

larly, the dipole moment surface determines the line intensities (3, 11).

In contrast to the *ab initio* calculation of line intensities, line positions from *ab initio* potential energy surfaces are still far from spectroscopic accuracy. Problems include the breakdown of the Born-Oppenheimer approximation (which enables the separation of nuclear and electronic motions) because of the light, rapidly moving hydrogen atoms. Furthermore, relativistic and quantum electrodynamic effects make nonnegligible contributions to the energy levels (7, 12).

Despite this somewhat discouraging news on the theoretical front, Partridge and Schwenke have shown that a small empirical adjustment of the high-quality *ab initio* potential surface yields calculated line positions to better than 0.05 cm^{-1} on average (11). This result is somewhat deceptive, however, because such spectroscopic accuracy is achieved for calculations within the range of data used to derive the correction only; the quality of the predictions deteriorates rapidly for line positions beyond the range of the available experimental data.

Callegari *et al.* (5) have devised a highly precise method for testing the reliability of dipole moment surfaces that does not require the measurement of line intensi-

ties. The trick is to apply an electric field to a sample of water vapor and monitor the splitting of the lines. The field causes each rotation-vibration level to split into components because of the partial removal of the degeneracy associated with spatial quantization of rotation.

This Stark effect is well known. Average dipole moments have previously been measured in the low-lying vibrational states of water (13). Callegari *et al.* (5) have now found a way to measure small Stark splittings in very weak water overtone transitions.

The authors use a sequence of three powerful laser pulses—the first to excite the overtone transition, the second to dissociate H_2O into $\text{H} + \text{OH}$, and the third to monitor the OH free radical—for the ultra-sensitive detection of visible water bands. Because the Stark splittings are relatively small compared to the laser line widths, the interference of two coherently excited molecular wave functions was monitored with the “quantum beat” method (14).

The work is a tour de force in modern chemical physics. As is often the case, the take-home message depends somewhat on your personal perspective. The good news is that the best current dipole moment functions for water are generally in good agreement with experimental observations.

However, discrepancies of up to about 10% indicate that there is room for improvement. Such work is already under way (5).

References

1. R. P. Wayne, *Chemistry of Atmospheres* (Oxford Univ. Press, Oxford, ed. 3, 2000), pp. 50–58.
2. V. Ramanathan, A. M. Vogelmann, *Ambio* **26**, 38 (1997).
3. D. W. Schwenke, H. Partridge, *J. Chem. Phys.* **113**, 6592 (2000).
4. R. B. Wattson, L. S. Rothman, *J. Quant. Spectrosc. Radiat. Transfer* **48**, 763 (1992).
5. A. Callegari *et al.*, *Science* **297**, 993 (2002).
6. W. Zhong, J. D. Haigh, D. Belmiloud, R. Schermail, J. Tennyson, *Q. J. R. Meteorol. Soc.* **127**, 1615 (2001).
7. P. F. Bernath, *Phys. Chem. Chem. Phys.* **4**, 1501 (2002).
8. F. Allard, P. H. Hauschildt, D. Schwenke, *Astrophys. J.* **540**, 1005 (2000).
9. O. L. Polyansky *et al.*, *Science* **277**, 346 (1997).
10. For example, the High resolution TRANsmission (HITRAN) Database of L. Rothman *et al.* (www.hitran.com).
11. H. Partridge, D. W. Schwenke, *J. Chem. Phys.* **106**, 4618 (1997).
12. P. Pytko, K. G. Dyall, A. G. Csaszar, O. L. Polyansky, J. Tennyson, *Phys. Rev. A* **63**, 024502 (2001).
13. S. Shostak, W. L. Ebenstein, J. S. Meunter, *J. Chem. Phys.* **94**, 5875 (1991).
14. This method detects splittings in lines by monitoring periodic time-dependent changes in a molecular property such as fluorescence caused by interference of the wave functions of the coherently excited energy levels.
15. M. Carleer *et al.*, *J. Chem. Phys.* **111**, 2444 (1999).
16. Thanks to R. Colin for assistance with the figure. Financial support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Space Agency (CSA), and NASA.

PERSPECTIVES: GENOMICS

Gene Duplication and Evolution

Michael Lynch

In the transmittance of genetic material from parent to offspring, accidents occasionally happen. Such accidents may result in the duplication of a chromosomal segment that then becomes separated from the original segment, ending up in a different chromosomal location. A number of human genetic disorders are known to be associated with the increased expression of genes contained within such duplications. However, evolutionary biologists have long been enthralled with the idea that duplicate genes could provide the ultimate substrate on which evolution could work. There are two ways in which gene duplication could generate a substrate suitable for adaptive evolution. Either one member of the duplicate gene pair could take on a new function, or two duplicate genes could divide the multiple functions of the ancestral gene between them, with natural selection then refining each copy to a

more restricted set of tasks (1). On page 1003 of this issue, Bailey and colleagues (2) contribute to our understanding of gene duplication by calculating the number of segmental duplications in the human genome.

The most common fate of duplicate genes appears to be the simple silencing of one member of the pair. The average time before silencing of one duplicate gene pair member is ~4 million years in animals (3, 4). By restoring the content of a genome to its original state, the silencing of duplicate genes has little direct effect on adaptive evolution. But, because either the ancestral or the descendant copy can be silenced, recurrent duplication of genes at unlinked chromosomal locations can passively give rise to small-scale chromosomal rearrangements (5, 6). When combined with geographical isolation, these small-scale gene rearrangements may contribute to the emergence of new reproductively isolated species.

Consider a pair of unlinked copies of an essential gene in an ancestral species. If, because of functional redundancy, one ran-

dom member of the duplicate gene pair is destined to become silenced in each population, there is a 50% chance that different copies will be silenced in two geographically isolated populations, thereby resulting in different chromosomal locations for the gene (see the figure). The contribution of this process to the evolution of genetic incompatibility between the two populations depends on the rate at which gene duplication takes place. Thus, it is noteworthy that all recently characterized eukaryotic genomes harbor substantial numbers of very young gene duplicates, many showing less divergence than gene copies (alleles) at the same chromosomal position (locus).

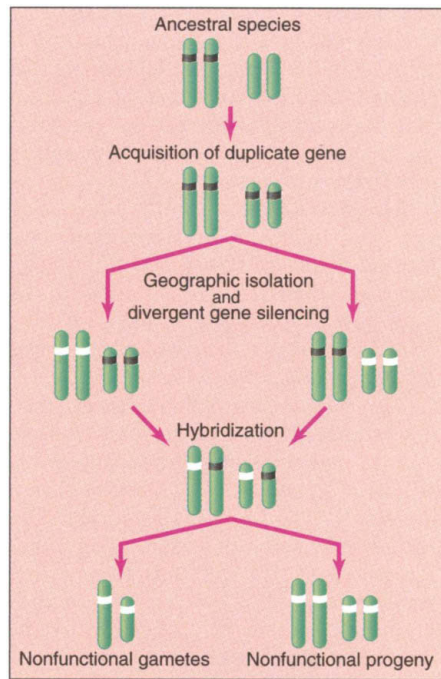
By applying demographic principles and genome sequence analysis to species-wide surveys of duplicate genes, scientists have calculated an average duplication rate of ~1% per gene per million years (3). This estimate holds up under a reanalysis of additional and better curated genome sequences (4). Using a new statistical approach to infer the presence of duplicated regions in pools of random human genome sequences, Bailey *et al.* estimate that at least 5% of the human genome consists of segmental duplications. They calculate that the span of the segmental duplications in the human genome ranges from tens to

The author is in the Department of Biology, Indiana University, Bloomington, IN 47405, USA. E-mail: mlynch@bio.indiana.edu

hundreds of kilobases, with >90% identity at the nucleotide level between ancestral and duplicated segments. Considering that roughly 50% of human gene duplicates have been silenced by the time the duplicated regions have diverged by 5% (3, 7), one can estimate that the high incidence of young gene duplicates in the human genome requires a duplication rate of at least 0.4% per gene per million years.

Given that DNA breakpoints occur randomly during the duplication process, some duplicated genes may lack key regulatory elements and/or regions of coding DNA. Thus, the rate of origin of functional gene duplicates may be somewhat less than 1% per gene per million years. Nevertheless, Bailey and colleagues show that a substantial fraction of human segmental duplications contain expressed genes, and another study suggests that at least 25% of the identifiable duplicate genes in the worm *Caenorhabditis elegans* are apparently still functional (8). So, we can now be fairly confident that many, if not all, eukaryotic genomes are subject to ongoing stochastic turnover of functional genes resulting from the origin and loss of small duplicated regions. The dynamics of this birth-death process are sufficiently pronounced that a random subset of 1 to 10% of the genes in a typical eukaryote are predicted to be in transient coexistence with their duplicates at any one time.

How does this level of genomic traffic affect the passive origin of incompatibilities between the genomes of different species? Let's be conservative and assume that only 0.1% of the genes in an ancestral species are in a duplicated state and subject to divergent silencing (that is, one member of the pair becomes silenced in one lineage, and its opposite becomes silenced in the other lineage). In this case, a genome containing ~20,000 genes (which approximates the number in many multicellular eukaryotes) would harbor the potential for ~10 changes in gene location (map changes) per genome over a period of a few million years. The continuous influx of new duplications keeps the process going, and with a duplication rate of only 0.1% per gene per million years, ~10 additional map changes are expected to accrue per population per million years of divergence. The consequences of these microchromosomal rearrangements for reproductive fitness (viability and fertility) of hybrid progeny depend on a number of factors, including the extent to which "absentee" alleles influence hybrid fitness, the degree of redundancy elsewhere in the genome, the involvement of autosomes versus sex chromosomes, and the level of gene expression in gametes (6). The salient point is that divergent silencing of just one pair of duplicates distributed on different autosomes results in



Divergent silencing of duplicate genes in sister species. The ancestral species may acquire a duplication of an essential gene (represented by the black bars on two pairs of parental chromosomes). Either copy of the duplicate gene pair (white bars) may be randomly lost in the two descendant populations. In this case, the first-generation (F_1) hybrid progeny of these two populations will contain two "absentee" alleles, one at each chromosomal locus. As a consequence of independent assortment, 25% of the gametes produced by these individuals will be entirely lacking an active copy of the original ancestral gene.

the complete absence of a functional allele in 25% of the gametes produced by a first-generation (F_1) individual (see the figure). And with just 10 divergently resolved loci, 94% of F_1 gametes will carry null genes at one or more loci. These conservative calculations, extrapolated from single-species analyses, suggest that chromosomal repatterning driven by gene duplication may be involved in the passive origin of isolating barriers on a time scale relevant to speciation (9).

More direct insight into the rate of microchromosomal repatterning will require fine-scale comparisons of closely related species. An impressive start in this direction has been made by Coghlan and Wolfe (10). By comparing the genomic sequences of *C. elegans* and its relative *C. briggsae*, they estimate that 4030 rearrangements have occurred over ~80 million years, implying ~25 rearrangements per genome per million years. Contributing to the total pool of rearrangements (in a 1:1:2 ratio) are reciprocal exchanges between two genomic locations, local inversions, and transpositions from one location to another. The vast majority of ge-

nomic rearrangements are quite small, involving segments that include five or fewer genes. Because the gradual replacement of one member of a duplicate gene pair by another does not impose a bottleneck in fitness on a population, many of these structural genomic changes may have arisen through an initial phase of gene duplication followed by loss of the ancestral locus. The potential for this process is quite high in *C. elegans*. With an estimated gene duplication rate of ~1.6% per gene per million years (3), this species experiences ~250 duplication events per genome per million years. Moreover, consistent with the types of rearrangements observed, a large fraction of the intrachromosomal duplications in *C. elegans* are inverted, and most duplication segments contain only one or two genes (11). Perhaps the high level of microchromosomal repatterning is peculiar to the *C. elegans* genome. However, extensive comparative genome mapping among the grasses paints a picture similar to that in worms (12), and an intermediate amount of small-scale rearrangement is apparent in the *Drosophila* genome (13).

In the early days of speciation genetics, models of chromosomal rearrangements were popular, particularly among plant biologists. However, recent studies (mostly in *Drosophila*) implicate changes within interacting genes that could be potential sources of genomic incompatibility between species (14). This shift in emphasis was motivated in part by the lack of obvious structural genomic rearrangements between many pairs of reproductively isolated species, as well as by perceived barriers to the full establishment of changes that cause meiotic problems in individuals heterozygous for large-scale chromosomal rearrangements. However, recent findings including those of Bailey *et al.* demonstrate how misleading large-scale chromosomal analyses can be. In terms of abundance, microchromosomal rearrangements are far more common than previously imagined, vastly outnumbering the macrochromosomal alterations favored by earlier studies. Containing only one or a few genes, such small-scale rearrangements are generally far too tiny to be revealed by visual inspection of chromosomes, and they may cause few, if any, problems during gamete production (meiosis).

Chromosomal rearrangements and nucleotide substitutions are by no means mutually exclusive contributors to the origin of reproductive isolation and the formation of new species. Indeed, the case has been made that chromosomal rearrangements can facilitate gene divergence (15), and the preservation of one member of a pair of duplicates by positive selection for a new function can promote the origin of changes in chromosomal location (16). Undoubtedly, some phy-

logenetic lineages will harbor genomic architectures that make them much more prone to speciation by genomic rearrangement than others (6), but at least one fundamental issue is now a bit clearer. With a rate of origin of duplicate genes on the order of 1% per gene per million years, the segmental duplication-rearrangement process is on a par with adaptive nucleotide changes within genes (genic evolution) as a mechanism for the origin of species incompatibilities. Consider that the average gene contains ~1000 nucleotides, that the mutation rate is ~0.1% per nucleotide per million years, and that the vast majority of mutations are either neutral or deleterious. In this case, the rate of nucleotide sequence changes per gene relevant to reproductive isolation and evolutionary adaptation could easily be less than the rate of gene duplication. Moreover, du-

plicated segments of DNA, even those not containing functional genes, contribute to chromosomal rearrangements in an indirect way—by serving as sites for nonhomologous recombination, thereby promoting secondary rearrangements (2, 7).

The recent observations of genome researchers raise two compelling issues for evolutionary biologists to ponder. First, discriminating between models of speciation that invoke negative epistatic interactions among genomes versus those that invoke microchromosomal rearrangements induced by gene duplication will require attention to the finest of details at scales of less than a few kilobases (the size of a typical rearrangement). Second, although the two big engines of evolution—adaptation and speciation—may be studied in isolation, they both are frequently interconnect-

ed through just one mechanism, that of gene duplication.

References and Notes

1. A. Force *et al.*, *Genetics* **151**, 1531 (1999).
2. J. A. Bailey *et al.*, *Science* **297**, 1003 (2002).
3. M. Lynch, J. C. Conery, *Science* **290**, 1151 (2000).
4. ———, in *Genome Evolution*, A. Meyer, Y. Van de Peer, Eds. (Kluwer Academic, Dordrecht, Netherlands, in press).
5. J. B. S. Haldane, *Am. Nat.* **67**, 5 (1933).
6. M. Lynch, A. Force, *Am. Nat.* **156**, 590 (2000).
7. R. V. Samonte, E. E. Eichler, *Nature Rev. Genet.* **3**, 65 (2002).
8. A. Mounsey *et al.*, *Genome Res.* **12**, 770 (2002).
9. J. A. Coyne, H. A. Orr, *Evolution* **51**, 295 (1997).
10. A. Coghlan, K. H. Wolfe, *Genome Res.* **12**, 857 (2002).
11. C. Semple, K. H. Wolfe, *J. Mol. Evol.* **48**, 555 (1999).
12. J. L. Bennetzen, W. Ramakrishna, *Plant Mol. Biol.* **48**, 821 (2002).
13. J. M. Ranz *et al.*, *Genome Res.* **11**, 230 (2001).
14. C.-I. Wu, *J. Evol. Biol.* **14**, 851 (2001).
15. L. H. Rieseberg, *Trends Ecol. Evol.* **16**, 351 (2001).
16. M. Lynch *et al.*, *Genetics* **159**, 1789 (2001).
17. I thank L. Rieseberg for helpful comments.

PERSPECTIVES: ASTRONOMY

Black Holes Reveal Their Innermost Secrets

A. C. Fabian and J. M. Miller

Black holes can be luminous sources of radiation. During accretion of matter from their surroundings, they release enormous amounts of gravitational energy as gas spirals through an accretion disk into the hole. Recent studies indicate that one component of this emission can provide unprecedented insights into the innermost regions of stellar mass black holes.

The emission spectrum of a black hole is dominated by a broad peak caused by thermal emission within the disk. Emissions extending to hard x-rays, which follow a power law, are produced by interactions between disk photons and hot electrons in a corona above the disk. This type of spectrum is observed for black holes of stellar mass (typically 5 to 15 times the mass of our Sun) in orbit around a normal star. It is also found for supermassive black holes (with typical masses of ~10⁷ solar masses) in the nuclei of galaxies. However, the temperatures in a disk around a supermassive black hole are cooler than around one of stellar mass.

A further feature, first clearly seen in the spectra of one class of accreting supermassive black holes, the Seyfert galaxies, is an iron emission line, produced by the

irradiation of the disk by the coronal radiation. In its simplest form, it is a fluorescent iron line at 6.4 keV. This line is distorted into a characteristically skewed frequency profile by the combined effects of Doppler shifts, relativistic beaming, and gravitational redshifts. The latter result from the emitting matter that orbits at high velocities deep in the gravitational potential well surrounding the black hole (1, 2). Such a line shape was originally used to describe the spectrum of a stellar mass black hole, Cygnus X-1 (see the first figure) (1, 3), but was first clearly resolved in the Seyfert galaxy MCG-6-30-15 (4, 5).

Until recently, evidence for such skewed emission lines from stellar mass black holes was very limited. Stellar mass black holes are much brighter because they are much closer than the supermassive black holes in Seyfert galaxies. But if the disk were highly ionized, little evidence for iron emission would be found. Now, however, the Chandra and XMM-Newton satellites have provided evidence for such lines in data from Cygnus X-1 (6) and other stellar mass black holes.

Reanalysis of archival data from recent

x-ray satellites such as BeppoSAX (6–9) is revealing further evidence for iron lines. Most previous workers misinterpreted these features as breaks in the continuum emission, or had serious problems with the detectors because the sources were too bright. It is of course possible that we are

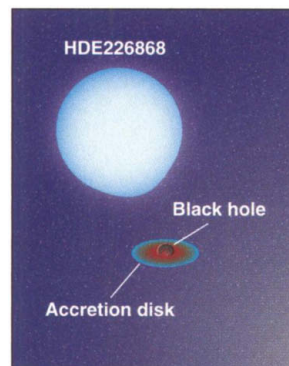
the ones who are misled. But the similarities between the stellar mass black hole lines now found and those from the Seyfert galaxies appear compelling.

Particularly exciting is the extent of the “red” (lower energy) wing on some of the lines, for example, in Cygnus X-1 and XTE J1650-500 (see the second figure). These lines indicate extreme gravitational redshifts, which can only arise when the emitting matter is extremely close [at about one gravitational radius (10)] to the black

hole’s event horizon, from beyond which no information can escape.

The iron lines are the most direct signature from the innermost region around a black hole. The frequency of quasi-periodic oscillations (QPOs) in the x-ray flux from some of these objects (11) may be tied to orbital motion and may also point to the inner regions. However, there is no clear consensus on what modulates the QPOs or why high-frequency QPOs are not observed continuously. They therefore do not yet represent as effective a diagnostic probe.

The shape of the broad iron line allows us to determine the inner radius of the emission and the emission profile on the disk. Sometimes it is steeply peaked inward (6, 12, 13), suggesting an additional



Artist's impression of the black hole Cygnus X-1 and its stellar companion.

A. C. Fabian is at the Institute of Astronomy, Cambridge CB3 0HA, UK. E-mail: acf@ast.cam.ac.uk
J. M. Miller is at the Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: jmm@space.mit.edu