GENOMICS

Global views of leukaemia

Todd R. Golub

Genomic characterization of a type of leukaemia has resulted in the identification of common genetic abnormalities that underlie the disease. The results constitute an advance on several fronts.

It is now widely acknowledged that cancer is a genetic disease, caused largely by the acquisition of mutations in somatic (non-germline) cells after birth — or in some cases during fetal development. A systematic search for these mutations was previously impossible, but progress in genomics technology holds the promise of making the complete characterization of the 'cancer genome' feasible.

An important advance has been the development of high-density DNA microarrays for detecting regions of genomic amplification or deletion. In principle, the application of these arrays might allow for both a global view of recurring abnormalities in genomic copy number, and a sufficiently precise mapping of those abnormalities to allow the gene(s) affected to be identified. On page 758 of this issue, Mullighan and colleagues[1] transform this concept into practice with their discovery of frequent deletions and loss-of-function mutations in the *PAX5* gene in childhood acute lymphoblastic leukaemia (ALL). This disease results from a defect in the differentiation of blood cells: an overproduction of immature B lymphocytes that overtakes the normal development of blood cells in the bone marrow.

By studying 192 ALL samples using DNA microarrays containing probes for about 350,000 genetic loci, the authors were able to rapidly identify deletions of the short arm of chromosome 9 in 57 patients (30%). Because of the high density of probes on these arrays, it was possible to narrow down the target of the deletion to a single gene, *PAX5*, which encodes a transcription factor. Some cases showed tiny deletions affecting only those regions of the gene that code for a particular part of the PAX5 protein, the 'transactivation domain'.

Most cases with a *PAX5* deletion lacked mutations in the remaining copy (allele) of the *PAX5* gene. This is strong evidence for 'haploinsufficiency', the situation in which the single functional allele cannot produce enough of the protein product. The authors hypothesized that some of the ALL cases might suffer from PAX5 loss of function by another mechanism, namely point mutation — the mutation of a single base. Indeed, they discovered *PAX5* point mutations in an additional 14 patients — all but one of the mutations occurred in only one of the two alleles, which again supports the notion of haploinsufficiency. The results with *PAX5* add to a growing number of examples of haploinsufficiency involving transcription factors[2].

Another striking finding[1] was the frequent co-occurrence of *PAX5* loss of function with particular molecular abnormalities — for example, a chromosomal translocation resulting in the fusion of two genes, *ETV6* and *RUNX1* (ref. 3). This is detectable at the time of birth in blood taken from the umbilical cord of infants who go on to develop ALL years later, suggesting that additional genetic hits are required to produce the consequences of *ETV6/RUNX1* fusion[4]. Consistent with this observation, mice genetically engineered to harbour only the *ETV6/RUNX1* defect do not suffer from the hallmarks of leukaemia[5]. So it seems that *PAX5* haploinsufficiency may represent a collaborating event in the development of *ETV6/RUNX1* leukaemias.

Mullighan and colleagues' study has implications for large-scale efforts that aim to characterize the cancer genome, including the Cancer Genome Atlas project sponsored by the US National Institutes of Health. Recent reports have led some to believe either that mutations in cancer genes are exceedingly rare, or that the molecular heterogeneity of cancer will lead to a mountain of genomic data that makes the problem hopelessly complex. But it is evident that when high-density approaches are applied to large numbers of patient samples, clear patterns of recurrent molecular abnormalities emerge.

The study[1] further shows that analyses of genomic copy number, which are relatively inexpensive, can be used to guide the selection of genes for resequencing: until it becomes possible to affordably sequence cancer genomes in their entirety (or at least sequence all protein-coding regions), such prioritization will be necessary. It also highlights the recurrent theme of a connection between cancer and cellular differentiation. Apart from abnormalities in *PAX5*, Mullighan *et al.* found defects in other genes — *EBF, IKZF1, IKZF3, LEF1, TCF3* and *BLNK* (although with much less statistical support) — thought to be involved in regulating lymphocyte differentiation. A systematic

GEOLOGY

Crystal-clear ideas

Caves stumbled upon deep in the workings of the Naica mine in Chihuahua, Mexico, are famed for the huge, elongated crystals of selenite that they contain. The most spectacular such discovery, the transparently named *Cueva de los Cristales*, came in 2000.

The crystals found there — at 300 metres' depth, a temperature of just under 60 °C and 100% humidity — are up to 11 metres long (see picture). Juan Manuel García-Ruiz and colleagues suggest the very specific conditions that were required for such huge crystals to form (J. M. García-Ruiz *et al. Geology* **35,** 327–330; 2007).

Selenite is the colourless crystalline form of calcium sulphate dihydrate ($CaSO_4 \cdot 2H_2O$), better known as gypsum. Gypsum's solubility in water is highest at around 58 °C. Above that temperature, however, it is thermodynamically less stable than the anhydrous form of calcium sulphate, known simply as anhydrite.

According to the authors' ideas, then, the ambient conditions of the *Cueva de los Cristales* would have been ideal for maintaining low-salinity cave-water slightly undersaturated in anhydrite (which would have originated from hydrothermal processes during the formation of the surrounding rock) and slightly supersaturated in gypsum over long periods. This tiny equilibrium imbalance would have encouraged the slow, sparse formation of very long gypsum crystals.

The authors backed up their theory with various hydrochemical and geochemical analyses of samples from the cave. But they point out that the effulgent result of the unique natural circumstances that they describe might not be visible for future generations. The Naica mines today number among the most important for lead and silver in the world. Once their booty is exhausted, however, and water-pumping of the workings stops, nature might reclaim the *Cueva de los Cristales*, reverting it to its erstwhile fully flooded state. **Richard Webb**

J. TRUEBA/GEOL. SOC. AM.