as in these highly oriented diamond films grown on silicon. Such heteroepitaxial films exhibit columnar growth and have low defect densities in the grains and shallow-angle grain boundaries.

Further improvements of the electronic properties can be anticipated. For example, Kobe Steel has claimed hole mobilities of up to 1400 $cm^2\ V^{-1}\ s^{-1}$ for homoepitaxial films. The use of substrates such as platinum may enable hole mobilities of natural single-crystal diamond (about 1800 to 2400 $cm^2\ V^{-1}\ s^{-1}$) to be reached. Very recently, biased enhanced nucleation on iridium was reported by Schreck *et al.* (*3*). The latter two techniques appear to result in coalesced grain boundaries, which may be "transparent" to the transport of charge carriers.

In contrast to p-type dopants, the incorporation of phosphorus and other n-type dopants (with the aim of producing a shallow donor) into diamond has been inefficient. This has been the main impediment to diamond bipolar devices. The main approaches have been in situ doping or cold implantation of phosphorus ions followed by rapid thermal annealing. Koizumi *et al.* used gas-phase doping in the fabrication of both regions of their p-n junction. While a Ti/Au layer was used to make electrical contact to the p-type layer, a patterned ohmic contact to the n-type region was formed by Ar^+ implantation to form a matrix of graphitic dots. Implantation of rare gas atoms into the diamond lattice yields n-type conductivity. The n-type behavior is, however, associated with the creation of defects in the band gap. This process must be minimized to avoid defect-induced conductivity, dopant compensation, or formation of dopant-vacancy complexes.

Okano *et al.* (*4*) first reported the production of n-type P-doped polycrystalline diamond films. They reported carrier mobilities of about 50 $cm^2\ V^{-1}\ s^{-1}$, slightly lower than the 60 $cm^2\ V^{-1}\ s^{-1}$ reported by Koizumi *et al.* In 1999, Sakaguchi and co-workers (*5*) announced shallow n-type doping using sulfur as a dopant. The results were initially difficult to replicate, but evidence for sulfur doping has now been reported by other groups. For example, Stutzman and co-workers report electron mobilities of 250 $cm^2\ V^{-1}\ s^{-1}$ at 290 K (*6*). There is ample scope for further improvements in light-emitting CVD diamond p-n junctions.

Natural diamonds are diverse in their individual properties compared to CVD diamond. Notwithstanding the difficulties of processing bulk diamond, the presence of inclusions, impurities, and—most important—nonuniformly distributed nitrogen defects degrades the electronic properties. Semiconducting type IIb crystals containing boron are rare, and, in any event, the boron dopant is not homogeneously distributed and predominates along growth fronts.

Relatively small single crystals of diamond may be grown reproducibly at high pressure and high temperature, but, until recently, contamination by the transition metal catalysts and by nitrogen complexes and aggregates was a problem in growing electronic-grade diamond by this route. The process has been improved by using nitrogen getters (materials used to remove impurities), but it is too early to predict whether high-pressure–high-temperature diamond will be used extensively in active electronics or whether the material is suitable for diamond light-emitting diodes.

To date, CVD is the most promising route to diamond films for use in diamond electronics, as demonstrated by the junction reported by Koizumi *et al.* (*2*). Predictions (*1*) of a market of $1.3 billion for GaN-based optoelectronic and electronic devices within a few years may even be conservative. Ultraviolet diamond lasers have a similarly promising future. For example, key applications in DVD data storage will be enhanced with such a laser. Numerous other applications are envisaged.

References

1. Blue Lasers: Lighting Up New Markets in Displays, Optical Storage, Medical Devices, and More, Report D219 (Technical Insights, Frost & Sullivan, NY, November 2000).
2. S. Koizumi, K. Watanabe, M. Hasegawa, H. Kanda, *Science* **292**, 1899 (2001).
3. M. Schreck, F. Hörmann, H. Roll, J. K. N. Lindner, B. Stritzker, *Appl. Phys. Lett.* **78**, 192 (2001).
4. K. Okano *et al.*, *Appl. Phys. A* **51**, 344 (1990).
5. I. Sakaguchi *et al.*, *Phys. Rev. B* **60**, 2139 (1999).
6. M. Stutzman, personal communication.

PERSPECTIVES: GENOMICS

Are There Bugs in Our Genome?

Jan O. Andersson, W. Ford Doolittle, Camilla L. Nesbø

For evolutionary biologists working on the exchange of genes between species (lateral gene transfer), the most exciting news from the human genome sequencing project has been the claim by the "public effort" (*1*) that between 113 and 223 genes have been transferred from bacteria to humans (or to one of our vertebrate ancestors) over the course of evolution. We, and probably many others wanting to test whether this result is really solid (*2*), have been beaten to the punch by Salzberg and colleagues (*3*). Their analysis, appearing on page 1903 of this week's issue, suggests that the actual number of bacterial genes in our genome may be lower than the predicted 223. These authors argue that there are other biologically plausible explanations besides lateral gene transfer that could account for the presence of bacterial genes in our genome.

Enhanced online at www.sciencemag.org/cgi/content/full/292/5523/1848

The claim for lateral gene transfer from bacteria into vertebrates (as exemplified by our own species' genome) was based on similarity searches. Such searches involve screening vertebrate genomes for sequences that are very similar to bacterial genes but are absent from other eukaryotic genomes. Genes shared by vertebrates and bacteria that are not found in other eukaryotes are considered to be probable bacteria-to-vertebrate transfers (BVTs). The 113 to 223 BVTs in question have significant similarity to bacterial sequences but no "comparable similarity" to genes in other completed eukaryotic genomes (*1*).

Salzberg *et al.* (*3*) now provide a careful reanalysis of these data (*1*) with a similar but more conservative approach that includes the addition of data from Celera's version of the human genome (*4*). As in the original study (*1*), the investigators' goal was to detect possible transfer of genes by analyzing gene distribution across taxa. They found 135 genes in the public effort's data set of 31,780 protein-encoding sequences (Ensembl proteome) and 89 genes in the Celera proteome of 26,544 proteins that were possible BVTs (*3*). This is similar to the public effort's conservative estimate of 113 possible bacterial genes in the human genome (*1*). Lateral gene transfer is not the only factor that could explain these results. For instance, differential gene loss (that is, random independent loss of genes in different eukaryotic lineages) may yield similar gene distribution patterns. The Salzberg *et al.* reanalysis demonstrates that the calculation of the number of bacterial genes in the human genome is highly dependent on how many nonvertebrate genomes were screened against the human genome. These authors found a downward trend in the number of BVTs observed when the human genome was screened against an increasing number of nonvertebrate genomes. Such a pattern is indeed

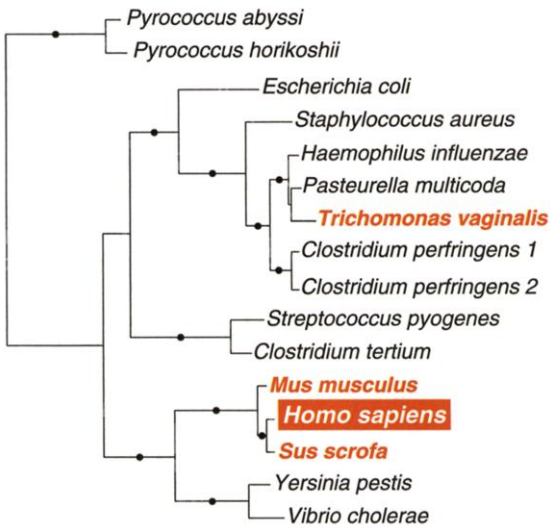
The authors are at the Canadian Institute for Advanced Research, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada. E-mail: joanders@is.dal.ca, ford@is.dal.ca, cnesbo@is.dal.ca

consistent with differential gene loss in eukaryotes, and it is reasonable to assume that the downward trend might continue as more nonvertebrate genomes become available for screening. Indeed, after removing possible BVTs that are more similar to genes in eukaryotic genomes for which only partial sequences are available, the number of possible BVTs drops to 114 and 68 for the Ensembl and Celera proteomes, respectively.

Another factor yielding false BVTs is differences in evolutionary rates among different vertebrate and nonvertebrate lineages. Salzberg *et al.* investigated the effect of differing evolutionary rates by relaxing the similarity criteria for inclusion of nonvertebrate homologs. Again, they found a reduction in the number of BVTs to 74 (Ensembl) and 56 (Celera). After additional trimming of the data—removing two mitochondrial genes, searching a further curated version of the Ensembl data, checking for annotation errors, and comparing the two data sets—the authors calculated the final number of possible BVTs to be 41 (Ensembl) or 46 (Celera).

So, the original description of 223 BVTs is probably overenthusiastic. But even 41 (or 46) BVTs is sufficient cause for excitement. The statistical arguments of the sort Salzberg *et al.* present can never eliminate the possibility that some of these BVT candidates (or others that they eliminated with their parsimonious broad-brush approach) really are true BVTs. In fact, although the Salzberg study criticizes conclusions about evolution based on similarity searches, this study, too, is based on a similarity search approach! The best way to determine whether some of the genes in the final list are real BVTs is to construct molecular phylogenetic trees for each of the possible instances. If a vertebrate gene sequence is nested within a robust cluster of bacterial sequences, the most probable explanation is that the vertebrate gene was laterally transferred from bacteria. Salzberg and colleagues mention that they have constructed phylogenetic trees for some genes (one is depicted in figure 2 of their paper), but they state that "most did not show patterns consistent with bacterial-to-vertebrate gene transfer." Yet they don't tell us how many "most" is, nor do they state which genes *do* show patterns consistent with BVTs.

We have prepared phylogenetic trees for seven of the genes listed in the supplementary information provided by Salzberg *et al.* Among these, we found one probable case of lateral gene transfer between bacteria and vertebrates: the gene encoding a putative *N*-acetylneuraminate lyase (see the figure). This gene was previously shown (very convincingly) to have been transferred from bacteria into the protozoan parasite *Trichomonas vaginalis* (*5*). Indeed, the gene found in *Trichomonas* is nested within a robust bacterial cluster, which would be expected for a gene that has undergone lateral transfer. The vertebrate version of this gene clusters unequivocally with *Vibrio cholerae* and *Yersinia pestis* genes, indicative of lateral gene transfer involving bacteria and eukaryotes that is independent of the bacteria-to-*Trichomonas* transfer. In this case, however, it is not possible to infer the direction of the transfer because neither the bacterial genes nor the vertebrate genes are nested within the other.



Gene swapping among friends and neighbors. Phylogeny of the gene encoding *N*-acetylneuraminate lyase (Ensembl ID IGI_M1_ctg1425_20). Phylogenetic relationships for this gene among prokaryotes, protozoans, and vertebrates were estimated using TREE-PUZZLE (*13*) from an amino acid alignment generated by CLUSTALX (*14*). Bullets on the nodes indicate that bootstrap support values—obtained using neighbor joining of PAM-based distances estimated in PHYLIP (*15*)—and puzzle support values were above 95%. Bacteria donated this gene to the protozoan parasite *T. vaginalis*. Vertebrates (human, mouse, and pig) together with two bacterial lineages (*Vibrio* and *Yersinia*) also show a branching pattern indicative of gene transfer, although it is not possible to infer the direction of the transfer.

Is lateral gene transfer between prokaryotes and eukaryotes an extremely rare event? Vertebrates are multicellular organisms, and so any evolutionarily stable incorporation of foreign DNA must take place in the germ cells that give rise to eggs and sperm. Unicellular eukaryotes, on the other hand, often live close to prokaryotes and frequently use them as food, which means that they have a much greater exposure to prokaryotic DNA than do vertebrate germ cells. Inevitably, this might lead to a gradual replacement of ancient eukaryotic genes with bacterial homologs (*6*). Indeed, pathogenic protozoa that live in environments rich in prokaryotes demonstrate that this does happen. *T. vaginalis*, a parasite of vertebrate epithelial cells, no doubt acquired its *N*-acetylneuraminate lyase gene from bacteria in its environment (see the figure) (*5*). The protozoan *Entamoeba histolytica* possesses fermentation enzymes that must have come from different anaerobic prokaryotes, and another protozoan, *Giardia lamblia*, encodes an enzyme in the mevalonate pathway that is unquestionably of bacterial origin (*7, 8*).

The fact that the ancestors of mitochondria and chloroplasts (DNA-containing cellular organelles) have contributed genes to the eukaryotic nucleus poses a serious problem for the detection of prokaryote-to-eukaryote lateral gene transfer. The endosymbiont bacteria that gave rise to mitochondria and chloroplasts have been a major source of bacterial genes in eukaryotic nuclear genomes, and their ancestral lineages are the α-proteobacteria and cyanobacteria, respectively. It seems sensible to infer that nuclear genes of similar ancestry are of endosymbiont origin, whereas those that cluster within other bacterial groups result from independent lateral transfers. However, independent transfers from α-proteobacteria and cyanobacteria subsequent to the original endosymbioses surely have also occurred, and these lineages themselves have been recipients of transferred genes (*9*). Reconstructing the history of prokaryote genes in the eukaryotic nucleus will be a formidable but exciting challenge (*10*).

To determine the extent to which lateral gene transfer is an important evolutionary force within eukaryotic evolution, we need to move beyond BLAST-based analysis to large-scale phylogenetic analysis. This is a realistic task: The methods are available, and several eukaryotic genome sequencing projects from a relatively broad range of eukaryotes are well under way (*11, 12*). We would not be surprised to see eukaryotes distributed in multiple places in a prokaryotic background. Such a finding could only be accounted for by multiple lateral gene transfer events (independent of genes donated by organelles). We expect, however, that the

vast majority of these transfer events happened before the evolution of multicellularity. Our multicellularity probably saved us from participating in the dirty business of lateral gene transfer so beloved by microbes.

References

1. International Human Genome Sequencing Consortium, *Nature* **409**, 860 (2001).
2. C. P. Ponting, *Trends Genet.* **17**, 235 (2001).
3. S. L. Salzberg, O. White, J. Peterson, J. A. Eisen, *Science* **292**, 1903 (2001); published online 17 May 2001 (10.1126/science.1061036).
4. J. C. Venter *et al.*, *Science* **291**, 1304 (2001).
5. A. P. de Koning, F. S. Brinkman, S. J. Jones, P. J. Keeling, *Mol. Biol. Evol.* **17**, 1769 (2000).
6. W. F. Doolittle, *Trends Genet.* **14**, 307 (1998).
7. J. Field, B. Rosenthal, J. Samuelson, *Mol. Microbiol.* **38**, 446 (2000).
8. Y. Boucher, W. F. Doolittle, *Mol. Microbiol.* **37**, 703 (2000).
9. T. Rujan, W. Martin, *Trends Genet.* **17**, 113 (2001).
10. C. G. Kurland, S. G. E. Andersson, *Microbiol. Mol. Biol. Rev.* **64**, 786 (2000).
11. T. Sicheritz-Pontén, S. G. E. Andersson, *Nucleic Acids Res.* **29**, 545 (2001).
12. TIGR Microbial Database (www.tigr.org/tdb/mdb/mdbinprogress.html).
13. K. Strimmer, A. von Haeseler, *Mol. Biol. Evol.* **13**, 964 (1996).
14. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, D. G. Higgins, *Nucleic Acids Res.* **25**, 4876 (1997).
15. J. Felsenstein, *Cladistics* **5**, 166 (1989).

Published online 17 May 2001;
10.1126/science.1062241
Include this information when citing this paper.

PERSPECTIVES: PLATE TECTONICS

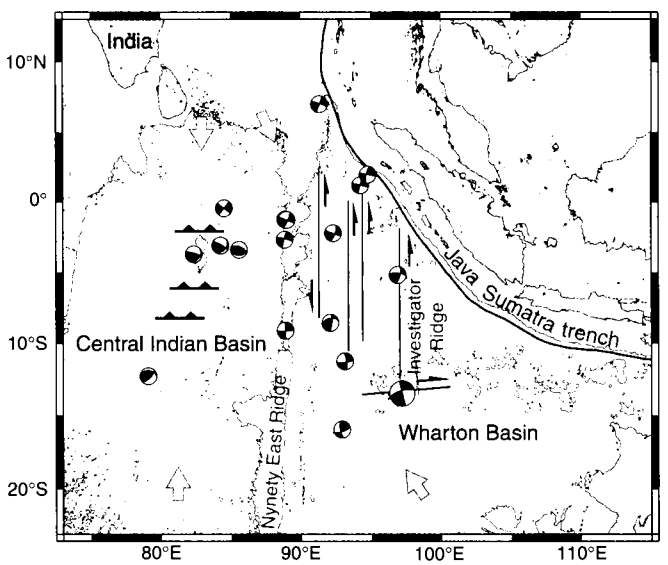
Indian Ocean Actively Deforms

Christine Deplus

According to plate tectonics, plates are rigid and deform only at their boundaries. There is, however, ample evidence of intraplate deformation in the Equatorial Indian Ocean. The deformation results from exceptionally high stresses in the oceanic lithosphere caused by the ongoing collision of India with Eurasia. Such intraplate deformation is commonly observed on the continents but not in the oceans, where deformation is usually localized at a narrow plate boundary rather than distributed over a wide area. The Indian Ocean thus offers a rare opportunity to study intraplate deformation of oceanic lithosphere.

Deformation in the Indian Ocean was first proposed in the 1970s (*1*) to resolve the difficulties in fitting global plate motion using a single rigid Indo-Australian plate. This traditionally defined plate is now considered as a composite plate, which includes three rigid or nearly rigid component plates and several deforming zones (*2*). The deforming zones may be described as diffuse plate boundaries, thus relaxing the fundamental assumption of plate tectonics that all oceanic plate boundaries are narrow (*2*, *3*). In the following, the term "intraplate" refers to earthquakes or deformation occurring within these large diffuse boundaries, with reference to the "classic" plate boundaries of the Indo-Australian plate.

The Equatorial Indian Ocean is known for its intraplate seismic activity and long-wavelength undulations in satellite-derived gravity data. Many earthquakes, with magnitude as large as 6 or 7, have occurred here during the last century (see the figure) (*4*). On 18 June 2000, an earthquake of magnitude 7.8 was registered in the Wharton Basin south of Cocos island (*5*). Earthquakes in the Central Indian Basin generally follow mechanisms different from those in the Wharton Basin. Their mechanisms indicate that the main compressive stress rotates from north-south in the Central Indian Basin to northwest-southeast in the Wharton Basin (*4*). In addition, the long-wavelength gravity undulations strike east-west in the Central Indian Basin and northeast-southwest in the Wharton Basin, in both cases roughly perpendicular to the compression axis. Numerical modeling of the Indo-Australian plate stress field (*6*, *7*) suggests that the rotation of the main compressive stress can be largely explained by the change in boundary conditions north of the Indo-Australian plate. The northward motion of India is resisted by the India-Asia collision, whereas the Wharton Basin freely subducts under the Java-Sumatra trench.



A large area of intraplate deformation. The compression axis in the Equatorial Indian Ocean (orange arrows) rotates from north-south in the Central Indian Basin to northwest-southeast in the Wharton Basin, yielding a different pattern of deformation in the two basins. Many large earthquakes have occurred during the last century [blue, Harvard CMT solutions (*13*); purple, from (*4*); pink, 18 June 2000 earthquake (*5*)]. In both basins, large-scale deformation may occur through buckling/folding perpendicular to the compression axis (long wavelength gravity undulations in light red). Brittle failure seems to occur along preexisting weakness directions (black), namely the north-south fracture zones and the east-west abyssal hill fabric.

The stress directions can thus be predicted reasonably well. The strain pattern is more difficult to assess, however, and requires field data from marine surveys or detailed source mechanism of recent earthquakes.

Most studies have focused on the Central Indian Basin. Here, tectonic deformation is characterized by long-wavelength (100 to 300 km) undulations of the oceanic basement—associated with the gravity undulations—and superimposed small-scale reverse faulting and folding of the crust and overlying sediments (*8*, *9*). We do not yet have enough data to determine the definitive cause of the long-wavelength undulations, but analog and numerical modeling suggests that it is caused by folding and buckling of the lithosphere (*9*, *10*).

Until recently, little was known about the deformation pattern in the Wharton Basin. The satellite-derived gravity undulations strike northeast-southwest and have wavelengths and amplitudes similar to those in the Central Indian Basin. They may again indicate folding and buckling of the lithosphere to accommodate northwest-southeast shortening. It remains unclear, however, whether the undulations are again associated with reverse faulting in the crust.

Some Wharton Basin earthquakes exhibit thrust mechanisms (indicative of reverse faulting), but most are strike slip. A marine

The author is at the Laboratoire Gravimétrie et Géodynamique, Institut de Physique du Globe de Paris and CNRS, 4, place Jussieu, 75252 Paris, France. E-mail: deplus@ipgp.jussieu.fr