Molecular approaches to dysmorphology

ALASDAIR IVENS, GUDRUN MOORE, AND ROBERT WILLIAMSON

From the Molecular Dysmorphology Group, Department of Biochemistry and Molecular Genetics, St Mary's Hospital Medical School, London W2 1PG.

SUMMARY The biochemical and physiological defects underlying human dysmorphic syndromes can now be approached using techniques of molecular biology. The genetic component of the causation of the dysmorphology can be studied in isolation from the environmental component by using large, rare families which exhibit the same phenotype as more complex multifactorial disorders, but inherit the mutation in a monogenic fashion. Such an analysis starts with the determination of linkage to a gene probe, followed by the use of newer techniques of molecular biology to enable cloning and sequencing of the mutated gene.

Analysis of the gene product by amino acid sequence homology to other known proteins, and tissue specific expression, may place the defect within the cascade of events associated with development and differentiation. Once cloned, the gene can also be manipulated in transgenic laboratory animals and the effect of its mutation studied directly. The use of techniques of molecular biology to study the genetic aspects of dysmorphic syndromes will allow insight to be gained both into normal fetal development and into the causes of congenital malformations.

The human genome has approximately 10^9 base pairs of DNA in 22 pairs of autosomes and the sex chromosomes. In addition to the several tens of thousands of genes which code for proteins, this complex of information also contains the instructions required for differentiation and development. It is only with the advent of recombinant DNA technology that we have been able to study the structure and function of the genes that control these processes at the molecular level. As a result of random mutations in the human genome, changes in the pattern of cleavage of restriction enzymes occur, which provide specific polymorphic markers, inherited in a Mendelian manner. The analysis of the patterns of inheritance of these markers in family studies provides a tool with which to map any gene and to study its cosegregation with any normal or abnormal variant phenotype.

It is now just over 10 years since the first mammalian gene, rabbit globin, was cloned. Since then, the advance in molecular genetics has been even greater than predicted by the most optimistic, with over 3500 markers identified. The mapping of the Duchenne muscular dystrophy locus provides a good example of the power of 'reverse genetics' in the analysis of gene disorders. The primary defect involved in Duchenne muscular dystrophy was first localised by linkage to a region of Xp and by a

combination of recombination studies and use of deletions and translocations for mapping and gene isolation, the DMD gene has now been characterised.^{1,2} The protein product of the gene, 'dystrophin', has been studied using polyclonal antibodies, and an animal model, the *mdx* mouse, has been validated³; indeed, if this mouse model is insufficient, further models may be made in the laboratory using transgenic techniques. This structure for the study of a complex single gene disorder may become a classic for similar conditions.

A necessary prerequisite for the study of a disease at the DNA level is a genetic model; a small proportion of dysmorphic syndromes appear to be simply monogenic or polygenic, with no environmental component. However, most fall into the category of multifactorial disorders, determined by the varying contribution of both intrinsic and extrinsic factors, such that any given case will be the result of the net effect of these factors, as in coronary heart disease.⁴ It is important to distinguish between multifactorial and polygenic inheritance; we use the latter term in a restricted sense with reference to conditions caused exclusively by a large number of genes and with minimal extrinsic involvement. We also note that conditions which are multifactorial in general may nonetheless be monogenic or polygenic in a given family, as will be seen for one large family with cleft palate that we have studied.

The role that chance plays in the occurrence of malformations has recently been studied in the form of a stochastic single gene model.⁵ Using computer simulations to describe endocardial cushion outgrowth and fusion, randomly walking cells were allowed to migrate, divide, and adhere with preset probabilities. The results were interesting: the outcomes are variable, despite multiple independently performed simulations using identical parameters (that is, identical genotypes and environments). Not only did the stochastic single gene model generate a continuous liability curve resembling that which is obtained from a multifactorial threshold model,⁶ but it also showed that the variability was intrinsic to the simulation and did not require external factors, such as changes at the genetic or environmental level. If a number of simulations were performed in which a single parameter was changed (inter-cell adherence),⁷ the liability curves could be shifted to the left (increased defect risk) or to the right (decreased risk) of a predetermined defect threshold value.

In terms of clinical applications of this model, it has been noted that in Down's syndrome patients, significant variability in the extent of congenital heart malformation occurs, with approximately half of the subjects exhibiting heart defects.⁸ Discordance for congenital heart defects between monozygotic twins has been observed⁹; a similar result was found in the murine model for Down's syndrome (trisomy 16).¹⁰ In summary, the segregation of a given malformation can be explained by the inheritance of a single defective gene that predisposes to, but does not necessarily result in, the malformation. This idea is embodied in the concept of 'reduced penetrance' of a disease, words which now must be regarded as giving way to more precise definitions of genetic causalities.

There are a variety of perturbations which can affect normal embryogenesis. The term 'developmental field' is used to describe those parts of the embryo in which the appropriate processes of development are controlled and coordinated spatially and temporally. Disturbances of normal field development may be corrected or may lead to anomalies. Evidence for the existence of fields comes from the observation of identical malformations resulting from more than one cause, whether genetic or environmental or both, thus identifying a single developmental unit. In this manner, early perturbation of a field may lead to an extensive defect, whereas a late perturbation will result in a lesser effect.

Malformations, on the other hand, can be defined as a morphological defect of an organ, part of an organ, or larger region of the body resulting from an

intrinsically abnormal developmental process. Postnatally, it may be impossible to determine whether a given anomaly is a malformation or a disruption, caused by extrinsic factors (for example, teratogens) on an otherwise normal development. For instance, radial asplasia may be a malformation, as in Holt-Oram syndrome, or a disruption, as in thalidomide embryopathy. Virtually all malformations investigated to date appear to be field effects.¹¹ In contrast, a deformation is the results of mechanical forces (for example, intrauterine restriction) and results in an abnormal form, shape, or position of a part of the body. (By definition, a disruption can not be inherited.)

One initial approach to the understanding of the molecular basis of a dysmorphology lies in the identification of pedigrees in which a simple disease phenotype manifests as a purely genetic entity, but exhibits all the features of the more common polygenic or multifactorial disorder. These 'model families' do exist for many disorders both in humans and in the mouse. This review is concerned with the technology available to identify and analyse a defective gene, some of the clues to candidate developmental genes, and future perspectives of molecular biology with respect to correction of the defect involved.

Techniques of gene analysis

Since the introduction of gene cloning 10 years ago, one of the goals of molecular biologists has been to understand the mechanisms that cause a single diploid cell to develop and differentiate into a human being. The application of both molecular and genetic analysis has brought us a great distance towards understanding the genetics which underlies both the structure and function of organisms such as *E coli* and *Drosophila melanogaster*. These techniques have been difficult to apply to mammals, however, primarily owing to the sheer size of the mammalian genome.

If a genetic marker which has been mapped to a particular chromosome is always associated with a disease, then the disease may be assumed to be on that chromosome. In studying linkage, one very considerable contribution of molecular biology has been the observation that, even more than proteins, DNA sequences show abundant polymorphism.¹² These markers are called restriction fragment length polymorphisms (or RFLPs for short). Whereas RFLPs are rarely part of the coding sequence of a gene, and therefore do not usually denote functional variation, it has been estimated that a few hundred spread evenly over the genome would provide sufficient genetic markers to enable any gene of

interest to be mapped.¹³ The recombinational map of the human genome is now virtually complete.¹⁴

The analysis of linkage between two markers is dependent on the availability of sufficient numbers of informative families; this is true whether marker/marker or marker/disease linkage is studied, and is a function of the distance between the markers and the informativeness of the polymorphism (or disease). The likelihood of linkage (cosegregation) between the disease and each polymorphic marker can be followed using the maximum likelihood estimate, usually with the help of a computer program such as LIPED.¹⁵ The alternative to pairwise linkage analysis is to follow the segregation of several loci on the same chromosome simultaneously. This is a much more powerful method for linkage analysis within a defined linkage group and is performed using a program package such as LINKAGE.¹⁶ This approach is, however, limited by the accuracy of our knowledge of the recombination distance between the markers and of their order.

Once linkage has been found, and flanking markers isolated for a gene which is mutated so as to cause pathology, it is usually necessary to walk the genome in order to define the precise nature of the basic defect. At present, standard cloning strategies are limited by the capacity of the vectors used. Relatively small distances can be analysed by the use of chromosome 'walking', involving the repeated use of fragments from the end of one clone to identify adjacent overlapping clones. Generally these walks are much smaller than would be expected theoretically; as a result, a walk of 1 million bases would probably require at least 50 steps, which is impractical.

To overcome this problem, 'jumping' clones have been developed to cope with the large distances separating markers and genes of interest.¹⁷ Only the ends of very large DNA fragments generated by rare cutting restriction enzymes are cloned. In this way, a directional jump of 500 000 bases can be attained. The use of rare cutting sites as start/end points of jumps allows the distance covered by a jumping clone to be estimated. This is made possible by the development of an extremely powerful method, pulse field gradient gel electrophoresis (PFGE), which separates DNA molecules up to several million base pairs in length.¹⁸

For this technique, Schwartz and Cantor¹⁸ modified standard electrophoresis by pulsing the DNA with roughly diagonal electric fields, allowing the resolution of DNA fragments ranging in size from 50 000 to over 7 million base pairs. In combination with rare cutting restriction enzymes and jumping clones, it has now become possible to determine a physical, as opposed to recombinational, map

around a gene/marker locus, and to compare the two. This is the method of choice for determining gene order when flanking markers are so close to a defective gene that recombination is rarely seen.²

AN APPLIED EXAMPLE

As most congenital dysmorphic syndromes, such as cleft palate and neural tube defects (NTDs), have an undefined site of lesion, they are not amenable to direct genetic analysis. One exception to this generalisation is being studied in our laboratories. It involves the localisation of a single gene defect responsible for secondary cleft palate and ankyloglossia in a large Icelandic pedigree. Clefting of the human secondary palate (CP) occurs during the seventh to ninth weeks of embryological development when the lateral palatine processes fail to fuse. This malformation shows little variation between populations; the incidence is approximately 1 in 1500.¹⁹ Normally, CP is more prevalent in females, owing in part to the palatine processes fusing one week later, with a concordantly longer period during which the fetus is vulnerable to teratogenic agents. In the pedigree studied, males exhibited both cleft palate and ankyloglossia, whereas the females, in almost all cases, had only ankyloglossia. This is the result of incomplete penetrance of the defect in females. There were no cases of father to son transmission of the mutant gene, which demonstrated X linkage. We regard this family as an example of a single, rare, monogenic disorder that is amenable to molecular genetic analysis, and which mimics the much more common multifactorial condition.

Accordingly, linkage analysis was performed on the family using polymorphic DNA probes. A significant linkage was detected at zero recombination with a probe localised to Xq21.1.²⁰ The acquisition of further samples from the family and the analysis of other polymorphic markers in the Xq21 region has enabled the defect to be more accurately localised to the distal portion of Xq21.²¹ The gene is flanked by two marker loci, *DXYS12* and *DXS17*, which are thought to be at a recombinational distance of 5% (5 cM).

Further corroboration of this result is provided by the molecular analysis of two deletion patients. The first deletion patient²² has an interstitial deletion of Xq13.3→Xq21.33 and has, among other defects, cleft palate. This patient is deleted for the locus *DXYS12* but not for *DXS17*. The second deletion patient has a slightly smaller interstitial deletion in this region of the chromosome (Xq21.1→Xq21.31) and does not have cleft palate.²³ We suggest that the cleft palate defect lies in the region outside the overlap of these two deletions, that is, Xq21.31 to

Xq21-33, a distance of approximately 1 million bases. This DNA sequence length can be mapped and coding sequences within the region identified. The techniques of HTF island cloning,²⁴ chromosome mediated gene transfer (CMGT),²⁵ jumping libraries,¹⁷ and PFGE¹⁸ are currently being applied to this region of the X chromosome.

As an extension of our work on X linked CP, we are interested in other midline defects, specifically spina bifida. Although much is known about neural tube defects in general, we still know very little about what causes them, and why spina bifida is becoming less frequent in most parts of the world. Once again, the vast majority (>90%) of defects such as anencephaly, spina bifida, and encephalocele are of multifactorial origin.²⁶ An additional problem in the analysis of such NTDs is that the prevalence of malformations at birth is usually far less than the true incidence; a factor of 1/10 has been proposed.²⁷ Comparison of the relative rates of NTDs in spontaneously aborted fetuses in areas of different prevalences suggested that the incidence of NTDs associated with chromosomal abnormalities may be uniform and that it is the incidence of defects with cytogenetically normal chromosomes that varies. The differences of prevalence noted between males and females again appear to be related to differences in rates of embryological development.²⁸

Spina bifida may be analysed at the genetic level in both humans and mice. A pedigree has been reported in which anencephaly and spina bifida show an apparent X linked mode of inheritance.²⁹ This is amenable to linkage analysis as detailed earlier. Mouse models of diseases, especially embryological defects, are useful as they enable one to circumvent the ethical restrictions on human embryo research. The pertinent model for spina bifida is the curly tail mouse, whose defect is associated with a malformation in the spinal cord at the lowest point of closure, the posterior neuropore. Many of these mice also suffer from spina bifida: severity of the lesion is inversely proportional to the speed of neuropore closure. Mechanisms for neuropore closure are currently being investigated; hyaluronic acid deposition is a candidate for the defect.³⁰

Homeo boxes: the key to developmental regulation?

An unanswered question of mammalian development is how an embryonic cell comes to follow a specific pathway of differentiation. A major obstacle to our understanding of differentiation at the molecular level has been the inability to identify the relevant genes and the genes that control them. Not so for *Drosophila melanogaster*, where a rigorous

genetic dissection has enabled the characterisation of many key genetic loci controlling development.³¹

Experiments have shown that a 180 base pair protein coding sequence, the homeo box, is a component of some of these *Drosophila* genes. The homeotic genes control the developmental fate of groups of cells with respect to their morphological characteristics; homeotic genes control pattern formation.³² In all instances studied, the homeo box is part of an open reading frame and the sequence has been detected in mRNA. A comparison of a number of homeo box genes indicated a helix-turn-helix DNA binding motif within the box domain³³; it is thought that the activity of these genes is mediated by direct binding to controlling/promoting sequences of important genes.

Using the *Drosophila* homeo box sequence as probes, it has been possible to identify homologous sequences in a wide variety of vertebrates, including man. Interestingly, the mammalian sequences show high degrees of homology to the *Drosophila* probe, typically 70 to 95%. In addition to their strong primary sequence homology, there are other common features. The majority of homeo box copies in both *Drosophila* and mammalian genomes are organised within physically linked clusters. At least three human homeo boxes are known to be localised in a 20 000 base pair stretch of DNA.³⁴ A similar situation is found in the mouse.^{35 36} Even more striking is the similarity between homeo box containing genes in mouse and man, which are more homologous to each other (>90%) than to the *Drosophila* sequences.³⁶

The two main mouse homeotic clusters, Hox-1 and Hox-2, have been localised to chromosomes 6 and 11, and contain six and at least four homeo boxes respectively.³⁷ In all cases, these murine sequences were shown to be parts of genes whose patterns of expression are specific for given stages of embryogenesis, or for restricted numbers of adult tissues or cell lineages.³⁵⁻³⁷ It appears that some are coordinately regulated, being expressed in the same embryonic structure but at different times.³⁸⁻⁴² Similarly, if homeo box genes direct specific morphogenetic functions during development, as for *Drosophila*, one could predict specific spatial expression. This is elegantly shown for the murine Hox-3 and the human HHO.c10, with expression limited to the central nervous system.^{38 39}

The homeo boxes found in several *Drosophila* genes control segmentation: do they have the same function in mammals? It has not been possible so far to analyse developmental mutants in mammals with the same thoroughness as *Drosophila*. A major drawback is the definition of comparable aspects of mammalian and *Drosophila* development, along

with the practical aspects of embryo manipulation. Analysis of the embryo is not technically possible at the required stage, although Gaunt⁴³ has recently been able to detect expression in the mouse at the gastrulation stage by *in situ* hybridisation.

How do the homeo box proteins function? We know a considerable amount about their expression and spatial localisation in the embryo but very little about how they exert their effects: the only clue we have is the presence of a putative DNA binding domain, similar to that found in yeast mating type $\alpha 2$ protein.⁴⁴ This suggests a regulation role. The conservation of homeo box sequences across a broad range of organisms indicates a conserved function; just what this is in mammals has yet to be shown.

Teratocarcinoma cells provide a useful model of differentiation *in vitro*.⁴² Normally growing as undifferentiated cells, they can be made to differentiate by adding a variety of agents (for example, retinoic acid) to the culture medium. Such cells can be used to assay homeo box containing genes. In differentiated F9 cells, a homeo box containing 2.7 kb RNA is expressed at a high level but its expression is dependent on the continued presence of retinoic acid.

Homeotic genes may well play an important role in the normal development of the embryo, but what controls their expression? Hormones and growth factors are known to act as inducers of gene expression, and are good candidates for morphogens. Their receptors are intracellular proteins that act as the mediators of transcription; the recent cloning and characterisation of the genes for the human glucocorticoid and oestrogen receptors has revealed that these molecules contain discrete DNA binding and ligand binding domains. Another striking observation is the extensive homology between the DNA binding domains of steroid and thyroid hormone receptors; this has enabled the isolation of other receptor molecules, one of which is the retinoic acid (RA) receptor.⁴⁵

RA is essential for normal epithelial cell proliferation and differentiation,⁴⁶ and plays a major part in pattern formation.⁴⁷ It was suggested that RA also mediates gene transcription via nuclear receptors.⁴⁸ The cDNAs of several steroid receptors have been cloned and appear to be structurally related and share sequence homology with the avian erythroblastosis virus *v-erbA* oncogene.⁴⁹ The DNA binding domain, a stretch of approximately 70 amino acids (region C), determines the specificity of the transcriptional activation by folding into two zinc stabilised fingers, each involving two pairs of highly conserved cysteine residues.⁵⁰ Region E is the hormone binding domain and is much less

conserved than region C. By creating chimeras of the various receptor domains, it has been possible to use the glucocorticoid receptor to induce genes that are normally enhanced by oestrogen, and vice versa.⁴⁵ The cloning of the RA receptor is important as it enables a more detailed molecular analysis of differentiation at the gene level.

DNA transfection

The analysis of gene expression in prokaryotes has been based, to a large degree, on the introduction of genes and other DNA segments into bacteria. A similar approach in higher eukaryotes was impossible until relatively recently, when transformation assays using co-precipitation of DNA and calcium salts was introduced.⁵¹

Several well defined selectable markers have already been used in eukaryotic cells; thymidine kinase (tk) is the most fully studied, but this requires a tk⁻ cell line. Two vectors have been developed to enable any cell line to be used; both use prokaryotic genes linked to eukaryotic promoters. An example of this is the SVgpt series of vectors⁵²; an alternative dominant acting vector, SVneo, consists of the neomycin gene ligated to the SV40 early region.⁵³ This region encodes a kinase that inactivates neomycin; hence, mammalian cells can be transformed to neomycin resistance. The existence of selectable markers led to the idea of introducing any gene into mammalian cells, by co-transfection of the cloned DNA with the selectable marker.

DNA can also be stably introduced into cells by direct microinjection; no selective pressure is needed. The ability to introduce virtually any piece of DNA into cells has enabled the introduction of DNA sequences that control the expression of certain eukaryotic genes. Generally, a gene introduced into eukaryotic cells continues to respond to the signals that control the expression of the gene *in vitro*.⁵⁴

INTRODUCTION OF FOREIGN GENES INTO FERTILISED MOUSE EGGS

Transfection of cell lines with purified genes is proving to be a valuable method of studying gene regulation. However, this approach is, in the majority of cases, limited to partially or fully differentiated cells. The molecular mechanisms responsible for the temporal and spatial regulation of genes driving embryogenesis remains masked. This can be circumvented by the introduction of foreign genes at the zygote or early embryo stage. Cloned genes are microinjected into the pronuclei of freshly fertilised mouse eggs and implanted into the uterus.

Integration of the cloned gene into the host genome appears to be random, but the gene does remain in both germline and somatic cells. Unfortunately, by virtue of the experiment itself, we are selecting for integration events that are compatible with normal embryonic development, those that do not involve important morphogenic genes, or those that remain silent or underexpressed.

An exciting development using fused genes, in which the 5' regulatory sequences of the mouse metallothioneine gene were attached to the coding sequences of the rat growth hormone gene, indicated the power of transgenic analysis. In this case, the gene was induced by heavy metals and resulted in growth hormone levels 100 to 800 times that of the controls, with concomitant rapid growth of the mice.⁵⁵

Other examples of important transgenic experiments include (1) the construction of a mouse model for Lesch-Nyhan syndrome by introduction of a mutant HPRT into mice using transgenic embryonal stem (ES) cell lines^{56, 57}; (2) cell lineage ablation in transgenic mice by cell specific expression of a toxin gene⁵⁸; and (3) the correction of dysmyelinating phenotype by the introduction of the wild type gene into shiverer mice by microinjection; correct temporal and spatial expression of the gene was achieved.⁵⁹ This raises the possibility of gene therapy in man and its prospects will, in the future, be the subject of continued debate.⁶⁰

However, from the point of view of this review, it is more interesting to note that there are now many models of single genes which, when introduced into cells or organisms, alter morphology in a complex yet defined fashion. The experimental imperatives are to gain precision in site of integration and to improve expression, which will enable us to study many genes important in defining morphology during embryogenesis and differentiation.

References

- Monaco P, Neve RL, Colletti-Feener C, Bertelson CJ, Kurnit DM, Kunkel LM. Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* 1987;**223**: 646-50.
- Forrest SM, Cross GS, Speer A, Gardner-Medwin D, Burn J, Davies KE. Preferential deletion of exons in Duchenne and Becker muscular dystrophies. *Nature* 1987;**329**:638-40.
- Hoffman EP, Brown RH, Kunkel LM. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* 1987;**51**:919-28.
- Karathanasis SK, Norum RA, Zannis VI, Breslow JL. An inherited polymorphism in the human apoA-1 gene related to the development of atherosclerosis. *Nature* 1983;**301**:718-20.
- Kurnit DM, Layton WM, Matthysse S. Genetics, chance and morphogenesis. *Am J Hum Genet* 1987;**41**:979-95.
- Falconer IS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* 1965;**29**:51-71.
- Edelman G. Cell adhesion and morphogenesis: the regulator hypothesis. *Proc Natl Acad Sci USA* 1984;**81**:1460-4.
- Kurnit DM, Aldridge JF, Matsuoka R, Matthysse S. Increased adhesiveness of trisomy 21 cells and atrioventricular canal malformation in Down syndrome: a stochastic model. *Am J Med Genet* 1985;**30**:383-99.
- Rehder H. Pathology of trisomy-21 with particular reference to persistent common atrioventricular canal of the heart. In: Burgio GR, Fraccaro M, Tiepolo L, Wolf U, eds. *Trisomy 21: an international symposium, Rapello, Italy, 8-10 November 1981*. Berlin: Springer, 1981:57-73.
- Miyabara S, Sugihara W, Yonemitsu N, Yun K. Comparative study of the phenotypic expression of mice trisomy 16 by different female strains: attempt at an animal model for human trisomy 21. *Congenital Anomalies* 1984;**24**:283-92.
- Opitz JM, Herrmann J, Pettersen JC, Bersu ET, Colacino SC. Terminological, diagnostic, nosological and anatomical-developmental aspects of developmental defects in man. *Adv Hum Genet* 1979;**9**:71-80.
- Solomon E, Bodmer WF. The evolution of a sickle variant gene. *Lancet* 1979;**i**:923-4.
- Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980;**32**:314-31.
- Donis-Keller H, Green P, Helms C, et al. A genetic linkage map of the human genome. *Cell* 1987;**51**:319-37.
- Ott J. Estimation of the recombination fraction in human pedigrees; efficient computation of the likelihood for human linkage studies. *Am J Hum Genet* 1974;**26**:588-97.
- Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA* 1984;**81**:3443-6.
- Poustka A, Lehrach H. Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics. *Trends Genet* 1986;**July**:174-9.
- Schwartz DC, Cantor CR. Separation of yeast chromosome-sized DNAs by pulsed field gel electrophoresis. *Cell* 1984;**37**: 67-75.
- Leck I. The geographical distribution of neural tube defects and oral clefts. *Br Med Bull* 1984;**40**:390-5.
- Moore GE, Ivens A, Chambers J, et al. Linkage of an X chromosome cleft palate gene. *Nature* 1987;**326**:91-2.
- Ivens A, Moore GE, Chambers J, et al. X-linked cleft palate: the gene is localised between polymorphic markers DXYS12 and DXS17. *Hum Gen* (in press).
- Tabor A, Andersen O, Lundsteen C, Niebuhr E, Sardemann H. Interstitial deletion in the 'critical' region of the long arm of the X chromosome in a mentally retarded boy and his normal mother. *Hum Genet* 1983;**64**:196-9.
- Schwartz M, Yang HM, Niebuhr E, Rosenberg T, Page DC. Regional localisation of polymorphic DNA loci on the proximal long arm of the X chromosome using deletions associated with choroideremia. *Hum Genet* 1988;**78**:156-60.
- Estivill X, Farrall M, Scambler PJ, et al. A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. *Nature* 1987;**326**:840-5.
- Pritchard CA, Goodfellow PN. Investigation of chromosome mediated gene transfer using the HPRT region of the X chromosome as a model. *Genes Dev* 1987;**1**:172-8.
- Holmes LB, Driscoll SG, Athus L. Etiological heterogeneity of neural tube defects. *N Engl J Med* 1976;**294**:365-9.
- Bell IE, Gosden CM. Central nervous system abnormalities—contrasting patterns in early and late pregnancy. *Clin Genet* 1978;**13**:387-96.
- Seller MJ. Neural tube defects and sex ratios. *Am J Med Genet* 1987;**27**:699-707.
- Jensson O, Arnason A, Gunnarsdottir H, Petursdottir I, Fosdahl R, Hreidarsson S. A family showing apparent X linked inheritance of both anencephaly and spina bifida. *J Med Genet* 1988;**25**:227-9.

- ³⁰ Anderson A. Curly tailed mice point the way to an early diagnosis of spina bifida. *New Sci* 1987;8 Jan:34.
- ³¹ Akam M. The molecular basis for metamer pattern formation in *Drosophila* embryos. *Development* 1987;101:1-22.
- ³² McGinnis W, Levine M, Hafen E, Kuriowa A, Gehring W. A conserved DNA sequence found in homeotic genes of the *Drosophila* antennapedia and bithorax complexes. *Nature* 1984;308:428-33.
- ³³ Porter SD, Smith M. Homeo domain homology in yeast MAT 2 is essential for repressor activity. *Nature* 1986;320:766-8.
- ³⁴ Hauser CA, Joyner AL, Klein RD, Learned TK, Martin GR, Tijan R. Expression of homologous homeo-box-containing genes in differentiated human teratocarcinoma cells and mouse embryos. *Cell* 1985;43:19-28.
- ³⁵ Colberg-Poley AM, Voss SD, Chowdhury K, Stewart CL, Wagner EF, Gruss P. Clustered homeo boxes are differentially expressed during murine development. *Cell* 1985;43:39-45.
- ³⁶ Hart CP, Awgulewitsch A, Fainsod A, McGinnis W, Ruddle FH. Homeo box gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis, and homology to a human homeo box locus. *Cell* 1985;43:9-18.
- ³⁷ Duboule D, Baron A, Machl P, Galliot B. A new homeo box is present in overlapping cosmid clones which define the mouse Hox-1 locus. *EMBO J* 1986;5:1973-80.
- ³⁸ Awgulewitsch A, Utset MF, Hart CP, McGinnis W, Ruddle FH. Spatial restriction in expression of a mouse homeobox locus within the central nervous system. *Nature* 1986;320:328-35.
- ³⁹ Simeone A, Mavilio F, Bottero L, et al. A human homeo box gene specifically expressed in spinal cord during embryonic development. *Nature* 1986;320:763-5.
- ⁴⁰ Gaunt SJ, Miller JR, Powell DJ, Duboule D. Homeo box gene expression in mouse embryo varies with position by the primitive streak stage. *Nature* 1986;324:662-4.
- ⁴¹ Mavilio F, Simeone A, Giampaolo A, et al. Differential and stage-related expression in embryonic tissues of a new human homeo box gene. *Nature* 1986;324:664-8.
- ⁴² Breier G, Bucan M, Francke U, Colberg-Poley AM, Gruss P. Sequential expression of murine homeo box genes during F9 EC cell differentiation. *EMBO J* 1986;5:2209-15.
- ⁴³ Gaunt SJ. Homeo box gene Hox 1-5 expression in mouse embryos; earliest detection by *in situ* hybridisation is during gastrulation. *Development* 1987;101:51-60.
- ⁴⁴ Johnson AD, Herskowitz I. A repressor (MAT $\alpha 2$ product) and its operator control expression of a set of cell type specific genes in yeast. *Cell* 1985;42:237-47.
- ⁴⁵ Petkovitch M, Brand NJ, Krust A, Chambon P. A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature* 1987;330:444-50.
- ⁴⁶ Shubeita HE, Sambrook JF, McCormick AM. Molecular cloning and analysis of functional cDNA and genomic clones encoding bovine cellular retinoic acid-binding protein. *Proc Natl Acad Sci USA* 1987;84:5645-9.
- ⁴⁷ Thaller C, Eichele G. Identification and spatial distribution of retinoids in the developing chick limb bud. *Nature* 1987;327:625-8.
- ⁴⁸ Green S, Chambon P. A superfamily of potentially oncogenic hormone receptors. *Nature* 1986;324:615-7.
- ⁴⁹ Green S, Walter P, Kumar V, et al. Human estrogen receptor cDNA: sequence, expression and homology to v-crb-A. *Nature* 1986;320:134-9.
- ⁵⁰ Green S, Gronemyer H, Chambon P. In: Sluysen M, ed. *Growth factors and oncogenes in breast cancer*. Chichester: Horwood, 1987:7-28.
- ⁵¹ Graham FL, Abrahams PJ, Mulder C, Meijneker HL, Warnaar S, de Fries FAJ. Studies on *in vitro* transformation by DNA and DNA fragments of human adenovirus and simian virus 40. *Cold Spring Harbour Symp Quant Biol* 1974;39:637-50.
- ⁵² Mulligan R, Berg P. Selection of animal cells that express the *E coli* gene coding for xanthine-guanine phosphoribosyl transferase. *Proc Natl Acad Sci USA* 1981;78:2072-6.
- ⁵³ Southern PJ, Berg P. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under the control of the SV40 early region promoter. *J Mol Appl Genet* 1982;1:327-42.
- ⁵⁴ Corces V, Pellicer A, Axel R, Meselson M. Integration, transcription and control of a *Drosophila* heatshock gene in mouse cells. *Proc Natl Acad Sci USA* 1981;79:7038-42.
- ⁵⁵ Palmiter RD, Brinster RL, Hammer RE, et al. Dramatic growth of mice that develop from eggs microinjected with metallothionine growth hormone fusion genes. *Nature* 1982;300:611-5.
- ⁵⁶ Hooper M, Hardy K, Handyside A, Hunter S, Monk M. HPRT deficient (Lesch-Nyhan) mouse embryos derived from germline colonisation by cultured cells. *Nature* 1987;326:292-5.
- ⁵⁷ Kuchn MR, Bradley A, Robertson EJ, Evans MJ. A potential animal model of Lesch-Nyhan syndrome: introduction of the HPRT mutations in mice. *Nature* 1987;326:295-8.
- ⁵⁸ Palmiter RD, Behringer RR, Quaife CJ, Maxwell F, Maxwell IH, Brinster RL. Cell lineage ablation in transgenic mice by cell-specific expression of a toxin gene. *Cell* 1987;50:435-43.
- ⁵⁹ Readhead C, Popko B, Takahashi N, et al. Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. *Cell* 1987;48:703-12.
- ⁶⁰ Departments of Health and Social Security. *Report of the Commission of Inquiry into human fertilisation and embryology*. Chairman: Dame Mary Warnock DBE, et al. July 1984. HMSO: CMND. 9314.

Correspondence and requests for reprints to Dr A Ivens, Department of Biochemistry and Molecular Genetics, St Mary's Hospital Medical School, Norfolk Place, London W2 1PG.