The Journal of Laryngology and Otology
January 1998, Vol. **112**, pp. 4–10

Molecular biology and the ENT surgeon in the millennium

Andrew P. Read, M.A., Ph.D., F.R.C.Path.

Welcome to this new monthly series of articles on molecular biology for otolaryngologists – head and neck surgeons. Why might you want to read them? Well, time-honoured although it is, the tactic of ignoring molecular biology in the hope that it will go away is looking increasingly threadbare. Molecular biology is the substratum underlying the whole of clinical understanding and practice. While it remains distant from surgical procedures, molecular insights increasingly influence the overall management of the patient. Two examples may illustrate the increasing role of genetic knowledge in patient management.

Molecular biology in the management of ENT patients

Distinguishing Type 2 neurofibromatosis from sporadic vestibular schwannoma (VS) is important for managing the patient and family. Making the distinction can be difficult when a patient presents with unilateral VS at an unusually young age, or with minor additional manifestations of NF2 without any family history. All VS are believed to develop from a single cell that has lost both copies of the *NF2* or Merlin gene. In sporadic VS both losses ('hits') occurred in that cell only, while in NF2 one hit is inherited and is present in every cell of the patient's body. Knowing the DNA sequence of the NF2 gene, we can check it for mutations. Proposed clinical criteria for making the diagnosis of NF2 can be tested against the objective evidence of mutation screening. The best results come from studying the tumour, rather than DNA from the patient's blood, because we can be reasonably confident that every tumour cell has mutations affecting both its copies of the *NF2* gene. If both mutations can be seen in the tumour, but neither is present in the blood, then we are not dealing with inherited NF2. There is however an intermediate group of mosaic patients. Mosaic individuals were normal at conception but acquired a mutation some time during embryonic development, so that they carry the mutation in a proportion of the cells of their body. Their children, but not their brothers and sisters are at risk of NF2. Sometimes direct evidence of mosaicism can be obtained by molecular testing. Thus DNA analysis can be an important tool for deciding the correct management of these families. Unfortunately, *NF2* mutations might be anywhere in quite a large stretch of DNA, and present protocols are unable to find all of them. Negative results are difficult to interpret – have we failed to find the needle in the haystack, or is there no needle? Over the next few years improved methods for mutation detection, perhaps using DNA chips, should make such investigations more informative and more routine.

A second example concerns the profoundly deaf child, born into a family with no history of hearing loss. The bewildered parents want to know why it happened, what they did wrong, and what the risk is of it happening again. Although congenital non-syndromic sensorineural hearing loss has many causes, most cases are believed to be genetic and autosomal recessive. Not surprisingly, given the complexity of the auditory apparatus, mutations in many different genes can cause hearing loss. Recently one such gene was identified as the gene (*GJB2*) that encodes connexin 26, a gap junction protein found in many tissues including the inner ear. One particular nucleotide sequence in the *GJB2* gene is prone to error when the DNA is replicated, and it seems likely that a significant proportion of all congenital profound hearing loss is caused by loss of the same single nucleotide from the *GJB2* gene. If further research confirms this, then it should be possible to devise a simple DNA test which, if positive, will identify the precise cause and recurrence risk.

Identifying human disease genes

Both examples above illustrate the advantages that flow from identifying the gene causing a disease. Over the past dozen years, geneticists have made extraordinary progress in locating and identifying human disease genes. David Bourn's articles in this series should help explain the technical background, but even without that background the general strategies are not hard to understand. There is no single standard way to identify a gene, but in Figure 1 I have tried to show the main strategies that have been used. To illustrate how these work, and as a

From the University Department of Medical Genetics and Regional Genetics Service, St Mary's Hospital, Manchester M13 0JH. Fax: 0161-276-6605; Email: andrew.read@man.ac.uk.

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press

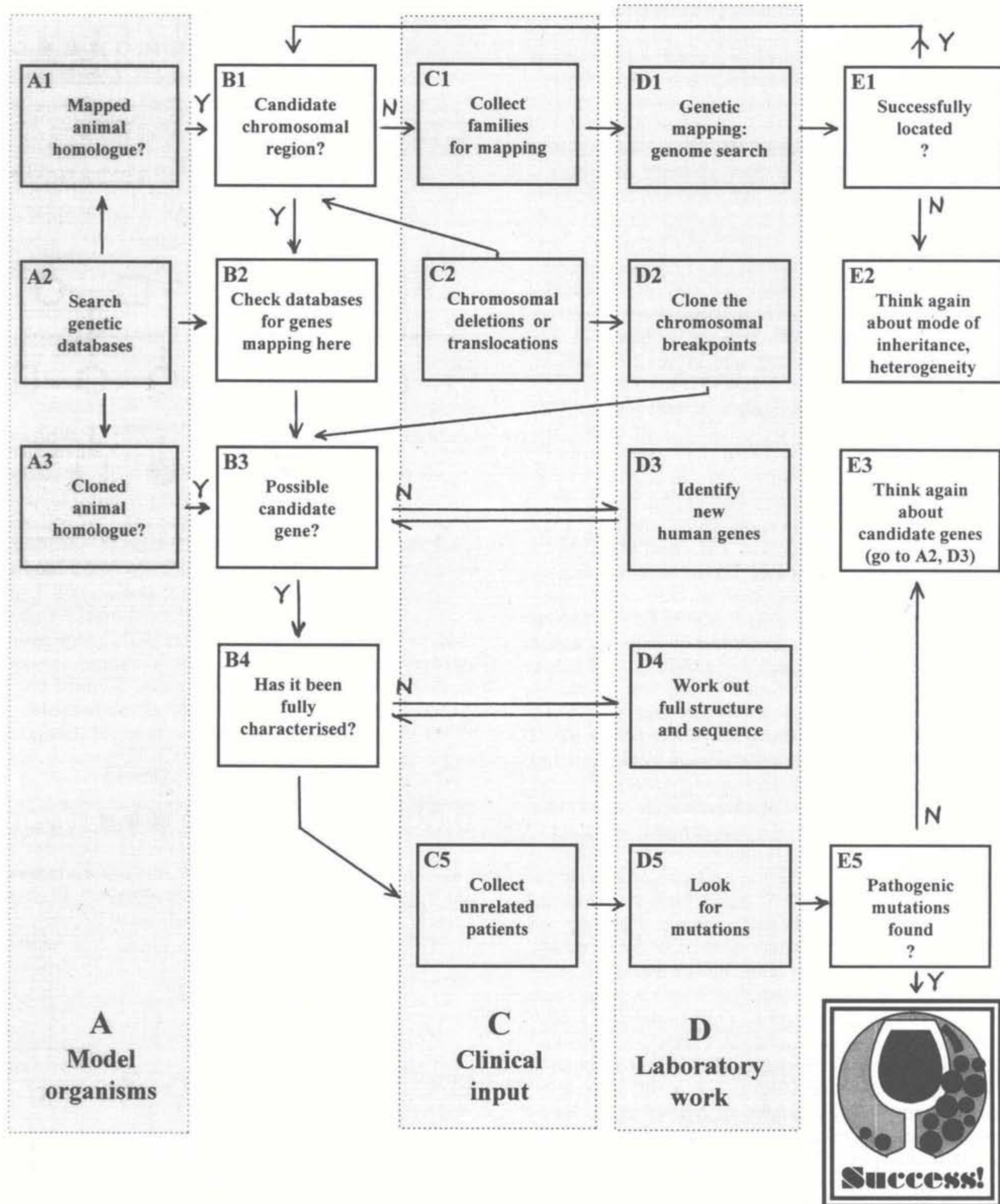


FIG. 1

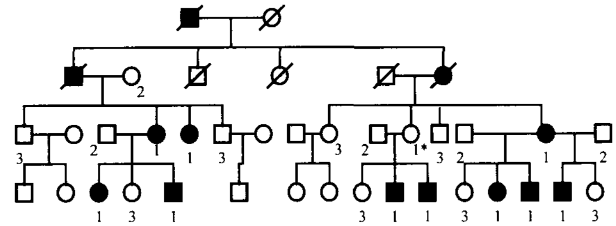
(Reproduced with permission from Bios Scientific Publishers, UK and John Wiley, USA).

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press

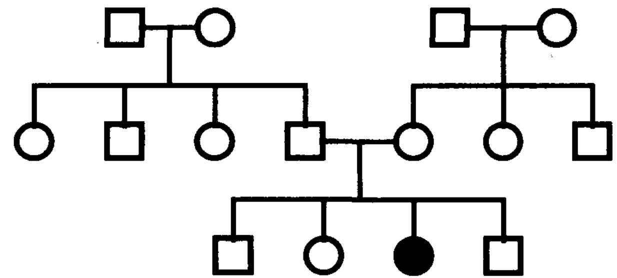
taster for the rest of the series, I will describe the work, by my group and others, that has led to the identification of four of the genes that cause various forms of Waardenburg syndrome (Newton and Read, 1997). Karen Steel and Tracy Bussoli will no doubt enlarge upon these topics in their articles in the series – but meanwhile, note how clinical investigations, laboratory molecular biology, computer searches and animal studies all played their part in the overall strategy.

Waardenburg syndrome (WS) is biologically interesting because the patchy absence of melanocytes from the eyes, hair, skin and stria vascularis (the latter causing hearing loss) points to a fault early in embryogenesis when melanocyte precursors are differentiating from the neural crest and migrating to their final positions. Firstly my clinical colleagues collected blood from a set of meticulously worked up families with Type I WS (Figure 1, box C1). As Fiona Macdonald will explain in a later article, genetic mapping (Figure 1, box D1) involves following the segregation of common non-pathological DNA variants (genetic markers) through the families until a marker is found that tracks perfectly with the disease. This can require testing up to 400 markers, which gets tedious and expensive. It helps to have a clue. Our clue came from a chromosome abnormality (Figure 1, box C2) described by a Japanese group in a poster I saw at a meeting in New Haven. They described a child with non-inherited WS and a *de novo* inversion of the bottom end of chromosome 2. Such things are often just coincidental, but it raised the possibility that one inversion breakpoint had caused the WS by slicing through the WS gene – in which case the break defined the location of the gene. We tested markers chosen from the breakpoint region, and they did indeed track with the disease in our families. So we had a candidate chromosomal region (Figure 1, box B1) without the labour of a full genome search.

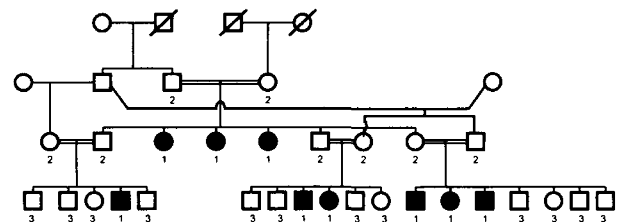
The whole human genome contains 3000 Mb (million base pairs) of DNA, and our linkage work had only narrowed the WS1 gene down to perhaps a 10 Mb region. Positional cloning (cloning an unknown gene, knowing only its approximate chromosomal location) can be desperately hard work, as Mike Dixon will show in his article on Treacher Collins syndrome. One looks for ready-cloned candidate genes, and increasingly the Human Genome Project is providing these. Unfortunately, checking the databases (Figure 1, box B2) at the time revealed no plausible candidate human gene. However, looking at genetic maps of humans and mice we saw that our part of human chromosome 2 had a close counterpart in mouse chromosome 1, with many homologous genes conserved across the two species. Among mouse mutants mapping to this region was *Splotch*. *Sp*/+ heterozygous mice have a white belly splotch (patchy absence of melanocytes), and *Sp/Sp* homozygous mice have lethal neural tube defects, showing that the *Sp* gene functions in the embryonic neural tube and neural crest. So we had a likely animal homologue for human WS1 (Figure 1,



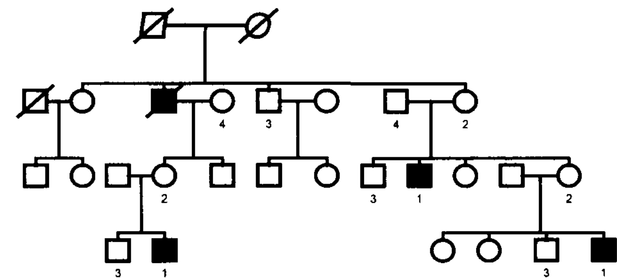
1. Typical autosomal dominant pedigree pattern, with priorities for sampling indicated. NB transmission by affected males rules out mitochondrial inheritance. *This person has an affected parent and an affected child, but is herself unaffected. This is an example of non-penetrance.



2. Typical autosomal recessive pedigree pattern, mimicking a sporadic condition. Such families are not useful for linkage studies, but may be useful for mutation detection



3. An example of a multiply-inbred pedigree with many cases of an autosomal recessive condition. Such rare families are exceptionally valuable for linkage analysis. Priorities for sampling are indicated.



4. X-linked recessive pedigree pattern, showing likely priorities for sampling

Fig. 2

Examples of pedigree patterns and priorities for sampling for research. 1 = highest priority, 2 = intermediate, 3 = low but still useful. (as published in HEAR Infoletter 2)

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press

box A1). The *Sp* gene had not been cloned; however, database searches of mouse genes identified a gene, *Pax-3*, that mapped in the right general area and was known to be expressed in the embryonic neural crest. *Pax-3* became an obvious candidate for both *Splotch* and WS1 (Figure 1, box B3). The human counterpart of *Pax-3* had been only partially characterized (Figure 1, box B4), but there was enough information for us to test our DNA collection for mutations (Figure 1, box D5). Soon we found mutations (Figure 1, box E5), as did Clint Baldwin's group in Boston, who had been following similar logic. Later Baldwin showed that the rare variant, Type 3 WS or Klein-Waardenburg syndrome (WS1 with limb abnormalities) was also caused by mutations in *PAX3*. Why some patients have this more severe manifestation is still not understood.

We knew that Type 2 WS (where there is no dystopia canthorum) was genetically distinct because our family data showed that it did not map to chromosome 2. We had a possible mouse model in the *microphthalmia* mouse. Despite the name, its main feature is a deficiency of melanocytes. *mi* was known to map to mouse chromosome 6 (Figure 1, box A1), but it proved hard to identify the human counterpart of this region. We tested several candidate locations with negative results, but we did not know whether this meant that *mi* was not the right homologue, or that we had not guessed the right location for human *mi*. Eventually we collaborated with Dr Anne Hughes in Belfast in a whole genome search (Figure 1, box D1) in a single large Type 2 family, and located the gene on chromosome 3. Meanwhile in the USA an apparently totally unrelated piece of work was about to provide a vital clue. A transgenic mouse had been made, containing a piece of foreign DNA. Transgenes integrate into the recipient chromosome at random, and occasionally they insert into an existing gene and inactivate it. In this case the transgenic mouse resembled the *mi* mutant, suggesting that the introduced DNA had fortuitously integrated into the *mi* gene. This lucky accident allowed the mouse *mi* gene to be cloned. Armed with the DNA sequence of the mouse gene, Masayoshi Tachibana's group in Bethesda, Maryland identified the human counterpart (named *MITF*) and mapped it. It mapped to precisely the location on chromosome 3 to which we, unknown to them, had just mapped WS2. We now had an excellent candidate gene (Figure 1, box B3). It was necessary to work out its structure (Figure 1, box D4) before we could proceed to mutation testing (Figure 1, box D5). Before long we had found mutations (Figure 1, box E5), and so another Waardenburg gene was identified.

The rare variant Shah-Waardenburg or Type 4 Waardenburg syndrome, is a neurocristopathy that combines features of WS2 with Hirschsprung disease. Genes underlying WS4 were identified by American, Dutch and French groups. In each case a transgenic mouse fortuitously showed the phenotype of a relevant animal model (Figure 1, box A3). The first gene, the endothelin receptor *EDNRB*, was a strong candidate for investigation because the human gene mapped to a previously identified candidate region for WS4 on chromosome 13 (Figure 1, box B1). Finding *EDNRB* mutations in a few families gave investigators the confidence to test directly for mutations in the second gene, endothelin 3, even without human mapping data to support its candidature, and again mutations were found in some families (Table I).

Collaboration between clinicians and scientists

Clinicians and scientists each have essential roles to play in identifying human disease genes. As the Waardenburg examples show, progress depended absolutely on the clinicians ascertaining suitable families and performing a fastidious work-up that involved many home visits and much weekend work. And without the Japanese patient, WS1 would still have been mapped, but it would have taken a lot longer.

Researchers always want to hear of patients and families suitable for research. What is required depends on what stage research has reached with the particular family condition. Initial studies use linkage analysis to map the gene, whilst later stages

TABLE I
HOW THE GENES UNDERLYING WAARDENBURG SYNDROME WERE IDENTIFIED

Gene	Path to identification of gene (see Figure 1)	References
PAX3 (Type 1 WS)	(C1, C2) – B1 – A2 – A1 –B3 – B4 – C5 – D5 – E5	Foy *et al.* (1990); Tassabehji *et al.* (1992)
MITF (some Type 2 WS)	C1 – D1 – E1 – B1 – A3 – B3 – B4 – D4 – D5 –E5	Hughes *et al.* (1994); Tachibana *et al.* (1994); Tassabehji *et al.* (1994)
EDNRB (some Type 4 WS)	C1 – D1 – E1 – B1 – A3 – B3 – B4 – C5 – D5 – E5	Hosoda *et al.* (1994); Puffenberger *et al.* (1994); Amiel *et al.* (1996)
EDN3 (some Type 4 WS)	A3 – B3 – C5 – D5 – E5	Baynash *et al.* (1994); Edery *et al.* (1996); Holfstra *et al.* (1996)

The middle column refers to boxes in Figure 1. The final common pathway always involved defining a candidate gene (B3) and then a successful search for mutations in patients (D5 – E5). The candidate gene was defined by using varying combinations of human linkage analysis (D1), human chromosomal aberrations (C2) and animal models (A1, A3). These pathways are typical of the various ways in which human disease genes are identified by clinicians and scientists working in partnership.

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press

move towards seeking mutations in a candidate gene. In all cases, good clinical descriptions are essential, and at least some family members must be willing to donate samples, usually of blood. Patients with a mendelian phenotype, but who also have additional features, are especially interesting. They sometimes have chromosomal abnormalities that can be enormously helpful for mapping or cloning a gene. For example, we found a boy with Waardenburg syndrome, growth retardation and mental retardation. He turned out to have a chromosomal deletion encompassing the *PAX3* gene. The guidelines in the Appendix, developed as part of a European Concerted Action on Genetics of Hearing Impairment, may help smooth the path of collaboration.

The future: gene therapy?

At present genetic knowledge assists management through improved diagnosis, prediction of prognosis and definition of risks to the family. Ultimately, of course, we would want to be able to use gene therapy to cure the patient. Here it is worth distinguishing between problems caused by a gene misbehaving, or failing to act, here and now in the patient in front of us (as with vestibular schwannomas) and problems caused by a gene having malfunctioned during the patient's early development, as in Treacher Collins syndrome. Problems of the first type may well be amenable to gene therapy. The time-scale is hard to predict. Gene therapy has been oversold as an imminent answer to all clinical problems; the reality is that the road leading from the first laboratory success to practicable treatment is long and uncertain. Already expression of an exogenous gene has been demonstrated after infusion of a genetically engineered DNA construct into the inner ear of guinea pigs (Lalwani *et al.*, 1996). Clinical trials may move from animals to humans quite soon. However, at least for the next few years, genetic advances can offer patients hope but not treatment. A later article in the series will explore the options for gene therapy in more detail.

Where will it all end up?

For better or for worse, we live in exciting times. Underlying present advances in diagnosis and future hopes for gene therapy is the extraordinary rate at which the human genome is being explored. Within the last decade the genes responsible for almost all the major mendelian (single-gene) human syndromes have been mapped. Increasingly the genes themselves are being identified, and the focus of research is moving on to working out the normal function of each gene in the cell. Other researchers are mounting sustained campaigns to unravel the genetic changes that convert a normal cell into a malignant tumour (to be described in articles by Helena Rowley and Richard Irving later in this series), and to identify the genetic factors governing susceptibility to diabetes, hypertension, schizophrenia and other major common diseases (as an example, Stephen Holgate will discuss the molecular biology of allergy). Progress involves a rich interplay between clinical family studies, laboratory investigations of cells and DNA, computer database searching and work on model organisms. Many an advance has started from an observation made by an alert clinician on a single patient.

As part of the Human Genome Project, massive automated DNA sequencing factories, such as the Sanger Centre in Cambridge, are adding tens of thousands of nucleotides to the Internet sequence databases every day. Already we know the total genomic sequence of many bacteria, of one eukaryote (yeast) and, in 1998, of the first multicellular organism, the nematode worm *Caenorhabditis elegans*. Dozens more genomes are being sequenced, including of course the human genome. All this activity is not just accumulating detail but revealing principles. For the first time, we can have total knowledge of an organism's genetic programme and work out exactly how it works. One message coming through loud and clear is the unity of life. The master gene governing formation of the eye in the *Drosophila* fruitfly performs the same task in humans, despite the radically different structure of an insect's compound eye. Even more amazing, the *Drosophila* mutant *Apterous*, which lacks wings, can be restored to normality by injecting the DNA of the human *Apterous* gene. What are we doing with a gene for making fly's wings? We don't use it for making wings, of course, but it is a dramatic demonstration of the fact that all organisms use essentially the same limited set of genetic pathways. We have the same *Apterous* pathway as *Drosophila*, but we use it for a slightly different developmental purpose – just like a computer subroutine for sorting numbers into order can be used in various different contexts. Studies of *Drosophila*, nematodes and yeast have a degree of relevance to human physiology that few would have dared to predict even five years ago.

These are indeed revolutionary times, but as with all revolutions, when the dust settles we will find life as we always knew it – maybe slightly refracted but still basically the same. Patients will still need operations, and physical anatomy will always be more important to surgeons than genetic anatomy. Genetic knowledge, however wonderful, can only explain those aspects of man that are genetically determined, and even those may not always be explained in a useful way. Consider, for example, the differences between humans and mice – two organisms that run on a virtually identical set of genes. The difference is the sum total of myriad individually trivial differences in the timing or level of expression of genes – but most of us feel that this sum total is not entirely trivial. Perhaps then the best reason for acquiring enough insight into molecular biology to keep up with the revolution is not so much its importance within its own domain, but the fact that a bit of knowledge gives the opportunity to watch some truly exciting science in the making. The

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press

biggest scientific party the world has ever seen is taking place in the laboratory – wouldn't it be a pity to stay locked outside the door?

References

Amiel, J., Attie, T., Jan, D., Pelet, A., Edery, P., Bidaud, C., Lacombe, D., Tam, P., Simeoni, J., Flori, E., Nihoul-Fekete, C., Munnich, A., Lyonnet, S. (1996) Heterozygous endothelin receptor B (ENDRB) mutations in isolated Hirschsprung disease. *Human Molecular Genetics* **5:** 355–357.

Baynash, A. G., Hosoda, K., Giaid, A., Richardson, J. A., Emoto, N., Hammer, R. E., Yanagisawa, M. (1994) Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* **79:** 1277–1285.

Edery, P., Attie, T., Amiel, J., Pelet, A., Eng, C., Hofstra, R. M. W., Martelli, H., Bidaud, C., Munnich, A., Lyonnet, S. (1996) Mutation of the endothelin-3 gene in the Waardenburg-Hirshsprung disease (Shah-Waardenburg syndrome). *Nature Genetics* **12:** 442–444.

Foy, C., Newton, V. E., Wellesley, D., Harris, R., Read, A. P. (1990) Assignment of WS1 locus to human 2q37 and possible homology between Waardenburg syndrome and the Splotch mouse. *American Journal of Human Genetics* **46:** 1017–1023.

Hofstra, R. M. W., Osinga, J., Tan-Sindhunata, G., Wu, Y., Kamsteeg, E. -J., Stulp, R. P., van Ravenswaaij-Arts, C., Majoor-Krakauer, D., Angrist, M., Chakravarti, A., Meijers, C., Buys, C. H. C. M. (1996) A homozygous mutation in the endothelin-3 gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah-Waardenburg syndrome). *Nature Genetics* **12:** 445–447.

Hosoda, K., Hammer, R. E., Richardson, J. A., Baynash, A. G., Cheung, J. C., Giaid, A., Yanagisawa, M. (1994) Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce megacolon associated with spotted coat colour in mice. *Cell* **79:** 1267–1276.

Hughes, A., Newton, V. E., Liu, X. Z., Read, A. P. (1994) A gene for Waardenburg syndrome type 2 maps to human chromosome 3p12–p14.1. *Nature Genetics* **7:** 509–512.

Lalwani, A. K., Walsh, B. J., Reilly, P. G., Muzyczka, N., Mhatre, A. N. (1996) Development of *in vivo* gene therapy for hearing disorders: introduction of adeno-associated virus into the cochlea of the guinea pig. *Gene Therapy* **3:** 588–592.

Newton, V. E., Read, A. P. (1997) Waardenburg syndrome. *Journal of Medical Genetics* **34:** 656–665.

Puffenberger, E. G., Hosoda, K., Washington, S. S., Nakao, K., deWit, D., Yanagisawa, M., Chakravarti, A. (1994) A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. *Cell* **79:** 1257–1266.

Tachibana, M., Perez-Jurado, L. A., Nakayama, A., Hodgkinson, C. A., Li, X., Schneider, M., Miki, T., Fex, J., Francke, U., Arnheiter, A. (1994) Cloning of MITF, the human homolog of the mouse microphthalmia gene and assignment to chromosome 3p14.1–p12.3. *Human Molecular Genetics* **3:** 553–557.

Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P., Strachan, T. (1992) Waardenburg syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355:** 635–636.

Tassabehji, M., Newton, V. E., Read, A. P. (1994) Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. *Nature Genetics* **8:** 251–255.

Appendix – Guidelines for collaborating with a genetics laboratory

How to make a family tree

The key point is to be systematic. Rule lines for the generations on a large sheet of paper. Start with the proband in the centre. Ask systematically about all first, second and third degree relatives, filling in the pedigree as you go. Mark name and maiden name, date of birth, age at death and cause, and relevant medical history. Ask specifically about infant deaths, stillbirths, abortions and consanguinity. Even with a dominant condition, ask about both sides of the family.

How to collect and send samples

For DNA extraction, a venous blood sample, preferably 10 ml, is taken into EDTA, stored at room temperature or 4°C but *not frozen* and despatched to the laboratory to arrive within 48 hours if at all possible. A laboratory may request blood for chromosome analysis for setting up a cell line. In this case blood, preferably 10 ml, is taken into lithium heparin, kept at room temperature and sent at room temperature to arrive at the laboratory within 48 hours. Tissues in addition to blood are particularly valuable for mitochondrial mutations, and the opportunities afforded by biopsies performed during the clinical work-up, e.g., muscle biopsies, should be borne in mind. Tissue samples should be frozen (−70°C) without fixation.

What samples to collect for linkage studies (see Figure 2)

It is best to contact the laboratory after drawing up the family tree but before taking samples. The following is a general guide. Linkage analysis requires samples from as many family members as possible. It is important to be sure who is affected and who is unaffected. Establishing a new linkage requires a much larger pedigree than confirming or refining an existing linkage.

(a) Autosomal dominant conditions: A minimum for establishing a new linkage would be ten individuals (affected or unaffected) each of whom has an affected parent.

(b) Autosomal recessive conditions: In general, only families with several affected individuals are useful. Minimum useful families would be:

i) three affected children born to unrelated parents.

ii) two affected children born to consanguineous parents.

Unaffected individuals (apart from parents of affected individuals) are of less value for analysis in recessive than in dominant conditions.

(c) X-linked inheritance: Any family with at least two affected males has some potential for linkage analysis, especially if they have different mothers. Offspring of unaffected males, and women with no affected descendants do not usually give useful information for linkage.

(d) Mitochondrial inheritance: A good family tree is needed to establish mitochondrial inheritance, but laboratory linkage analysis is not useful.

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press

What samples to collect for mutation analysis.

Once the condition has been mapped, the emphasis of genetic research moves towards mutation detection. The priority is a sample from a single well-described affected individual, from each separate family. With an autosomal dominant condition, if an affected person appears to be the first in the family and has unaffected parents, then samples from that person and those parents are valuable. With X-linked conditions, samples from affected males are more valuable than samples from females, even obligate carriers. If mitochondrial inheritance is suspected, samples from a single clearly affected individual, and preferably also the person's mother, should be taken. Analysis of mitochondrial mutations is a specialized area, and it is important to discuss plans in advance with the laboratory.

https://doi.org/10.1017/S0022215100139817 Published online by Cambridge University Press