it mainly consists of a mosaic of very long (>>200 Kb) DNA segments, the isochores (see Science 228, 953, 1985 for a review). These (i) belong to a small number of classes characterized by different GC levels and by a fairly homogeneous base composition; and (ii) seem to correspond to the DNA segments present in Giemsa and Reverse chromosomal bands. The families of DNA molecules derived from isochore classes can be separated and used to study the genome distribution of any sequence that can be probed. This approach has revealed (i) that the distribution of genes in the genomes of warm-blooded vertebrates is highly biased towards the GC-richest isochores; (the latter are absent or poorly represented in the genomes of most cold-blooded vertebrates); (ii) that the human genome presents some specific features; and (iii) that the GC levels of both coding and non-coding sequences, as well as those of individual codon positions, show a linear dependence on the GC levels of the isochores harboring the sequences.

The latter observation prompted a study of the compositional constraints operating on over 60 genomes from a very wide range of organisms. This work has led to some general conclusions concerning (i) the effects of compositional constraints on codon usage as well as on the structure, stability and function of DNA, RNAs and proteins; (ii) the causes of compositional constraints and the processes underlying the fixation of mutations; (iii) the neutralist-selectionist controversy; and (iv) the nature of the organismal phenotype.

POPULATION GENETICS OF α THALASSAEMIA AND THE MALARIA HYPOTHESIS: A.V.S. Hill[1], J. Flint[1], D.K. Bowden[2], S.J. Oppenheimer[3], D.J. Weatherall[1] and J.B. Clegg[1], [1] M.R.C. Molecular Haematology Unit, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, U.K; [2] Department of Anatomy, Monash University, Victoria, Australia; [3] Department of Tropical Paediatrics, Liverpool School of Tropical Medicine, Liverpool, U.K.

Whereas both epidemiological and <u>in vitro</u> studies provide convincing support for the hypothesis that the sickle mutation provides protection against malaria, there is no such evidence that carriers of the extremely common forms of α^+ thalassaemia are similarly protected. Because of the very mild phenotype of this condition it was impossible to perform reliable population surveys before the advent of DNA analysis. We have studied the association between present-day frequencies of α-thalassaemia and historically-recorded prevalences of malaria in the Southwest Pacific. These island populations are particularly suitable for such analysis because of the rarity of other interacting haemoglobinopathies, their high prevalences of α-thalassaemia and marked geographical variation in malarial endemicity.

Within the large island of New Guinea very high frequencies of α-thalassemia were found in coastal populations with hyperendemic malaria, whereas in the highland populations, which have historically been free of malaria, α thalassaemia is rare. From the northwest to the southeast of Island Melanesia there is a marked fall in the gene frequency of α thalassaemia and this cline shows a clear correlation with historically recorded preintervention endemicities of malaria. Several factors support the proposal that these correlations are due to selection by malaria rather than genetic drift. Analysis of multiple unlinked DNA polymorphisms failed to reveal inter-population differences similar to those found for α thalassaemia. The cline in Island Melanesia is composed of varying proportions of three different single α-gene deletions. Finally restriction enzyme haplotype data support local origins for the α deletions rather than their introduction by migration. Limited data documenting the absence of single α gene deletions in populations historically free of malaria lend further support to the malaria hypothesis for α thalassaemia. The magnitude of the selective advantage and possible mechanisms involved will be discussed.

RECENT DEVELOPMENTS IN THE STUDY OF PRIMATE α GLOBIN GENE FAMILY: DISCOVERY OF NOVEL NEW MEMBERS AND RECONSTRUCTION OF HUMAN α THALASSEMIA-2 GENOTYPES IN MONKEY CELLS:

C.-K. James Shen, J. Marks, J. P. Shaw, W. S. Hu, M. Ayres, and Celia Shen, Department of Genetics, University of California, Davis, California.

For several years, the human α globin gene family has been known to consist of five members tightly clustered on chromosome 16 with the arrangement 5'-ζ(embryonic)-ψζ-ψα-α2(adult)-α1(adult)-3'. A recent study has located an additional pseudogene between ζ and ψα. Within the cluster, gene duplications and various other DNA recombination processes can be inferred.

To study these processes, we have cloned and characterized the adult α globin gene regions from several anthropoid primates. Unexpectedly, during these studies, we have discovered the existence of a new α globin gene subfamily: θ. One member of this subfamily, θ1, is located approximately 3 kb downstream from the α1 gene, and has been cloned from the human, orangutan, and baboon. It contains all the known essential DNA elements required for an expressible gene, and is well conserved among the anthropoid primates. The predicted θ1 protein is also 141 amino acids, and differs from the α at nearly as many aa (55) as does ζ (59), suggesting an ancient history of θ1 gene. Furthermore, genomic blot hybridization and molecular cloning have demonstrated that multiple θ1-homologous sequences exist in the primate genomes. Their chromosomal locations are being mapped.

We have also reconstructed the human α-thalassemia 2 genotypes in monkey cells. The adult α duplication units have been cloned in SV40 ori-containing vector and transfected into COS-7 cells. Newly replicated plasmid DNA were analyzed by Southern blot-hybridization, and by rescuing from bacterial strains. High frequency of homologous DNA recombination (10-20% per kb of homology), possibly generated by intrachromosomal crossing over, has resulted in both the rightward and the leftward α-thalassemia 2 genotypes. Different recombination sites have been mapped within the X block recombinants. Removal of the 5' end of either one of the X blocks greatly affects the frequencies of both deletions. Our data suggest that the homologous recombination of the α globin locus in primate somatic cells may be modulated by remote DNA sequence and by spacial arrangement of the recombination units.

COMPARISONS OF β-THALASSEMIA AND HEMOPHILIA A MUTATIONS: NEW LESSONS FROM A GIANT GENE. HH Kazazian[1], SE Antonarakis[1], C Wong[1], H Youssoufian[1], D Boehm[1], S-z Huang[1], S Goff[2], SH Orkin[2]. [1]Dept. of Peds., Johns Hopkins Sch. of Med., Baltimore, MD, [2]Dept. of Peds., Harvard Medical School, Boston, MA.

We have compared mutations affecting expression of two different genes, β-globin and Factor VIII:C (FVIII:C). The β-globin gene is small (1.5 kb), ribosomal, and mutations affecting its expression (β-thal genes) are subject to positive selection in regions endemic for malaria. In these regions many β-thal genes attain a frequency of 1% or more. Thirty-seven different point mutations and only 4 deletions producing β-thal have been characterized, many using a strategy of haplotyping β-thal chromosomes in each affected ethnic group and sequencing one β-thal gene from each haplotype. We estimate that 90% of the common mutations have now been observed; 7 have had more than one independent origin.

In contrast, the FVIII:C gene is very large (186 kb), X-linked, and mutations producing hemophilia A have a finite life span of 100-150 years in the population. For these reasons the FVIII:C gene provides a less biased view than β-globin of the frequency of occurrence of various mutations affecting gene expression. Upon screening 70 hemophilia A genes by restriction analysis, we found 7 different deletions (10%) and 6 point mutations (8%). The remaining 57 genes presumably contain undiscovered point mutations. Of the known point mutations, two are identical, a CGA-TGA change in codon 1960. Two independent origins of this mutation have been proven by family studies. Two other mutations are located at the Taq I site in exon 22 and, while still uncharacterized, are potential CGA-TGA mutations. These data along with those of Gitschier et al suggest that 1) the frequency of deletions as a cause of deleterious mutations depends upon the target size of the relevant gene and 2) as suggested previously, CpG dinucleotides are significant mutation hotspots. Based on the frequency of mutations in Taq I sites and the fraction of coding CpG dinucleotides which are present in Taq I sites, we estimate that up to 1/3 of all mutations producing hemophilia A are CpG-TpG or CpG-CpA changes. These mutations do not appear in β-thal genes because the β-globin gene lacks CpG in the first two or second two residues of codons.

THE T CELL SURFACE PROTEIN, T4: POSSIBLE ROLES IN THE CELLULAR IMMUNE RESPONSE AND IN THE PATHOGENESIS OF AIDS. Richard Axel[1,2] and Paul Maddon[1], [1]Department of Biochemistry, and [2]Howard Hughes Medical Institute, College of Physicians and Surgeons, Columbia University, New York, New York.

The surface glycoproteins, T4 and T8, define different functional subsets of T lymphocytes and may act as recognition molecules mediating appropriate interactions between the T cell and its target. We have isolated the cDNAs and genomic clones encoding the T4 and T8 molecules using gene transfer and subtractive hybridization. T4 and T8 share only 28% homology, but both molecules share significant sequence and structural homology with the immunoglobulin variable domains and are members of the immunoglobulin supergene family. Experiments in collaboration with Denise Gay and Philippa Marrack suggest that T4 interacts with a component of the target cell surface and this interaction governs the specificity and efficiency of T cell functions. The T4 molecule also serves as a receptor for the family of AIDS viruses, HTLV-III/LAV. In collaborative studies with J. Steven McDougal and Robin Weiss, we have shown that the virus binds efficiently to the T4 molecule on the surface of $T4^+$ lymphocytes as well as $T4^+$ non-lymphoid cells. Moreover, the introduction of the T4 molecule into non-lymphoid human cells allows for the efficient infection and replication of AIDS virus. Thus, a T helper cell receptor thought to function as a recognition molecule in the cellular immune response also serves as a receptor for AIDS virus.

MOLECULAR FEATURES INVOLVED IN CELL SURFACE RECEPTOR FUNCTION AND THEIR ROLE IN ONCOGENESIS: A. Ullrich[1], H. Riedel[1], J. Lee[1], A. Gray[1], L. Coussens[1], T. Dull[1], M. Waterfield[2], I. Verma[3], L. Williams[4], and J. Schlessinger[5], [1]Genentech, Inc., South San Francisco, California; [2]Imperial Cancer Research Fund, London, England; [3]The Salk Institute, La Jolla, California; [4]University of California, San Francisco, California; [5]The Weizmann Institute of Science, Rehovot, Israel.

Growth factors and their receptors are involved in the regulation of cell proliferation, and a variety of recent findings suggest that they may also play a key role in oncogenesis. Of approximately twenty identified oncogenes, the three that have been correlated with known cellular proteins have each been found to be related to either a growth factor or a growth factor receptor.

The receptor-related oncogenes v-erbB (EGF receptor) and v-fms (CSF-1 receptor) are members of a gene family by virtue of the fact that they possess tyrosine-specific protein kinase activity and are associated with the plasma membrane. Such features are also shared by several other polypeptide hormone receptors, including those for insulin, platelet-derived growth factor (PDGF), and insulin-like growth factor 1, which suggests that more connections may be found between tyrosine kinase growth factor receptors and tyrosine kinase oncogene products.

To investigate the structural features involved in the normal biological function and oncogenic potential of cell surface receptors bearing tyrosine kinase activity, we characterized cDNA clones coding for the human EGF receptor, the human homolog of the neu oncogene HER2, the human insulin receptor, IGF-1 receptor, c-fms/CSF-1 receptor and the mouse PDGF receptor. This structural analysis revealed the existence of three classes of receptors possessing tyrosine kinase activity, and provided clues for the construction of receptor mutants to investigate receptor structure-function relationships. These experiments strongly suggest that intracellular, carboxy-terminal sequences play a key role in the normal and oncogenic function of cell surface receptor tyrosine kinases.

ORGANIZATION AND EXPRESSION OF THE HUMAN TRANSFERRIN RECEPTOR GENE:

F. H. Ruddle, W. K. Miskimins and M. P. Roberts,
Department of Biology, Yale University, New Haven, CT

In our earlier studies, we cloned (Kuhn et al., Cell 37:95,'84) and then characterized the molecular structures of the human transferrin gene (McClelland et al., Cell 39: 267,'84). Our current investigations relate to the regulatory function of the 5' regions of the gene especially in relationship to cellular proliferation. Using plasmid constructs in which various segments of this sequence are linked to the CAT gene, we have defined several important control regions. A 365 base pair fragment which includes the first exon, part of the first intron, and 116 base pairs of flanking sequence proved to contain a promoter. This fragment was sequenced and shown to be highly GC-rich. It contains a normal TATA box and, just upstream, a unique sequence containing four tandem repeats of GGGGC. A similar series of repeats is found just upstream of the TATA boxes of the human dihydrofolate reductase gene and the mouse interleukin 3 gene. The expression of both of these genes, like the transferrin receptor gene, is dependent on mitogen stimulation of cell proliferation. These data suggest a common mode of regulation for these and other genes in this class. Using a protein blotting procedure (Southwestern blotting: Miskimins et al., PNAS 82:6741,'85), we have identified a group of nuclear proteins which bind with high affinity to this region. One of these proteins (88 kd) binds preferentially to a 140 base pair fragment of flanking sequence which contains the TATA box and GC repeats. The activity of this protein is present at very low levels in quiescent 3T3 cells but is rapidly induced by serum stimulation, reaching a maximum about 6 h after stimulation. The increase in this 88 kd protein precedes a similar increase in transferrin receptor message. We have further studied DNA-protein interaction at the transferrin receptor promoter by electrophoresis of complexes on low ionic strength polyacrylamide gels and by DNase I footprinting. We have identified several protein binding sites just upstream of the TATA box. Two protected sites overlap the GC repeat region. One of these is centered around a concensus SP1 binding sequence and protection is likely to be due to this protein. Two additional binding sites, which have homology to each other, are located upstream.

HOW HUMAN T LYMPHOCYTES GROW: A MOLECULAR ANALYSIS OF THE HUMAN INTERLEUKIN-2 RECEPTOR: Warner C. Greene, Metabolism Branch, National Cancer Institute, NIH, Bethesda, MD.

The human receptor for interleukin-2 (IL-2) plays a pivotal role in the growth of human T cells. IL-2 receptors are not present on resting T cells, however antigen activation induces both IL-2 and IL-2 receptor gene expression which mediate T cell proliferation. The IL-2 receptor is encoded by a single structural gene composed of 8 exons and 7 introns located on chromosome 10p band 14-->15. Following antigen activation, as many as 12 different IL-2 receptor mRNAs are produced differing due to the use of two discrete transcription initiation sites, variable post-transcriptional splicing of the fourth exon, and polyadenylation at three different sites. Expression of IL-2 receptors is unsustained during T cell activation and thus may contribute not only to the development of the T cell immune response but also to its subsequent termination. Based on the sequence of a full length cDNA, the IL-2 receptor protein is composed of 251 amino acids and contains an unexpectedly short (13 residues) intracytoplasmic domain. Protein studies demonstrate that the receptor is a densely glycosylated, sulfated, and phosphorylated 55,000 M_r structure which may exist in both high (functional) and low (nonfunctional) affinity states. The high affinity but not low affinity form of the IL-2 receptor undergoes receptor mediated endocytosis. Retroviral introduction of IL-2 receptor cDNA into human T cells results in both high and low affinity IL-2 receptor display, however in mouse fibroblasts only low affinity IL-2 receptors are expressed.

Leukemic T cells infected with the HTLV-I or HTLV-II retroviruses are characterized by the continuous expression of large numbers of high and low affinity IL-2 receptors. Deregulated IL-2 receptor expression in these cells is not produced by translocation, rearrangement, or amplification of the IL-2 receptor gene. Recent studies, performed in collaboration with W.A. Haseltine and colleagues, have demonstrated that the transactivator gene (tat) of HTLV-II is capable of inducing IL-2 receptor gene expression in Jurkat T cells but not in Raji B cells or mouse fibroblasts. In addition, the tat gene also induces expression of the cellular IL-2 gene. These data suggest the possibility that these retroviruses may subvert normal T cell growth control by inducing coexpression of the cellular genes encoding a critical T cell growth factor and its receptor.

EXPRESSION AND FUNCTION OF THE HUMAN INTERLEUKIN 2 RECEPTOR: T. Honjo, A. Shimizu, S. Kondo, H. Sabe, N. Ishida, M. Kinoshita, U. Saito, N. Suzuki, H. Kanamori, N. Matsunami and Y. Yaoita, Department of Medical Chemistry, Kyoto University Faculty of Medicine, Kyoto, Japan.

Interaction between interleukin (IL)-2 and the IL-2 receptor plays an essential role in antigen-stimulated proliferatiaon of T cells. The IL-2 receptor is induced temporally on stimulated T cells. We have analyzed the promoter function of the IL-2 receptor gene by in vivo and in vitro transcription systems. We identified the 5' 200-bp region of the gene responsible for cell-specific expression.

Comparison of the structures of the human and murine receptors allowed us to identify several conserved regions localized to exons 2 and 4, cytoplasmic portion and the transmembrane portion, which might be important for the functions of the IL-2 receptor. The human IL-2 receptor, which was expressed on an IL-2 dependent murine T cell line, CTLL-2, by cDNA transfection, was shown to be functionally active by blocking the endogenous mouse IL-2 receptor with monoclonal antibodies. On the other hand, the human IL-2 receptors expressed on non-lymphoid cells were functionally inactive. They were unable to mediate the growth signal, were of low affinity species and aberrant in internalization. We postulated that the dysfunction of the IL-2 receptors in non-lymphoid cells would be due to the absence of the putative "converter" protein which is expressed specifically in lymphoid cells. Since the human IL-2 receptor is active in the murine T cell, the converter may interact with the receptor at the portions conserved between man and mouse. Function of the IL-2 receptor was also analyzed using EGF/IL-2 chimeric receptors, and monoclonal antibodies directed against different determinants of the IL-2 receptor. Based on these studies we proposed the affinity conversion model that explained the high affinity state of the receptor by the ternary complex formation between IL-2, the IL-2 receptor and the converter.

STRUCTURE AND FUNCTION OF OESTROGEN AND GLUCOCORTICOID RECEPTORS.

P. Chambon, S. Green, A. Krust, V. Kumar, P. Walter, M. Devic, M.V. Govindan and U. Stropp, LGME/CNRS et U184/INSERM, Institut de Chimie Biologique, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg-Cédex, France.

We have previously isolated and sequenced cDNAs corresponding to the oestrogen (hER) (1) and glucocorticoid (hGR) (2) receptors of the human MCF-7 breast cancer cell line and to the chicken oestrogen receptor (cER) (3). Cytosol extracts of HeLa cells, transiently expressing a vector containing the hER cDNA sequence, contain an oestrogen-binding protein with a similar affinity for oestradiol (K_D= 0.4-0.5nM) and the same molecular weight (66 kDa) as the MCF-7 cell ER receptor. The human and chicken ER amino acid sequences share 80% homology. There are three highly conserved regions. Using site-directed mutagenesis we have demonstrated that two of these regions correspond to the DNA and hormone binding domains. The DNA binding domain is characterized by its high content of cysteine and basic amino acids and the hormone binding domain by its overall hydrophobicity. The integrity of the DNA binding domain is required for tight binding of the receptor to the nucleus, but not for steroid hormone binding. Regions homologous to these two domains are found in the hGR, in the chicken progesterone receptor and in v-erbA, the oncogene potentiator gene of the avian erythroblastosis virus (AEV). The cellular counterpart (c-erbA) of v-erbA is therefore likely to belong to a multigene family of transcriptional regulatory factors which bind steroid-related ligands. The other conserved ER region is not present in v-erbA, but some homology is found in the N-terminal portion of the hGR. The implication of these results for steroid receptor and erbA protein structure and function will be discussed.

1. S. Green, P. Walter, V. Kumar, A. Krust, J.M. Bornert, P. Argos and P. Chambon (1986). Nature 320, 134-139.
2. M.V. Govindan, M. Devic, S. Green, H. Gronemeyer and P. Chambon (1985). Nucl. Acids Res. 13, 8293-8304.
3. A. Krust, S. Green, P. Argos, V. Kumar, P. Walter, J.M. Bornert and P. Chambon (1986). The chicken oestrogen receptor sequence : homology with v-erbA and the human oestrogen and glucocorticoid receptors. EMBO J., in press.

STRUCTURE AND EXPRESSION OF HUMAN GLUCOCORTICOID RECEPTOR cDNA: A TRANS-ACTING FACTOR RELATED TO THE c-erbA PROTO-ONCOGENE FAMILY. R.M. Evans[1], C. Weinberger[1], S. Hollenberg[1], Vincent Giguere[1], and M.G. Rosenfeld[2]; [1]Howard Hughes Medical Institute, Gene Expression Laboratory, The Salk Institute, San Diego, California; [2]School of Medicine, University of California, San Diego, California.

We are interested in developmental and inducible regulation in the neuro-endocrine system. In animals, the endocrine and nervous systems produce a diverse set of molecules which interact to exert profound effects on development, physiology and behavior. All these molecules exert their effects by binding specific membrane or intracellular receptors. Our interest in steroid receptors is that these molecules serve as a paradigm for sequence-specific DNA binding proteins that are capable of directly regulating gene transcription. To investigate the molecular mechanisms by which it regulates gene transcription, we have cloned receptor cDNA sequences. Analysis reveals two human glucocorticoid receptor (hGR) variants (alpha, 777 amino acids; and beta, 742 amino acids), differing at their carboxy termini. The proteins contain a cystine-lysine-arginine rich region which may define the DNA-binding domain. Expressed full-length hGR is immunoreactive, possesses intrinsic steroid binding activity and exerts potent transcriptional enhancement on steroid responsive genes. Analysis of the receptor sequence reveals it to be distinct from other previously characterized DNA binding proteins and suggests that it may be part of a new class of regulatory molecules. Furthermore, the hGR as well as the recently characterized estrogen receptor are related to the oncogene v-erbA from the Avian Erythroblasosis Virus (AEV) suggesting that steroid receptor genes and the c-erbA proto-oncogene are derived from a common primordial regulatory gene. Oncogenicity by AEV may result, in part, from the inappropriate activity of a truncated steroid receptor or related regulatory molecule encoded by the v-erbA gene. This suggests a mechanism by which trans-acting factors may facilitate or potentiate transformation. To further characterize these relationships we have cloned a human and rat c-erbA protooncogene. This analysis reveals extensive homology and suggests common functional domains.

GENETIC MAPPING OF THE HUMAN X CHROMOSOME : LINKAGE ANALYSIS OF THE q27-qter REGION, INCLUDING THE FRAGILE X LOCUS AND ISOLATION OF EXPRESSED SEQUENCES.

J.L. Mandel[1], B. Arveiler[1], G. Camerino[2], A. Hanauer[1], R. Heilig[1], M. Koenig[1] and I. Oberlé[1], [1]LGME/CNRS et U184/INSERM, Faculté de Médecine, 67085 Strasbourg-Cédex, France and [2]Dipartimento di Genetica e Microbiologia - Universita di Pavia.

We are mapping the Xq27-q28 region which includes several important loci : Hemophilia A and B, Adrenoleukodystrophy, and the mental retardation with a fragile X. This latter disease is the most frequent of all X-linked diseases, and a major cause of mental retardation. It presents several important problems : cytogenetic analysis is insufficient especially for carrier detection, segregation of the disease has unusual characteristics (incomplete penetrance in males, high mutation rate, etc.), and the relationship between the cytogenetic and pathologic manifestations are not understood. All these would benefit from analysis with closely linked probes and ultimately from the cloning of the fragile X locus itself. We have mapped several markers, including the hypervariable St14 locus (DXS52) and the coagulation factor IX gene. Our results suggest the order cen - DXS100 - DXS86 - DXS144 - DXS 51 - F9 - FraX - (DXS52, DXS15). Recombination frequency between St14 or F9 and FraX is about 15%, and there is some evidence for heterogeneity in the genetic distance between F9 and FraX in different families. Closer markers are needed, but genetic mapping of additional probes is impaired by the low frequency of RFLPs, an apparent characteristic of X linked probes.

As a complementary approach towards mapping of the human X chromosome, we have searched for expressed X linked sequences since this might allow one to correlate a disease locus with a sequence located in the same region and expressed in the appropriate tissue. We have partially characterized several X linked genes and pseudogenes. One of the most powerful way of finding expressed sequences in randomly cloned fragments is to search for homologies with rodent genomes. Comparison of the homologous human and rodent sequences allows one to characterize putative protein coding regions, even though the corresponding mRNA has not yet been detected. This sould be useful in "genome walking" strategies designed to isolate a disease gene.

MOLECULAR GENETICS OF MIC2 : AN X AND Y CHROMOSOME LOCATED EXPRESSED GENE: P.N. Goodfellow[1], P.J. Goodfellow[1], J. Wolfe[2], G. Banting[1], B. Pym[1], C. Pritchard[1] and S.M. Darling[1], [1]Laboratory of Human Molecular Genetics, ICRF, London, UK; [2]Department of Human Genetics, University College, London, UK

The Y-located gene MIC2 encodes a cell surface antigen defined by the monoclonal antibody 12E7. In addition to the Y-located gene, the human genome contains an X-located MIC2 gene which has been shown to escape X-inactivation. The X and Y chromosomal forms of MIC2 have been compared by immunochemical analysis of their products, by molecular cloning and by family studies. No difference was found in size or charge in the antigenic products of the MIC2 genes and restriction mapping the MIC2 loci on the X and Y chromosomes suggested that the two loci were either identical or very similar. The genomic sequences corresponding to the MIC2 loci are highly polymorphic and identity between the two loci has been formally shown by studying segregation of DNA polymorphisms in families. Exchange between the sex chromosomes at the MIC2 locus occurs at a low frequency (less than 5% recombination) implying that MIC2 lies at the centromeric end of the pseudoautosomal region. If current genetic maps of the Y chromosome are correct then MIC2 flanks the region containing the sex determining gene(s).

Using the 12E7 antibody as a selective agent in conjunction with the fluorescent activated cell sorter we have isolated hybrids containing fragments of the Y chromosome. These hybrids provide a source material for saturation cloning of the sex determining region of the human Y chromosome.

THE PSEUDOAUTOSOMAL REGION OF THE HUMAN SEX CHROMOSOMES

F. Rouyer, M.C. Simmler, G. Vergnaud, C. Johnsson and Weissenbach, INSERM U163, CNRS UA271, Institut Pasteur, Paris, France.

Strictly homologous DNA loci shared between the human X and Y chromosomes have been characterized by several laboratories. They have been assigned by in situ hybridization and/or deletion mapping with X-autosome translocations to the short arm tip of both the X and Y chromosomes. A number of these DNA loci exhibit RFLPs with numerous allelic forms shared by both sex chromosomes. In family studies we have observed that each paternal allele can segregate both in male and female progeny demonstrating the occurrence of recombination between the X and Y chromosomes. The term pseudoautosomal has been coined (Burgoyne, 1982) to denote the pattern of inheritance of sex chromosomal loci displaying no absolute sex linkage.

In order to define the general genetic features of the pseudoautosomal region, a quantitative analysis of 6 pseudoautosomal DNA loci (including a telomeric locus) has been performed on about 100 meioses (DNA samples kindly provided by CEPH). This analysis shows that the loci do not segregate independently, as illustrated by the fact that no double recombination has been observed. The recombination events can therefore be ascribed to a crossing-over mechanism. In male meiosis, the telomeric locus recombines with sex phenotype at a frequency of 50%, whereas the other pseudoautosomal loci recombine less frequently and define a gradient of sex linkage. Furthermore the empiric additivity rule of recombination frequencies applies to the pseudoautosomal region for values up to 50%. Taken together these results strongly support the idea that a single, but not uniquely localized, crossing-over between X and Y takes place during XY synapsis.

The recombination frequency between pseudoautosomal loci is 10-fold greater in male meiosis than that obtained from female meiosis. It has been proposed (Koller & Darlington, 1934) that at least one chiasma occurs in each meiotic bivalent to ensure proper segregation. This 10-fold increase in recombination frequency in the male pseudoautosomal region is then a direct consequence of a chiasma having to be formed in such a limited region.

SEX-REVERSAL: DELETION MAPPING THE MALE-DETERMINING FUNCTION OF THE Y CHROMOSOME. David C. Page, Laura Brown, Kristina Bieker, Jonathan Pollack. Whitehead Institute, Cambridge, Massachusetts.

In humans, gonadal sex is determined by the presence or absence of the Y chromosome. There are a number of "sex reversal" syndromes that appear to contradict this rule--syndromes in which testes develop in the apparent absence of the Y chromosome (XX males, XO males, XY hermaphrodites), or in which ovaries develop in the presence of the Y chromosome (XY females). However, DNA hybridization studies reveal the presence of certain Y-DNA sequences in the genomes of most XX males, and the absence of certain Y-DNA sequences from the genomes of many XY females. In total, we tested 50 individuals for the presence of as many as 140 Y-DNA loci by hybridization to Southern transfers. Most of these persons tested were XX or XO males, or XY females, or had, as judged by cytogenetics, a structurally abnormal Y chromosome. On the basis of the results obtained, we have constructed a 10-interval deletion map of the Y chromosome. (An earlier, 7-interval map is described in Vergnaud, Page et al., Am. J. Hum. Genet. [1986] 38:109.) This map is oriented with respect to the long and short arms of the Y chromosome, and the position of the centromere is known. The testis-determining factor (TDF) is unambiguously mapped to one of the six deletion intervals on the short arm; this interval in present in all XX and XO males in whom we have detected any Y DNA, and is absent from all XY females with detected Y deletions. Our results suggest that XX males and XY females may be reciprocal products of similar anomalous exchanges between the X and Y chromosomes. The deletion map may be dimorphic, i.e., there may exist two structurally distinct Y chromosomes in the normal population. TDF appears to be interposed between two substantial blocks of X-Y homology on the short arm of the Y chromosome. Proximal to TDF is a large block of X-Y homology that appears to be the result of a recent transposition from Xq13-q21, and that is not subject to X-Y exchange during male meiosis. Distal to TDF is the pseudoautosomal domain, a block of X-Y homology that undergoes frequent X-Y exchange during male meiosis. We are using pseudoautosomal restriction-fragment-length polymorphisms to test whether XX males have received the telomere of their fathers' Y chromosomes, as predicted by the X-Y interchange model.

VARIABILITY AT THE TELOMERES OF THE HUMAN X/Y PSEUDO-AUTOSOMAL REGION:

H.J. Cooke[1] and B.A. Smith[1], [1]MRC Mammalian Genome Unit, King's Buildings, West Mains Road, Edinburgh, Scotland.

The telomeric regions of the pseudoautosomal region of the human sex chromosomes shows four levels of variability. A centromere proximal region 40kb from the telomere end consists of 18kb of tandem repeats found only at this location. This block of repeats shows extensive variability at the RFLP level, one common allele has no TaqI sites, another has 13. More distal is a short non-variable region followed by an area defined by repeated sequences present in 1 to 3 copies per chromosome which show extensive RFLP variability with all enzymes tested. The number of alleles in the population is probably in excess of 100. This variability may be a reflection of the high recombination frequency in this region of the genome in male meiosis.

The telomeres themselves show evidence of variable lengths in DNA from a single individual suggesting the presence of different numbers of copies of a short terminal repeat. This variability is much reduced or absent in cloned cell populations - e.g. a B cell leukaemia or cloned lymphoblastoid lines. The germ line carries 5kb of extra DNA at these telomeres in comparison to somatic tissues.

MOLECULAR BIOLOGY OF Y CHROMOSOME AND ANOMALIES OF SEX DETERMINATION IN MAN :

M.Casanova[1], E. Seboun[1], P. Leroy[1], C. Disteche[2], E. Magenis[3], C. Bishop[1] and M. Fellous[1].

[1]Immunogénétique Humaine et U.INSERM 276,Institut Pasteur Paris, [2] Department of Pathology,University of Washington Seattle, [3] Department of Medical Genetics, Crippled children's Division, Oregon Health Sciences University Portland

In the human XX maleness syndrome, the 46XX caryotype does not corroborate with male sexual phenotype. Using Y specific probes from a human Y library, DNA of 40 XX males has been studied. 66% of patients reacted with the probes used but to a different extend. Comparison of the hybridization patterns with DNA from patients bearing abnormal Y chromosomes allows us to assign most of the Y specific bands detected in XX males to the short arm of the Y chromosome. Direct in situ hybridization strongly supports the occurence of Y short arm translocation to the distal short arm of one of the X chromosome in the XX male syndrome.

A counter part of such exchange was observed in few cases of 46 XY female (Pure gonadal dysgenesis syndrome) with absence of corresponding Yp DNA sequences.

We have also isolated a probe named p49f. This probe detected an autosomal (chromosome n° 3) and Y sequences (Y_q11-23). This probe hybridizes with a 4kb testis specific mRNA, present in different species including man. The corresponding cDNA has also been characterized. Using in situ hybridization, the corresponding transcript are detected in male germ cells, late during spermatogenesis.

SEX REVERSAL: GENETIC AND MOLECULAR STUDIES ON 46,XX AND 45,X MALES.

Albert de la Chapelle, Department of Medical Genetics, University of Helsinki, Finland

The first significant finding that provided a hint as to the etiology of maleness without a Y chromosome was that some XX males had inherited their father's Xg allele while others had not. This finding led to two hypoteheses, (1) at least some XX males arise through meiotic interchange of genes between the X and Y chromosomes at paternal meiosis (2) the mechanisms are not the same in all cases. Subsequent findings have corroborated both hypotheses. Based on X chromosomal RFLP's, most XX males have one paternal and one maternal X chromosome. In one case serological evidence indicated that the father had transmitted to his XX male son an X chromosome without the Xg but carrying the "Yg" gene that normally occurs in an Xg-like locus on the Y. DNA sequences specific for the Y chromosome occur in some 60% of XX males while in some 40% none of the Y-derived probes so far tested occur. Finally, by in situ hybridization Y chromosome-derived sequences occur on the tip of the short arm of one X chromosome in some XX males. The testis determining factor (TDF) which is normally located on the Y chromosome has not yet been characterized or isolated. Nevertheless data on XX males are consistent with the hypothesis that TDF is carried by the paternal X chromosome in some XX males. In others, including the rare familial cases, TDF may be carried on another chromosome. Alternatively another testis-determining mechanism is involved. In 45,X males the etiology must be different at least in those cases where the single X chromosome is maternal. In at least some 45,X males low-grade mosaicism involving a 46,XY cell line has been detected by cytogenetic and molecular studies, while in others, Y chromosome-derived material occurs elsewhere in the genome. Thus the mechanisms giving rise to 46,XX and 45,X maleness are different and heterogeneous.

MOLECULAR GENETICS OF HUMAN B AND T CELL NEOPLASIA:
Carlo M. Croce, The Wistar Institute, Philadelphia, PA

Specific chromosomal translocations are involved in more than 80% of human B cell neoplasms. In all these cases the neoplastic phenotype is apparently the consequence of reciprocal chromosomal translocations involving the loci for the human immunoglobulin chains and either well described cellular proto-oncogenes or putative proto-oncogenes. The juxtaposition of the proto-oncogenes to the immunoglobulin loci results in their transcriptional deregulation, because of their proximity to genetic elements within the human immunoglobulin loci capable of activating gene transcription in cis over considerable chromosomal distances. Sequence analysis of the translocation breakpoints has provided important insights concerning the molecular mechanisms involved in chromosome translocation in B cells. It appears that the reciprocal translocation contributing to B cell neoplasia are catalyzed by the same enzymes that are involved in physiological immunoglobulin gene rearrangements.

The analysis of human B cell leukemias and lymphomas has also provided considerable information concerning the possible scenarios for B cell neoplastic transformation. It is clear that the Epstein Barr virus does not play a direct role in neoplastic transformation, but it may contribute by increasing the number of B cells at risk of developing chromosome translocations during immunoglobulin gene rearrangements.

Cytogenetic and molecular genetic analysis of T cell malignancies is beginning to provide a very similar scenario for neoplastic transformation. The locus for the α chain of the T cell receptor is directly involved, and it apparently juxtaposes to proto-oncogenes or to putative proto-oncogenes leading to their transcriptional deregulation. It seems quite likely that the enzyme system involved in rearrangements of the genes for the T cell receptor plays a crucial role in the causation of these chromosomal translocations. Thus, the genetic basis of many human B and T cell neoplasias may be quite similar. For the future, the challenge resides in trying to characterize specifically the role of both old and new proto-oncogenes in B and T cell proliferation, normal and neoplastic.

T AND B CELL RECEPTOR GENE REARRANGEMENTS AND LYMPHOID TUMOUR AETIOLOGY

T.H. Rabbitts, R. Baer, L. Bulawela and L. Mengle-Gaw.
Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, England.

The majority of tumours exhibit detectable chromosomal abnormalities such as translocations, inversions or interstitial deletions. Furthermore, the genes which undergo rearrangements in the normal course of B and T cell differentiation (i.e. immunoglobulin and T cell receptor genes respectively), map to the chromosomal breakpoints of certain types of lymphoid tumours. It seems likely that the activation of genes by the proximity to B and T cell receptor genes, resulting from these gross chromosomal abnormalities, has a causative role in the formation of these tumours. To test this hypothesis, we have studied the molecular structure and consequences of chromosomal breakpoints in human B and T cell tumours. Two types of abnormality will be discussed. Firstly, rearrangements of immunoglobulin genes in B cells involving the c-myc proto-oncogene on chromosome 8, and secondly, association of immunoglobulin and T cell receptor genes in the characteristic inversion of chromosome 14 (q11;q32) in human T cell chronic lymphocytic leukaemia.

In variant Burkitt's lymphomas (t2;8) (p12;q24) the kappa light chain genes are often joined at some distance downstream of the c-myc gene. The position with respect to the c-myc gene at which this can occur (analogous to 'pvt' described by S. Cory and J. Adams) has also been found in human T cell leukaemias and this observation will be discussed with reference to tumour specificity.

The T-cell leukaemia specific inversion 14 can result in the formation of an IgT gene which consists of a productive join between an immunoglobulin heavy chain variable region gene and a T cell receptor alpha chain joining/constant region gene. Such a hybrid molecule may be involved in the chronic proliferation of the pre-malignant T cell leukaemia cell prior to the formation of the malignant cell in which further gene aberrations have occurred.

JUXTAPOSITION OF TRANSFERRIN (TF) AND TRANSFERRIN RECEPTOR (TFR) GENES IN THE INV(3) AND T(3;3) IN ACUTE NONLYMPHOCYTIC LEUKEMIA (ANLL): M.M. Le Beau[1], R. Espinosa[1], F. Yang[2], C. Schneider[3], R.A. Larson[1], J.D. Rowley[1], and M.O. Diaz[1], [1]Joint Section of Hematology/Oncology, Univ. of Chicago, IL; [2]Department of Cellular and Structural Biology, Univ. of Texas Health Science Center at San Antonio, TX; [3]Laboratoire Européen de Biologie Moléculaire, Heidelberg, Germany.

Recent data indicate that the genes located at the breakpoints of the recurring chromosomal rearrangements in human tumors play an integral role in tumorigenesis; during the past few years, transforming sequences as well as metal ion binding proteins have been mapped to these breakpoints. In ANLL, rearrangements of chromosome 3 involving both bands q21 and q26 [t(3;3)(q21;q26), inv(3)(q21q26)] are associated with thrombocythemia and abnormal megakaryocytopoiesis. Genes mapped to these regions include TF (3q21-25) and TFR (q26-ter). TF, the major iron-binding protein in serum, is a growth factor that is required for the proliferation of cells both in vivo and in vitro.

By using the technique of in situ chromosomal hybridization (ISH) to examine leukemia cells from ANLL patients with an inv(3) or t(3;3), we have demonstrated that the breakpoints at q21 and q26 occur within the TF and TFR genes, respectively. ISH of a TF cDNA probe to metaphase cells with an inv(3) or t(3;3) resulted in labeling of the two breakpoint junctions of the inv(3), and of the junctions of the 3q+ and 3q- chromosomes in the t(3;3). Hybridizations using the 5' and 3' regions of the TF probe revealed that the 3q21 breakpoint occurred near the middle of the TF gene, close to the boundary of the two homologous domains, and that this gene is oriented 5' to 3' on 3q.

Hybridizations using coding and 3' untranslated regions of the TFR gene revealed that the q26 breakpoint in these rearrangements occurs within the TFR gene, near the boundary of the coding sequences and the 3' untranslated region. These hybridizations also indicate that TFR is oriented 3' to 5' on 3q.

Our studies demonstrate that parts of the TF and TFR genes are juxtaposed at the junctions of the rearranged chromosomes in the inv(3) and t(3;3). Thus, the altered function of the TF and/or TFR gene may play a role in the pathogenesis of this type of leukemia.

ONCOGENE ACTIVATION BY CHROMOSOMAL TRANSLOCATION IN CHRONIC MYELOCYTIC LEUKEMIA:

J. Groffen[1], K. Stam[1] and N. Heisterkamp[1], A. de Klein[2] and G. Grosveld[2] [1]Oncogene Science, Inc., 222 Station Plaza North, Mineola, New York 11501, [2]Erasmus University, Rotterdam, The Netherlands.

We have previously identified a DNA region on chromosome 22, designated bcr, which contains the chromosomal breakpoint of the Ph'-translocation in all Ph+ CML patients studied to date. Bcr is part of a gene, oriented with its 5' end towards the centromere of chromosome 22. Restriction enzyme and sequence analysis of bcr cDNA and genomic DNA demonstrates that coding regions of the bcr gene are spread over at least 66 kb of genomic DNA including bcr and exons 5' and 3' to it. As a consequence of the Ph' translocation, part of the bcr gene, the 3' end, has been translocated to chromosome 9, while the 5' part remains on the Ph' chromosome. The remaining bcr sequences act as an acceptor of a chromosome 9 gene, the c-abl oncogene: the c-abl oncogene is fused in a head-to-tail fashion to the chromosome 22 sequences. Such genomic configuration could result in the transcription of a novel chimeric mRNA consisting of 5' bcr sequences and 3' c-abl oncogene sequences. In RNA isolated from K562, a cell line derived from a CML patient, and in five CML patient RNAs we have demonstrated a chimeric bcr/c-abl transcript. Abnormally sized c-abl protein has been demonstrated in the cell line K562 and in leukemic cells from patients. Our results indicate that this protein represents the translational product of the chimeric mRNA. The role of the bcr part of the fusion protein remains unknown: computer searches for bcr homologous proteins shows no significant homology with other known proteins, precluding the assignment of a cellular function. However the bcr moiety could alter the structure of the c-abl protein and unmask its tyrosine kinase activity. In analogy to the gag/v-abl polyprotein, the CML specific bcr/c-abl protein might have transforming activity and play an essential role in the generation and/or maintenance of CML.

THE MYC-FAMILY OF CELLULAR ONCOGENES: Frederick W. Alt, Ronald DePinho, Edith LeGouy, Kathy Zimmerman, and Perry Nisen, Department of Biochemistry and Molecular Biophysics, College of Physicians & Surgeons of Columbia University, New York, New York.

The myc-family of cellular oncogenes contains three well-defined members: c-myc, N-myc, and L-myc; all lie on distinct human chromosomes. Our current results suggest that this gene family contains additional members. Activated expression of the c-myc gene, by a variety of different mechanisms, has been implicated in the development of many different types of human tumors. Activated expression of the N-myc and L-myc genes has been found to occur in a much more restricted set of human tumors including neuroblastomas, retinoblastomas, and small cell carcinomas of the lung. Activated expression of N-myc and L-myc has been shown to occur, to date, by gene amplification. We have demonstrated that the N-myc and c-myc genes have a similar oncogenic potential as assayed in vitro, that the two genes are quite similar in their overall organization, and that they share several anomalous structural features that are probably related to overlapping regulatory strategies. The L-myc gene appears related in organization. All three genes encode proteins that show considerable homologies in amino acid sequence.

Despite the similarities in structure and apparent function (i.e. transforming activity) of the myc-family genes, high level expression of the N-myc and L-myc genes is very restricted with respect to tissue and developmental stage while the c-myc gene is more generalized. These genes are differentially expressed during the progression of particular cell lineages, including B and T lymphoid lineages. The expression patterns of the genes in normal development helps to predict the types of tumors in which they are expressed or activated; in particular, activated N-myc expression may be a characteristic of a variety of tumors which derive from embryonic tissues. Our initial transfection experiments suggest that sequences which confer specific expression patterns to the N-myc gene are located within or near the gene. In addition, we find that high level expression of the N-myc gene may cross-regulate expression of the c-myc gene. Together, our findings suggest that differential, or perhaps combinatorial, myc-family gene expression may play a role in normal development.

STUDIES ON THE REGULATION AND FUNCTION OF THE HUMAN C-MYC GENE: P. Leder, J. Chung, T. Halazonetis, J. Sarid and E. Schmidt, Department of Genetics, Harvard Medical School, Boston, Massachusetts.

The human c-myc gene has been implicated in a number of human malignancies including Burkitt lymphoma, a malignancy in which this gene is translocated into one of the immunoglobulin loci. Detailed studies of these translocated c-myc oncogenes strongly suggest that they are activated by a disturbance of their regulated expression, although alterations of their coding potential occur as well. We have created an array of site-directed mutations in both the putative regulatory and coding portions of the human c-myc gene in an effort to identify regions essential for its normal function. In so doing we have identified several regulating segments near the c-myc gene that appear capable of interacting with trans-acting positive and negative regulatory elements. We also have been able to identify protein coding segments of the c-myc gene that may be uncoupled from the transforming function of the gene--a finding that suggests that they may also have a regulatory function. Finally, we have used the human c-myc gene under the control of a powerful immunoglobulin heavy chain enhancer to produce strains of transgenic mice that invariably develop pre-B and B-cell lymphomas. These results indicate that the otherwise normal human c-myc, when deregulated, can conribute to the development of malignancy in the context of a living mouse. Since the lymphomas are monoclonal, they offer further evidence that the c-myc gene must function with other genetic elements in order to cause malignancy.

ASSOCIATION OF THE MYC ONCOGENE PROTEIN WITH SNRNP'S: IS MYC INVOLVED IN RNA PROCESSING?

D.L. Spector[1], N.F. Sullivan[1] and R.A. Watt[2]. [1]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724 and [2]Smith Kline and French Laboratories, Swedeland, Pennsylvania 19479.

The immunolocalization and functional associations of the myc oncogene protein were examined using antibodies directed against bacterially produced human c-myc protein. The nuclear distribution of $p110^{gag\text{-}myc}$ in virally transformed quail cells (Q8) as well as the distribution of c-myc in a human colon carcinoma cell line (COLO 320) paralled the distribution of snRNP's as determined by double-label immunofluorescence with a monoclonal anti-Sm antibody. In addition, human c-myc protein microinjected into the cytoplasm of REF-52 cells migrated into the nucleus and localized at snRNP complexes. Immunoelectron microscopy revealed the immunoreactive regions to form a reticulum which occupied a portion of the nucleoplasm. While the immunostained regions often came into direct contact with the nucleolar surface they rarely were observed to be contiguous with the nuclear lamina and/or envelope. Nuclear extracts fractionated on sucrose density gradients provided a 40S RNP peak which exhibited immunoreactivity for both the Sm and c-myc antigens by immunoblotting. RNA labeled with [^{32}P]-orthophosphate from fractions of a similar sucrose gradient was immunoprecipitated with the anti-Sm and anti-c-myc antibodies. The anti-c-myc antibody immunoprecipitated a series of RNA's with identical electrophoretic mobilities as the U1,U2,U4,U5 and U6 snRNA's immunoprecipitated with the anti-Sm monoclonal antibody. These data show that c-myc and Sm-snRNPs are members of common nuclear macromolecular complexes and raise the possibility that they may be functionally related in one or more aspects of RNA processing.

PROTO-ONCOGENES FOS AND FMS. Inder M. Verma, Charles Van Beveren, Richard Mitchell, Wiebe Kruijer, Jacqueline Deschamps and Frits Meijlink, Molecular Biology and Virology Laboratory, The Salk Institute, San Diego, California 92138.

Proto-oncogene fos is a multifaceted gene, which is expressed during cell growth, cell differentiation and development. The viral homologue, v-fos, was identified as the resident transforming gene of FBJ-murine osteosarcoma virus which induces bone tumors in mice. Due to an in-frame deletion during the biogenesis of the v-fos gene, the products of viral and cellular fos proteins differ at their C-termini. The first 332 amino acids are nearly identical (5 amino acid changes), but the remaining 48 amino acids of the c-fos protein are different from the v-fos encoded product. Despite different C-termini, both fos proteins are nuclear in their location, and can transform fibroblasts in vitro. However, transformation by c-fos gene requires removal of a 67 base pair sequence from the 3' non-coding domain.

Proto-oncogene fos is expressed at higher levels in amnion and bone marrow (macrophages). It is a highly inducible gene in response to a variety of growth factors and differentiation-specific inducers. During differentiation of monocytic cells to macrophages, c-fos gene is rapidly induced with maximal levels accumulating by 30-60 min, which then decline by 4-5 fold and remain unchanged for the next 100 hr. However, c-fos protein is synthesized for only 90-120 min. In contrast, when quiescent fibroblasts are stimulated with platelet-derived growth factor or PC12 cells are induced by NGF to differentiate to neurites, c-fos expression is observed for only 30-60 min. The transcriptional enhancer and the inducible element of the fos gene have been mapped along with R sequences in the 3' non-coding domain which may influence the stability of the c-fos transcripts. The complex regulation of the c-fos gene will be discussed.

This work was supported by the National Institutes of Health and the American Cancer Society.

MARROW TRANSPLANTATION AND GENE THERAPY: E. Donnall Thomas, Fred Hutchinson Cancer Research Center, Seattle, Washington.

The bone marrow is a reasonable target for therapeutic gene transfer since the marrow is an organ that can be obtained readily, it contains pluripotent stem cells and it is affected by many genetically determined disorders. Bone marrow transplantation from a histocompatible donor is now accepted therapy for a variety of malignant and non-malignant diseases. Bone marrow transplantation for genetic disorders is best illustrated by the experience with B-thalassemia major. The first patient was transplanted more than four years ago. He is hematologically normal and growth and development have been normal illustrating that the disease can be cured by the transfer of normal stem cells. More than 100 such transplants have now been carried out in several transplant centers. Approximately 15% of the recipients have died of graft-versus-host disease, a problem that could be avoided by gene transfer into the patient's own stem cells. Despite intensive chemotherapy or irradiation designed to destroy the patient's abnormal marrow, approximately 10-15% of the patients have had their own marrow regenerate and again have thalassemia major, illustrating the difficulty of completely eradicating the patient's own abnormal marrow.

Current research in gene therapy addresses problems such as the efficiency of gene transfer, maintained gene expression, species differences and the co-transfer of a gene conferring a selective advantage. As it becomes practical to attempt to transfer normal genes into human marrow stem cells, the technology developed for bone marrow transplantation will play an essential role including destruction of the patient's abnormal marrow, procedures for procurement, in vitro manipulation and reinfusion of marrow, and the intensive supportive care of the patient during periods of marrow aplasia.

GENE TRANSFER INTO HEMATOPOIETIC STEM CELLS: IMPLICATIONS FOR GENE THERAPY:

A. Bernstein[1], J. Dick[1], C. Magli[2], and R.A. Phillips[2], [1]Mt. Sinai Hospital Research Institute, Toronto, [2]Hospital for Sick Children Research Institute, Toronto.

The blood-forming system of mice and men consists of a heterogenous population of cells, ranging from rare pluripotent stem cells with extensive proliferative and developmental potential to fully differentiated end cells. In an attempt to understand the molecular mechanisms that underlie the developmental decisions of stem cells, we have devised high efficiency protocols for the insertion of new genetic information into stem cells derived from either mice or humans. By using the random integration site of retrovirus vectors into these cells, we have monitored the developmental potential of primitive stem cell clones following bone marrow transplantation into genetically anemic W/W^v mice. As well, the ability of stem cells and their progency to express new genetic information was assayed by determining the proportions of colony-forming cells expressing a selectable drug-resistance gene. These studies have led to several observations with relevance to gene therapy in Homo sapiens: First, only a small number of clones (1-3) participate in the reconstitution of these W/W^v mice. Second, both myeloid and lymphoid lineages can be repopulated by the identical stem cell, providing direct evidence for gene transfer into a pluripotent myeloid-lymphoid stem cell. Third, the levels of expression of the transduced gene appears to decline as primitive stem cells differentiate into committed progenitor cells. Fourth, the protocols we have developed for gene insertion into mouse stem cells can be successfully applied to human bone marrow. Fifth, there appears, however to be marked differences in the levels of gene expression in human vs. murine stem cells with certain retrovirus vector constructs. (Supported by MRC and NCI of Canada).

GENE TRANSFER INTO MULTIPOTENT HEMOPOIETIC STEM CELLS:
E. Wagner[1], C. Paige[2] and G. Keller[2], [1]European Molecular Biology Laboratory, Heidelberg, Fed. Rep of Germany; [2]Basel Institute for Immunology, Basel, Switzerland

The transfer of foreign genes with retroviral vectors to precursors of the hemopoietic system provides a unique opportunity to analyze the expression and function of a variety of different genes in a well-characterized developmental system that can be manipulated both in vitro and in vivo. To develop a model system for high efficiency gene transfer, the first experiments involved infection of bone marrow cells with helper-free recombinant viruses that contained only the selectable neo gene[1]. Following virus infection, cells were either assayed in methyl cellulose cultures for G418 resistant colony-forming cells in vitro or injected into irradiated mice to determine if primitive precursors (stem cells) capable of long-term reconstitution had been infected. It was found that the neo gene was efficiently expressed in various lineages of the hemopoietic system in all reconstituted animals. Analysis of the viral integration sites from the DNA of different organs, as well as from B- and T-cell hybridomas and from factor-dependent mast cells derived from reconstituted animals indicated that multipotent stem cells capable of generating myeloid and lymphoid progeny had been infected with the recombinant virus[1].

Presently, bone marrow cells are being infected with recombinant viruses carrying growth control genes (viral oncogenes) in addition to the neo gene. The infected cells are being assayed again in vitro as well as following reconstitution of mice in vivo to study gene expression and the consequences of expression of these genes on the differentiation of the different precursor cells. These experiments should also establish the potential of the system for future gene therapy experiments.

[1]Keller, G., Paige, C., Gilboa, E. and Wagner, E.F. (1985) Nature 318 149-154

TARGETTED MODIFICATION OF HUMAN CHROMOSOMAL GENES:
R. G. Gregg and O. Smithies; Laboratory of Genetics, University of Wisconsin, Madison, WI 53706

Targetted modification of the human chromosomal β-globin locus has recently been demonstrated (Smithies et al., 1985). The specific modification was the insertion between the adult δ and β globin genes of plasmid-derived sequences as a result of homologous crossing-over between a 4.6 kb fragment of human DNA carried on the plasmid and its chromosomal equivalent. An assay for the desired event, based on "rescuing" predicted DNA fragments from the treated cells, showed that the ratio of targetted to non-targetted insertions was about 10^{-3}.

In order to attempt modification of another natural chromosomal locus by targetting, and also to permit a more rapid exploration of methods aimed at improving its frequency, we have initiated analogous experiments with the X-linked chromosomal locus coding for HPRT (hypoxanthine-phospho-ribosyl-transferase). This locus is particularly favourable because it normally occurs as a single active copy, and because it is possible to select both for (with HAT-medium) and against (with thioguanine) cells expressing the locus.

We have constructed a plasmid containing approximately 4.5 kb of the human HPRT gene including the third exon which codes for a non-integral number of amino acid residues. A single XhoI site in this exon permits the plasmid to be cut within the region of homology to provide recombinogenic ends. The plasmid also carries the Neo gene that confers resistance to G418 on mammalian cells. Successful targetting with this plasmid will inactivate the HPRT gene by adding a non-integral number of coding triplets to the transcript, and so will make modified cells resistant to thioguanine. At the same time the modified cells should become G418 resistant as a result of their acquisition of the Neo gene.

The cut plasmid was introduced into human cells by electroporation. A two-stage selection procedure was used to identify colonies in which loss of the HPRT function, leading to resistance to thioguanine, is associated with gain of the Neo function and resistance to G418. DNA from these doubly-resistant colonies will be analyzed by Southern blots to determine in what proportion of them targetted modification of this X-linked human chromosomal locus has been achieved.

HIGH FREQUENCY TARGETING OF GENES TO SPECIFIC SITES IN THE MAMMALIAN GENOME: Mario R. Capecchi and Kirk R. Thomas, Department of Biology, University of Utah, Salt Lake City, Utah 84112.

We corrected a defective gene residing in the chromosome of a mammalian cell by injecting into the nucleus copies of the same gene carrying a different mutation. We determined how the number, the arrangement, and the chromosomal position of the integrated gene, as well as the number of the injected molecules influence the gene-targeting frequency. Recombination between the newly introduced DNA and its chromosomal homolog occurred at a frequency of 1 in 10^3 cells receiving DNA. Correction events were mediated by either double reciprocal recombination or gene conversion. This resulted in sequences in the genome being replaced by sequences of the introduced DNA or, in separate experiments, sequences in the incoming DNA being replaced by chromosomal sequences. Both point mutations and deletion mutations were corrected; however, the nature of the mutation carried by the respect sequence influenced whether the integrated or injected sequence was corrected.

In addition to obtaining corrected genes via homologous recombination, we also obtained a frequent but unexpected class of corrected genes where the chromosomal gene still retained the original mutation but acquired a compensating mutation. We show that the introduction of the compensating mutation appears to result from the incorrect repair of a heteroduplex formed between the introduced sequence and the homologous genomic sequence.

MOLECULAR GENETICS OF DOWN'S SYNDROME: OVEREXPRESSION OF TRANSFECTED HUMAN Cu/Zn-SUPEROXIDE DISMUTASE GENE AND THE CONSEQUENT PHYSIOLOGICAL CHANGES. Y.Groner, O.Elroy-Stein, Y.Bernstein, N.Dafni, D.Levanon, E.Danciger, A.Neer. Dept. of Virology, Weizmann Inst. of Sci., Rehovot 76100, Israel.

The human Cu/Zn-superoxide dismutase (CuZnSOD) a key enzyme in the metabolism of oxygen free-radicals is encoded by a gene residing on chromosome 21 at the region 21q22 known to be involved in Down's syndrome (D.S.). Overexpression of the CuZnSOD gene, due to gene dosage, may disturb the steady state equilibrium of active oxygen species within the cell, resulting in oxidative damage to biologically important molecules. To investigate a possible involvement of CuZnSOD overproduction in the etiology of D.S. we first cloned its cDNA, then isolated the gene and studied its molecular structure and expression in normal individuals vs. D.S. patients. The gene, which is present as a single copy per haploid genome, spans 11 Kb of chromosomal DNA. It contains five rather small exons and four introns that interrupt the coding region. At the 5' end of the gene there are the 'TATA' and 'CAT' promoter sequences as well as four copies of the -GGCGGG- hexanucleotide. Two CuZnSOD mRNAs of 0.7 and 0.9 Kb were found in a variety of human cells. Both of them are transcribed from the unique CuZnSOD active gene and sequences at the 3'-untranslated region account for the size difference between them. Quantitative RNA analysis have indicated that the gene is overexpressed between 1.5 to 2-fold in a variety of tissues (brain, lung, heart) of aborted D.S. fetuses.

To examine the consequences of CuZnSOD overproduction, apart from dosage effects exerted by other genes residing on chromosome 21, the CuZnSOD gene and the cDNA were introduced into mouse L-cells and human HeLa cells respectively, as recombinant plasmids containing the neo^R selectable marker. Human and mouse transformants were obtained that expressed elevated levels (up to six-fold) of authentic, enzymatically active human CuZnSOD. Cell clones that overexpressed the transfected CuZnSOD had altered properties; they were more resistant to paraquat - an oxygen free-radical generator - than the parental cells and showed an increase in lipid peroxidation. Enhanced lipid peroxidation may alter the structure and function of cell membranes and thus contribute to some of the clinical symptoms associated with D.S. An "anti-sense SOD" plasmid was constructed and used to relieve the transfected cells from symptoms caused by the excess of CuZnSOD.

TRANSFER OF GENES INTO HUMAN SOMATIC CELLS USING RETROVIRUS VECTORS: A. Dusty Miller, Randy A. Hock and Theo D. Palmer, Department of Molecular Medicine, Fred Hutchinson Cancer Research Center, Seattle, Washington 98105

Insertion of genes into somatic cells of humans has been widely discussed as a possible treatment for genetic disease. Retrovirus vectors provide a highly efficient method for introduction of genes into cultured cells and animals and thus appear to be ideal vehicles for gene transfer into humans. We have generated high-titer retrovirus vectors carrying dominant selectable markers, and have used them to test our ability to infect human somatic cells.

The vectors were generated in the absence of helper virus using the previously described retrovirus packaging cell line PA12, and a newly developed line called PA317. PA12 cells were made using a helper virus from which the signal required for packaging viral RNA into virions was removed. In addition to this deletion, the helper virus used to make PA317 cells contained several other deletions in regions required in *cis* for virus replication. Although helper virus production can be detected following introduction of certain retroviral vectors into PA12 cells, we have not detected helper virus production from PA317 cells.

Using these vectors, we have been able to confer drug resistance to human bone marrow cells and normal human diploid skin fibroblasts *in vitro*. Both are good targets for gene therapy because of the ease of obtaining and reintroducing these tissues, and their long term persistence following engraftment. Following co-cultivation of mononuclear bone marrow cells with virus producing fibroblasts, 5-25% of the committed progenitor cells, including CFU-GM, BFU-E and CFU-MIX, were converted to drug resistance. Greater than 50% of human skin fibroblasts were converted to drug resistance following exposure to virus-containing medium. These experiments demonstrate the usefulness of retroviruses for gene transfer into human somatic cells.

RETROVIRUS-MEDIATED TRANSFER AND EXPRESSION OF HUMAN ADENOSINE DEAMINASE (ADA) SEQUENCES IN HEMATOPOIETIC CELLS: SH Orkin[1,2], DW Williams[1,2], B Lim[1], M Dexter[3], and RM Mulligan[4], Div of [1]Hematology, Children's Hospital and [2]Howard Hughes Medical Institute, Harvard Medical School, Boston MA; [3]Christie Hospital, Manchester, England; [4]Whitehead Institute, MIT, Cambridge, MA.

Retrovirus-mediated transfer of ADA sequences into bone marrow is a model for somatic gene therapy. We have examined ADA expression in cells, both in vitro and in vivo, following infection with recombinant retroviruses.

Initial work centered on a DHFR-vector into which ADA cDNA was introduced alone (Zip-DHFR-ADA) or with an SV40 promoter (Zip-DHFR-SVADA). Upon infection of NIH3T3 cells only the latter virus led to expression of substantial human ADA. Fur- ther studies employed Zip-DHFR-SVADA. Human ADA enzyme was readily detected following infection of murine pre-B and T cells and ADA-deficient human B-cells, indicating activity of the SV40 promoter in lymphoid cells in culture.

Lethally-irradiated mice were reconstituted with bone marrow cocultivated with Zip-DHFR-SVADA. 15% of CFU-S contained intact proviral DNA. No human ADA enzyme was detected. RNA analysis revealed variable, very low expression of RNA from LTR elements and none from the SV40 promoter. These findings domonstrate a dramatic difference in expression between tissue culture and primary hematopoietic cells.

Murine FDC cells, continuous stem cells that undergo differentiation in vitro, were infected with Zip-DHFR-SVADA. Expression of human ADA was abundant in undifferentiated cells and in all CFU-GM colonies, indicating that (1) FDC cells are beyond a stage in differentiation at which repression occurs or (2) other regulatory factors differ between them and primary hematopoietic cells.

Features of recombinant viruses that affect expression in primary hematopoietic cells are being examined in additional ADA-retroviruses.

MOLECULAR GENETICS OF THE HUMAN LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR: David W. Russell, Michael S. Brown, and Joseph L. Goldstein, Department of Molecular Genetics, University of Texas Health Science Center, 5323 Harry Hines Blvd., Dallas, TX 75235

The LDL receptor is the focal point of a metabolic pathway that supplies mammalian cells with cholesterol. Mutations in the gene for the receptor cause the genetic disease, familial hypercholesterolemia (FH). In FH, a reduction in the number of functioning cell surface LDL receptors results in increased plasma LDL-cholesterol levels, which in turn cause premature atherosclerosis and heart attacks. Approximately 1 in 500 individuals inherit a mutant allele at the LDL receptor locus. Biochemically, mutations in the LDL receptor protein have been classified into four distinct groups. Null alleles are the most prevalent and are defective in the synthesis of receptor protein. Transport-deficient alleles encode proteins blocked in transport to the cell surface. Proteins derived from binding-deficient alleles cannot bind LDL and those encoded by internalization-defective alleles cannot internalize bound ligand.

We have isolated cDNA and genomic clones derived from the human LDL receptor locus on the short arm of chromosome 19. The receptor gene spans 50 kb and is divided into 18 exons. Using these probes, we have cloned and sequenced at least one representative gene from each of the above mutation groups. A null allele has been shown to be caused by a 5 kb deletion that removes exon 12 through intervening sequence 15. Two mutations that block transport of the receptor are due to small deletions in the cysteine-rich ligand binding domain of the protein. This ligand binding domain is composed of 7 cysteine-rich repeat units of 40 amino acids each. Deletion of one repeat (exon 5) results in the loss of LDL binding. Three internalization- defective alleles have alterations in exons of the gene that encode the cytoplasmic domain of the protein.

A striking finding associated with all of these mutations is that they can be caused by recombination events involving Alu sequences. The mutations involve Alu repeats in either opposite orientation (intrastrand) or the same orientation (interstrand). Thus, repetitive DNAs in human genes can have dramatic consequences for genetic disease.

MOLECULAR GENETICS OF APOLIPOPROTEINS AND CORONARY HEART DISEASE: A.G. Motulsky, S. Deeb, A. Failor, D. Walker, J. Brunzell, J. Albers; Departments of Medicine and Genetics, University of Washington, Seattle

Hyperlipidemia is an important risk factor for coronary heart disease. Study of the molecular genetics of apolipoproteins may elucidate specific defects. DNA probes for apolipoproteins at the loci for AI-CIII-(AIV) on chromosome 11, AII on chromosome 1, CII-(CI-E) on chromosome 19, and B on chromosome 2 were studied in a) normal populations of various racial origins, b) a Caucasoid population with coronary heart disease (n > 100), and c) families with hyperlipidemia.

Restriction enzyme polymorphisms (RFLP) at a given apolipoprotein locus determine characteristic haplotypes. A minisatellite polymorphism with high information content was discovered at the apo B locus. Linkage equilibrium and disequilibrium of RFLPs at the AI-CIII locus were not related to physical distance in three major racial groups.

Associations of apolipoprotein RFLPs (and RFLP haplotypes) with coronary heart disease in comparison with control populations of matched ethnic origin will be presented and correlations with lipid parameters will be shown. Associations may be found with one RFLP marker but not with another RFLP at the same lipoprotein locus. Apolipoprotein haplotype and interaction of several apolipoprotein loci are likely to affect susceptibility to hyperlipidemia and coronary heart disease.

Linkage studies of DNA variants at the apo B locus in families with elevated apo B levels and familial combined hyperlipidemia are being carried out in order to search for a major cosegregating gene affecting apo B levels.

These strategies may lead to the identification of specific lipid related genes involved in coronary heart disease susceptibility. The conceptual approaches and their practical applications to coronary heart disease will be discussed.

CELLULAR MOLECULAR BIOLOGY OF LIPOPROTEIN METABOLISM: CHARACTERIZATION OF LIPOPROTEIN RECEPTOR-LIGAND INTERACTIONS. Robert W. Mahley, Gladstone Foundation Laboratories, Departments of Pathology and Medicine, University of California, San Francisco, California

Cholesterol homeostasis is dependent upon the action of specific lipoprotein receptors on various cells (especially within the liver) that mediate lipoprotein uptake and degradation. There are at least two distinct hepatic lipoprotein receptors: the apo-B,E(LDL) receptor and the apo-E receptor, which is postulated to be the chylomicron remnant receptor. They are independently regulated and display unique ligand specificity. In addition, the apo-B,E(LDL) and the apo-E receptors differ in apparent molecular weight (130,000 and 56,000, respectively). However, both receptors possess a region within their amino acid sequence that is enriched in acidic amino acids (glutamate and aspartate) and that is postulated to be the ligand binding domain. The tentative identification of acidic amino acids within the receptors as being essential in mediating ligand interactions is based primarily on data showing that basic amino acids (lysine and arginine) of the ligands (apo-E and apo-B) mediate receptor binding. The receptor binding domain of apo-E has been localized to basic amino acids in the vicinity of residues 140 to 160. Naturally occurring variants of apo-E with single amino acid substitutions in this region of the molecule are defective in their abilities to bind to the lipoprotein receptors and predispose affected individuals to develop type III hyperlipoproteinemia. Recently, it has been possible to more clearly define essential residues within this region of the apo-E molecule by using site-specific mutagenesis. A natural extension of the studies defining the receptor binding domain of apo-E has been the identification of two basic regions of apo-B that appear to mediate the interaction of this apoprotein with lipoprotein receptors. Eventually, it will be possible to develop a model describing ligand-receptor interactions; presumably, these occur through the interaction of key acidic amino acids of the receptors with the key basic amino acids of apo-B and apo-E.

THE GENES FOR HUMAN INSULIN AND ITS RECEPTOR: SELF SPECIFIC SYNTHESIS AND TRANSMEMBRANE SIGNALLING: W. Rutter, University of California, San Francisco

IDENTIFICATION OF DNA REPAIR GENES IN THE HUMAN GENOME.

D. Bootsma, M. van Duin, J. Hoeijmakers, A. Westerveld and A. Yasui. Department of Cell Biology and Genetics, Erasmus University, Rotterdam.

To identify human DNA repair genes we have transfected human genomic DNA ligated to a dominant marker to excision repair deficient xeroderma pigmentosum (XP) and CHO cells. This resulted in the cloning of a human gene (ERCC-1) that complements the defect of a UV- and mitomycin-C sensitive CHO mutant 43-3B. Extensive efforts to clone XP-correcting genes using the same approach have failed so far. The ERCC-1 gene has a size of 15 kb, consists of 10 exons and is located in the region 19q13.2-q13.3. Its primary transcript is processed into two mRNAs (1.1 and 1.0 kb) by alternative splicing of an internal coding exon. These transcripts encode largely identical polypeptides of 297 and 273 aminoacids (AA) respectively. Only the cDNA of the 1.1 kb messenger is able to correct the defect in 43-3B cells. A putative DNA binding protein domain and nuclear location signal could be identified.

Significant AA-homology is found between ERCC-1 and the yeast excision repair gene RAD10. Experiments are in progress to identify ERCC-1 homologous sequences in Drosophila.

Transfection of the gene into representative UV-sensitive cell lines of the 5 known CHO complementation groups revealed that only the defect in complementation group 2 could be corrected. Introduction in cell lines of different XP complementation groups (XP-A, F and G) did not result in correction of the defects in these cell lines. These results suggest that ERCC-1 is probably not involved in the genetic defects of these XP complementation groups. Southern blot analysis did not reveal gross rearrangements or (partial) deletions of ERCC-1 in cell strains of the 6 other XP complementation groups nor in 3 Fanconi's anemia cell lines representing at least 2 complementation groups.

MAPPING OF GENES INVOLVED IN GROWTH CONTROL: U. Francke, Department of Human Genetics, Yale University School of Medicine, New Haven, Connecticut.

On one hand, an increasing number of genes are being cloned whose products are involved in normal or malignant cell growth; and on the other hand, numerous non-random chromosomal aberrations, translocations and deletions, are being identified in human tumor cells. In an attempt to predict which growth factors or protooncogene might play a role in which type of cancer, we are mapping the chromosomal location of these genes and are correlating the results with cytogenetic abnormalities reported in malignant tissue.

My laboratory has determined the chromosomal sites of genes for (1) growth factors EGF, NGFβ, IGFI, IGFII, TGFα, TGFβ, for (2) growth factor receptors of the tyrosine kinase family INSR, EGFR, EGFR variant or NEU, PDGFR and for (3) the serine/threonine-specific Ca/phospholipid-dependent protein kinase type α, β and γ. The probes were mostly cDNAs of various mammalian origins cloned in the laboratories of A. Ullrich or R. Derynck (Genentech) and R. Weinberg (Whitehead Institute).

Mapping was done by two independent approaches: Southern blot analysis of rodent x human somatic cell hybrid lines with a subset of normal and rearranged human chromosomes and in situ hybridization of ^{3}H-labeled probes to prometaphase chromosomes. The combined results provide subregional gene assignments, often to a single chromosome band. In several instances the homologous genes were also mapped in the mouse using Southern analysis of Chinese hamster x mouse somatic cell hybrids. The mapping data will be reviewed and a number of possible correlations with human tumors will be proposed. Comparison of mosue and human mapping data adds to the delineation of conserved chromosome regions.

THE CHROMOSOME 11 GENE MAP; GENES AND MARKERS IN GROWTH, DEVELOPMENT, AND CANCER: Thomas B. Shows, Department of Human Genetics, Roswell Park Memorial Institute, Buffalo, New York.

Chromosome 11 is second only to the X chromosome in the number of gene markers assigned to a chromosome. On this chromosome, there are currently four major locations associated with abnormal growth, development, and cancer. They have been identified at p15 for the Beckwith-Wiedemann syndrome and familial renal carcinoma; p13 for embryonal Wilms' tumor of the kidney, associated with the aniridia, urogenital, mental retardation syndrome (WAGR) involving overlapping deletions; q13 for cancer chromosome translocations associated with different leukemias and a fragile site; and at q23 for additional cancer chromosome translocations associated with developmentally earlier forms of leukemias than those at q13 and a fragile site. We have mapped such control genes as insulin, insulin-like growth factor-2 (IGF2), parathyroid hormone, and the HRAS1 oncogene at p15. These assignments have been reported by others also. At the p13 WAGR common deletion region, we have mapped genes for follicle stimulating hormone-B, a hepatitis B virus insertion site, several arbitrary DNA segments, and the previously mapped catalase gene. Because of the clinical phenotype and chromosomal rearrangement patterns, genes for development of the kidney, iris, and urogenital tract must also be located at 11p13. The tumor types, reduction to homozygosity of 11p15 alleles, and expression of IGF2 in certain childhood tumors suggest a pathogenetic developmental mechanism involving genes in p15 and p13. At q13 we have mapped the pepsinogen gene complex and an arbitrary DNA segment; and at q23 the ETS1 oncogene and a gene probe have been mapped by others. These markers, and DNA sequences now being isolated from chromosome 11 libraries and from the 11p regions associated with WAGR deletions, are being used to characterize these growth disorders and, particularly, the WAGR deletions. Mapping these markers has stimulated an extensive fine structure physical map of chromosome 11. The markers being isolated show promise for identifying chromosome breakpoint junctions, mutant genes, genes in development, and in the prognosis and diagnosis of disease.

MOLECULAR GENETICS OF HUMAN FAMILIAL CANCER: W. Cavenee[1,2], M. Hansen[1], H. Scrable[2] and A. Koufos[2], [1]Ludwig Institute for Cancer Research, McGill University, Montreal, Canada; [2]Department of Microbiology and Molecular Genetics, University of Cincinnati College of Medicine, Cincinnati, Ohio.

Mutant alleles of the RB1 locus at 13q14 have previously been linked to the development of retinoblastoma, a tumor of embryonal neural retina. In the bilateral form of the disease, inheritance of a germinal mutation at this locus predisposes each retinal cell to a subsequent second somatic event which results in the formation of the retinoblastoma tumor. Spontaneous unilateral tumor occurrence appears to result from two sequential somatic events. Survivors of the bilateral form of this eye tumor are at substantially increased risk for the subsequent development of second primary cancers, particularly osteosarcoma, a bone cancer. Conversely, survivors of unilateral retnoblastoma exhibit the same likelihood of development of osteosarcoma as the general population.

In order to determine whether osteosarcoma and retinoblastoma share a common pathogenetic mechanism, we determined restriction fragment length alleles at loci on chromosome 13 in DNA from constitutional tissues of bilaterally affected retinoblastoma patients as well as from their osteosarcoma tumors. These analyses indicated that the loss of constitutional heterozygosity in the tumors occurred specifically for chromosome 13 and appeared to involve the same chromosomal region as that previously identified in retinoblastoma tumors. Similarly, specific loss of constitutional chromosome 13 heterozygosity was also apparent in sporadic osteosarcomas. These results suggest that both diseases are due to the mitotic unmasking of pleiotropic recessive mutant alleles at the RB1 locus, and that the clinical occurrence of mixed cancer families such as these may be due to the differential expression of a single recessive mutation.

These findings have led to the development of accurate premorbid predictors for this disease. They have also provided a theoretical basis for defining the number and genomic location of genes whose recessive mutant forms predispose to cancer. Our examination of several other tumor types indicates that there are a limited number of this type of gene and that each has multiple tissue specificities. Advantage has been taken of this to provide for accurate discrimination of tumors of questionable histopathology.

GENETIC ANALYSIS OF THE WILMS' TUMOR-ANIRIDIA REGION OF HUMAN CHROMOSOME 11p

D. Housman[1], T. Glaser[1], D.S. Gerhard[1], G. Bruns[2] and C. Jones[3]

[1]Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA
[2]Genetics Division, Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, MA
[3]Eleanor Roosevelt Institute for Cancer Research, University of Colorado Medical Center, Denver, CO

The fine structure of human chromosome 11p has been analyzed by several complementary techniques involving somatic cell genetics and meiotic linkage. A focus of this analysis has been the identification and characterization of DNA sequences that are altered in Wilms' tumor and aniridia. Patients with heterozygous deletions involving a segment of band 11p13 exhibit a set of disorders (known as the WAGR complex) which may include aniridia (the absence of a normal iris), genito-urinary malformations, mental retardation and bilateral Wilms' tumor. This tumor, an embryonal malignancy of the kidney, apparently arises via the homozygous loss of function of a recessive oncogene in a target nephroblast. We have identified a series of DNA segments and surface antigen genes that lie within the WAGR region through the development of somatic cell hybrids. These cell lines carry (1) chromosome 11s with WAGR deletions and (2) translocation chromosomes with breakpoints interrupting this region. Using six hybrid cell series of this type, we have identified four DNA segments that are deleted in one or more WAGR patients. Our results suggest the following organization of the WAGR region:

CEN - D11S9 - CAT - (WAGR,FSHB) - HBVS1 - TELO.

Two surface antigen genes (MIC1 and MIC4), identified by monoclonal antibodies, are also deleted in WAGR patients. Like FSHB, the MIC1 marker maps within the smallest interval that includes the Wilms tumor and aniridia loci. Independent data, obtained from 11p deletions that were selected in vitro, suggest that MIC1 lies between CAT and FSHB. The position of the Wilms' tumor and aniridia genes in relation to these close genetic markers--FSHB, CAT and MIC1--will be discussed.

MOLECULAR AND CELLULAR BIOLOGY OF HUMAN LUNG CANCER:

J.Minna, J.Battey, B. Brooks, J. Carmichael, F. Cuttitta, J. DeGreves, J. Gu, D. Ihde, A.M. Lebacq-Verheyden, I. Linnoila, B. Johnson, J. Mulshine, M. Nau, E. Sausville, E. Seifter, M. Vinocour, A. Gazdar: NCI-Navy Medical Oncology Branch, National Cancer Institute and Naval Hospital, NIH, Bethesda, Maryland 20814

New methods to prevent, diagnose, stage, and treat lung cancer are needed. With these goals in mind we have been studying the molecular and cellular biology of human lung cancer. We have found: 1) expression of neuroendocrine and some hematopoietic markers in small cell lung cancer (SCLC); 2) the existence of a common endodermal derived stem cell for all lung cancer; 3) the production of several peptide hormones by SCLC and at least one of these, gastrin releasing peptide (GRP, also called bombesin) can act in an autocrine fashion to stimulate tumor growth. Thus, manipulation of GRP or its receptor represents a new way to treat lung cancer; 4) the GRP prohormone has associated peptides and these peptides can exist in several different forms determined by alternative splicing of the GRP mRNA; 5) the tumor cell lines exhibit in vitro drug sensitivity and resistance that mirror the clinical behavior in the patient. Thus, the tumor cell lines are being used in clinical trials to search for new drugs and therapy for individual patients as well as a source for isolating drug resistance genes; 6) a specific chromosomal deletion (del 3p14-23)is found in SCLC suggesting the presence of genetic defects similar to those in retinoblastoma and Wilm's tumor; 7) cellular oncogenes are frequently deregulated and expressed to high levels in lung cancer cells including those for c-myc, N-myc, L-myc, p53, and the ras family. A common method of deregulation is gene amplification and there appears to be great specifity of expression of oncogenes in individual patient's tumors. High levels of expression of c-myc by SCLC significantly influences the growth phenoof the cancer cells, and is associated with decreased patient survival. These findings suggest the possibility of classifying tumors by their oncogene expression pattern and then designing and directing therapy against these products, as well as focusing the search for inherited predisposition to malignancy among these genes.

HUMAN ONCOGENES INVOLVED IN THE GROWTH FACTOR ACTIVATED PATHWAYS OF CELLULAR PROLIFERATION: Stuart A. Aaronson, Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland

Proto-oncogenes that have given rise to the oncogenes of acute transforming retroviruses have been implicated as frequent targets for genetic alterations leading to neoplasia. The normal functions of some proto-oncogenes are becoming better understood. The v-sis oncogene encodes a protein closely related to human platelet-derived growth factor (PDGF), a potent connective tissue mitogen. V-sis transforming activity is specific for cell types that possess PDGF receptors, implying that its actions are mediated directly by interaction with the PDGF receptor. The human c-sis/PDGF-2 coding sequence can also be activated as a transforming gene by expression in cells possessing PDGF receptors, indicating that activation of normal c-sis/PDGF-2 expression in human tumors of connective tissue origin may contribute to their malignant properties. We have identified regulatory sequences of the c-sis/PDGF-2 proto-oncogene as well as domains within its encoded product that are important for its actions. Evidence will also be presented concerning functions of other human proto-oncogenes which act in pathways by which growth factors induce cell proliferation.

THE MOLECULAR BASIS OF AN INHERITED IMMUNODEFICIENCY DISEASE AND OF ITS POTENTIAL CURE:

D. W. Martin, Jr., Steven R. Williams, and Scott McIver,
Genentech, Inc., South San Francisco, California

The inherited deficiency of purine nucleoside phosphorylase (PNP) results in a severe immunodeficiency characterized by a lack of thymocytes and mature T cells. The ubiquitous PNP consists of a homotrimer of 30,000 d subunits that catalyze the phosphorolysis of Ino, dIno, Guo and dGuo. The human cDNA for PNP has been cloned by expression in PNP deficient E. coli, and the human genomic clones have been isolated from a lambda library. The gene consists of six exons within 13 kbp. The first PNP deficient patient described was the product of a consanguineous mating, and her genomic PNP gene has been cloned. The sequencing of all exons and exon/intron junctions revealed only one nucleotide change. That G to A transition results in codon 89 being mutated from the wild type Glu to mutant Lys. This change explains the previously described peptide maps of the mutant protein isolated from the rbc of the heterozygous parents.

Synthetic 19-mer probes containing the mutant or wild type sequence were used to demonstrate that the patient was homozygous for the G to A transition and that her parents were both heterozygous. In vitro mutagenesis was used to revert the mutant A in codon 89 to the wild type G, and the wild type and reverted genomic clones were expressed in mouse cells. The expressed mutant gene produced catalytically inactive, immunoreactive PNP with the predicted abnormally alkaline PI. The reverted genomic gene produced wild type catalytic and immunoreactive activities.

The structure of wild type human PNP has been solved by Ealick et al. Amino acid 89 is adjacent to the active site.

Retroviral vectors have been constructed which are capable of infecting murine cells in vivo and in vitro and expressing human PNP catalytic activity in vitro.

EXPRESSION AND FUNCTION OF SUPPRESSOR tRNA GENES IN MAMMALIAN CELLS: Y.W. Kan[1], Y.S. Ho[1], P. Palese[2], G. Norton[2], M.C. Cheung[1], J. Chang[1], [1]Howard Hughes Medical Institute and Department of Medicine, University of California, San Francisco, California; [2]Department of Microbiology, Mt. Sinai Medical Center, New York, New York.

Nonsense mutations are one of the many diverse genetic lesions causing β thalassemia. We have explored the use of suppressor tRNAs to correct the defect in β-globin synthesis in this class of disorders. Previously, two suppressor tRNA genes were constructed by site-specific mutagenesis. When injected into Xenopus oocyte nuclei, these tRNA genes directed synthesis of tRNAs that suppressed the amber codon in mRNAs derived from thalassemic patients who have the β^{17}AAG(Lys)$\rightarrow$UAG and β^{39}CAG(Glu)$\rightarrow$UAG mutations. We have now tested the properties of these tRNA genes in mammalian cells.

The tRNA genes were cloned into the late region of SV40, transfected into CV1 cells, and the level of suppression assayed by superinfection with an influenza virus. A nonstructural protein (NS-2) which terminates with a UAG codon is elongated if suppression occurs. With this assay, we observed suppression of up to 50%. To study the effects of stable transformants on mammalian cells, we cloned the tRNA gene into a vector containing a DHFR cDNA, and used it to transfect CHO DHFR$^-$ cells. The cells were cultured in the presence of increasing doses of methotrexate, and suppression assayed by infecting the resistant cells with influenza virus. Only up to 5% suppression occurred, even in cell lines in which the tRNA genes were amplified to 500 fold and abundant tRNAs were synthesized. The low level of suppression observed was caused by the lack of charging of the suppressor tRNAs in the cell lines containing high copy number of the suppressor tRNA gene. The ability of suppressor tRNA genes to suppress the nonsense mutation in bone marrow cells is currently being investigated by introducing the tRNA gene in a retrovirus vector. The vectors will be used to infect hybrid MEL cells containing a human chromosome 11 derived from a patient with the β^{39} nonsense mutation. Production of human β-globin would indicate the ability of the suppressor tRNA to function in an erythroid cell environment.

CORRECTION OF THE DEFECT OF GAUCHER DISEASE FIBROBLASTS AND LYMPHOBLASTS BY RETROVIRAL MEDIATED GENE TRANSFER:
J. Sorge and E. Beutler, Scripps Clinic and Research Foundation, La Jolla, California.

Patients with Gaucher disease are deficient in the enzyme glucocerebrosidase. We have isolated a complete cDNA clone of the human glucocerebrosidase gene. This cDNA was inserted into a retrovirus vector, using the virus LTR as a promoter, and the recombinant virus was used to infect type I Gaucher fibroblasts. The recombinant virus completely corrected the enzymatic defect, whereas a virus lacking the glucocerebrosidase gene did not affect enzyme activity. Full enzymatic activity was present in the cells infected with the recombinant virus after nine cell passages. The presence of virus did not affect cell growth or morphology. EBV-transformed lymphoblasts, from a type I Gaucher patient were superinfected with the glucocerebrosidase retrovirus. Again glucocerebrosidase enzymatic activity was fully restored.

The stability of the retroviral construction is fairly satisfactory if the number of viral passages is kept low. Using replication-defective helper cell lines, viral passage is kept down to one cycle. Under these conditions, genomic rearrangement has not been a problem.

GENE TRANSFER AND EXPRESSION BY A TRANSCRIPTIONALLY INACTIVATED RETROVIRAL VECTOR CONTAINING AN INTERNAL PROMOTER: J.-K. Yee[1], D. Jolly[2], J. Moores[1], J. Respess[1] and T. Friedmann[1], [1]Department of Pediatrics, University of California at San Diego, School of Medicine, La Jolla, California; [2]Lab Hormones Inserm, Paris, France.

We have constructed an amphotropic, MLV-based retroviral vector carrying the full length cDNA for human hypoxanthine guanine phosphoribosyltransferase (HPRT) and fully disabled LTR transcriptional functions. The enhancer sequences and CAT and TATA boxes of the 3' LTR of the cloned parent vector pLPL were removed by site-directed mutagenesis. The splice donor site was also deleted to reduce the possibility of aberrant splicing and the human metallothionein (MT-II) promoter was introduced in front of the HPRT cDNA. Transmissible virus was produced by transfection of the modified vector plasmid into LNPA1 cells constitutively expressing the packaging helper provirus pPAM. HAT-resistant colonies were isolated from infected rat HPRT-deficient cells, and the newly expressed HPRT activity was found to be regulated by heavy metals and dexamethasone. Infection of HAT-resistant cells with helper viruses failed to produce infectious HPRT virus, indicating that all HPRT activity is expressed from the internal metallothionein promoter. Such a disabled vector may help to reduce the likelihood of inadvertant and unwanted changes in the expression of cellular genes close to an integrated provirus.

RETROVIRAL-MEDIATED GENE TRANSFER OF HUMAN HPRT AND ADA:
C. Thomas Caskey, Steve Chang, John Belmont, Jenny Henkel-Tigges, Karen Wager-Smith; Institute for Molecular Genetics and Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas.

Defective ecotropic and amphotropic retroviral vectors containing the human HPRT and adenosine deaminase (ADA) cDNA's have been developed for high efficiency gene transfer and high level cellular expression. The production of high titer helper cell producer cell lines was achieved by a simplified method which minimizes cell culture effort. Initial titers (10^3-$5x10^4$) of defective SVX HPRT B in ψ2 cells, a vector containing both HPRT and neomycin, were increased 3-10 fold by cocultivation with the amphotrophic helper cells, PA-12. Transfection or infection by MoLV of 3T3 cells which carried a single copy of SVX HPRT B allowed development of higher viral titer ($8x10^5$-$7.5x10^6$) stocks. Using SVX HPRT B defective virus we have partially corrected (4-56% of normal) the HPRT deficiency of cultured cells derived from rodents and Lesch-Nyhan patients. Instability of HPRT expression was detected in some clonal isolates. The instability is of two general types: 1) retention of an apparently unaltered SVX HPRT B sequence and 2) loss of the SVX HPRT B sequence. In the circumstance of the first type, cells which lost HPRT express ($6TG^r$) also lost $G418^r$ indicating a cis-acting down regulation of expression.

A related retroviral shuttle vector, pZIP NEO-SV(B), has been used for gene transfer of human ADA. Transfected and infected ψ2 and PA12 cell lines were selected in 11AAU plus deoxycoformycin (dCF) (Alanosine, Adenosine, and Uridine) media. This selection system stringently requires high expression of ADA. A cell line 4.2T, was isolated that produces moderate titers of defective retrovirus (~ $2x10^4$ cfu-Neo^R/ml). This retroviruses have been used to transfer the gene for human ADA to mouse bone marrow cells. Transfer and expression of Neo and ADA in murine CFU's was demonstrated by in vitro colony formation in the presence of G418 and XylA plus dCF respectively. Isoenzyme analysis also showed human ADA expression in mouse bone marrow. Injection of infected marrow cells into lethally irradiated recipients demonstrated expression of human ADA in spleen colonies.

GENE TRANSFER AND EXPRESSION IN NON-HUMAN PRIMATES USING RETROVIRAL VECTORS: W. French Anderson[1], Philip Kantoff[1], Martin Eglitis[1], Jeanne McLachlin[1], Evelyn Karson[1], James Zwiebel[1], Arthur Nienhuis[1], Stefan Karlsson[1], R. Michael Blaese[1], Donald Kohn[1], Eli Gilboa[2], Donna Armentano[2], Alfred Gillio[3], Claudio Bordignon[3], and Richard O'Reilly[3]. [1]National Institutes of Health, Bethesda, MD 20892; [2]Princeton University, Princeton, NJ 08544; [3]Memorial Sloan-Kettering Cancer Center, NY, NY 10021.

The correction of an inborn error of metabolism by the insertion of a functioning gene into a patient's defective cells is a potential mode of therapy which is now approaching realization. A likely initial disease candidate for human gene therapy is adenosine deaminase (ADA) deficiency, a cause of severe combined immunodeficiency (SCID). A retroviral vector, called SAX, containing the human cDNA for ADA as well as a selectable gene, Neo^R phosphotransferase (NPT), has been constructed and shown to be effective in correcting the metabolic defect in human ADA(-) T and B cell lines treated in vitro. Using an autologous bone marrow transplant (BMT) protocol with non-human primates, the ADA gene was introduced into bone marrow (BM) cells via retroviral infection with the SAX vector. The donor animals were then lethally irradiated and their vector-treated cells reinfused. Peripheral blood and BM cells were analyzed starting 3-6 weeks later for the presence of SAX vector DNA, human ADA activity, and NPT activity. Approximately 0.1-1% of the BM cells in one animal 6 weeks after transplant contained vector DNA which restriction digests showed to be intact. The lysate of Ficoll-separated mononuclear blood cells was fractionated on an FPLC Mono-Q column using a KCl gradient. Human ADA, primate ADA, and NPT elute in different regions. Low levels of activity in the human ADA and NPT regions were detected in several animals. In one animal, using the most recent protocol, activity for putative human ADA was found at 69d and 104d post-transplant. The highest level was 20-fold higher than previously seen (now reaching approximately 0.5% the endogenous monkey activity). Both monkey and human BM cells are being studied by the CFU-GM assay using the SAX vector in order to determine optimal conditions for retroviral vector infection. Our results suggest that this autologous BMT/gene transfer protocol may be applicable as a model for human gene therapy.

THE HUMAN ARGININOSUCCINATE SYNTHETASE (AS) LOCUS: GENE REGULATION AND POTENTIAL FOR GENE THERAPY: A.L. Beaudet[1,2], M.J. Jackson[1], P.A. Wood[1], and W.E. O'Brien[1,2]; [1]Department of Pediatrics, Baylor College of Medicine and [2]Howard Hughes Medical Institute, Houston, Texas.

Genetic abnormalities in the structural gene for AS cause citrullinemia, which is a disorder of the urea cycle in humans. Two types of gene regulation can be studied in cultured human cells. First, expression of AS is repressed by arginine in wild type cells; second, AS is overproduced in canavanine resistant (Can^r) cell variants without gene amplification. Expression of AS in Can^r cells is not repressed by arginine. Results using a minigene linking the 5' end of the AS gene to the chloramphenicol acetyltransferase (CAT) gene indicate that 1) the construction contains promoter activity and is regulated by arginine in wild type cells; 2) the CAT minigene is not overexpressed in Can^r cells; 3) the CAT minigene is regulated by arginine in Can^r cells, although the endogenous AS gene is not similarly regulated in these cells; and 4) Can^r cells are in the repressed state with regard to arginine regulation.

In order to explore the potential for gene therapy for citrullinemia, the cDNA for AS was introduced into various retroviral expression vectors with and without the selectable neo marker. Recombinant AS-retrovirus transfers AS activity to AS^- rodent cells, and ecotropic viral titers of $>10^5$ CFU/ml were obtained using citrulline (arginine-free) medium or G-418 medium to select for expression of AS or neo respectively. A rapid assay for detection of AS-retrovirus was developed based on incorporation of [^{14}C]citrulline into protein after virus infection of AS^- 3T3 or XC cells. Introduction of AS activity into bone marrow cells or into hepatocytes might benefit patients with citrullinemia.

GENE TRANSFER: A POTENTIAL NEW THERAPY FOR THE TREATMENT OF GENETIC DISEASE. Richard C. Mulligan, Whitehead Institute for Biomedical Research and Department of Biology, M.I.T., Cambridge, Mass.

For several years, our laboratory has been interested in determining the feasibility of using gene transfer methods to treat selected disorders of the hematopoietic system in man. Most of our work in this direction has focused on determining whether retrovirus vectors can be used to transfer new genetic information into hematopoietic cells of the mouse. Specifically, we have asked whether the totipotent hematopoietic stems present in freshly explanted bone marrow can be infected by recombinant retroviruses in vitro, and subsequently used to permanently and completely engraft lethally irradiated recipients via bone marrow transplantation.

In order to characterize the developmental potential and general properties of the transduced hematopoietic stem cells, we have used retroviral-mediated gene transfer to clonally mark stem cells in vitro, and have tracked the fate of those cells after their introduction into lethally irradiated recipients via bone marrow transplantation. In order to efficiently transduce hematopoietic stem cells, a protocol was developed that involved the co-cultivation of bone marrow cells from 5-Fluoro-uracil (5FU) treated mice with virus producing cell lines in the presence of a source of IL-3. This protocol resulted in the transduction of 50-100% of both day 14 CFU-S and stem cells with long term reconstitution capacity, regardless of the titer of viral stocks used. To assess the contribution of individual transduced stem cell clones to hematopoiesis in completely and permanently engrafted transplant recipients, hematopoietic cells from the recipients were fractionated into pure cell populations representative of a particular lineage or anatomical location, and DNA isolated from these cell populations was analyzed in a quantitative Southern blot assay. Classes of stem cells with totipotent or apparently restricted developmental potential have been defined based on detection of specific proviral integrants in the various fractionated cell populations. These include cells whose progeny repopulate all lineages and anatomical locations as well as stem cells that appear

to contribute only to certain lineages, subsets of lineages or to specific organs and anatomical locations. We further describe a large class of stem cells which although totipotent displays quantitative differences in degree of contribution to various cell populations. The predominance of this class of cells suggests that apparent developmentally restricted stem cells may represent extreme examples of unequal lineage or organ repopulation by a totipotent cell.

The Southern blot analyses also indicated that surprisingly few (one or two) clones of transduced stem cells often account for the majority of mature hematopoietic cells in a recipient. Coupled with the finding that a substantial number of transduced cells that do not actively contribute to hematopoiesis in a primary recipient, can be activated to do so by retransplantation, the results suggest that a mechanism exists in vivo for the control of stem cell utilization. Studies involving the periodic sampling of hematopoietic tissue from a primary transplant recipient suggest that significant changes in the utilization of different stem cell clones can naturally occur. These studies underscore the dynamic nature of hematopoiesis, and may suggest that normal hematopoiesis results from the sequential activation of different stem cell clones, rather than from an averaged contribution of the entire stem cell pool.

We have also undertaken more practical experiments to determine the feasibility of 'genetic therapies' for two specific genetic diseases: severe combined immunodeficiency (due to ADA deficiency) and β-thalassemia. A major focus of these latter studies has been to examine the expression of various constructs carrying the ADA or β-globin gene. For expression of the ADA gene, a number of recombinant genomes have been constructed (In collaboration with Dr. Stuart Orkin's laboratory at Harvard) which utilize various viral and cellular promoters to drive expression of a human ADA cDNA. Experiments with the β-globin gene have involved the generation and characterization of recombinant genomes encoding a complete intact human β-globin gene. Data will be presented regarding the expression of these recombinant genomes after their introduction into mice via infection of bone marrow cells in vitro and subsequent long term reconstitution of transplant recipients.