evolutionary tree, based on nucleotide sequence data, for the singleton RNASE1 genes of 15 primates, plus the two RNASE1 genes of doucs. The langur genes are each other's closest relatives. The simplest explanation is that the RNASE1B gene in doucs arose as a recent duplication of the monkey's RNASE1 gene. The two langur genes each have an intron, and their introns are nearly identical. This suggests that the duplication arose by unequal crossing over.

Zhang and colleagues found nine amino acid substitutions distinguishing the proteins encoded by the douc langur's two RNASE1 genes. To see whether the enzymes had diverged in function, the researchers tested their RNase activity at a variety of pHs. The reaction norms appear in Figure 5.27b. RNASE1B is more active at the relatively low pHs characteristic of the douc langur small intestine. RNASE1, on the other hand, retains an optimal pH similar to that of other primate RNASE1s. In addition, the researchers found that douc RNASE1B has lost the ability to break down double-stranded RNA, a capacity seen in other RNASE1s and thought to play a role in defense against viruses.

These results are consistent with the hypothesis that following the duplication creating RNASE1B, the new gene evolved to encode an enzyme specialized for the digestive demands of the douc langur's unusual diet, while the parental gene retained its ancestral, generalist function. Similar duplications and specializations of RNASE1 genes appear to have occurred independently in ruminants (Zhang 2003) and African leaf-eating monkeys (Zhang 2006; Yu et al. 2010).

Genes that are duplicated within a genome and later diverge in function, like RNASE1 and RNASE1B in douc langurs, are described as **paralogous**. Paralogous genes can be contrasted with **orthologous** genes. These are genes that are derived from a common ancestral sequence and separated by a speciation event, like RNASE1 in douc langurs and RNASE1 in humans.

A New Gene Generated by Retroposition

An example of a new gene created by retroposition comes from the work of Heidi Parker and colleagues (2009) on dogs. The researchers sought to find the gene or genes responsible for chondrodysplasia, the short-legged condition characteristic of corgis, dachshunds, bassets, and a variety of other breeds **(Figure 5.28a)**. Scanning the dog genome for differences between breeds with and without chondrodysplasia, Parker and colleagues found, on chromosome 18, a duplicate copy of the gene for fibroblast growth factor 4 (*fgf4*). Possession of the duplicate is strongly associated with chondrodysplasia. Figure 5.28b sorts breeds with and without the condition into three categories. In most breeds all individuals tested either carried the duplicate on both copies of chromosome 18 (*DD*), or on neither copy (*NN*). A few breeds were polymorphic, meaning that different individuals had different genotypes. (The breeds classified as Other in which all dogs carried two copies of the duplicate—cairn terrier, Norwich terrier, and shih tzu—have short legs but do not meet a stringent definition of chondrodysplasia.)

The duplicate copy of *fgf4* lacks introns and is located some distance from the parent copy. These and other clues indicate that the duplicate arose by retroposition. Parker and colleagues found that in dogs that carry it, the duplicate gene is expressed in the joint-forming cartilage in the long bones of puppies, as is the parent copy. It appears that the duplicate acquired a promoter by serendipitously inserting into the middle of a transposable element. Fine control over when and where the encoded protein is actually made likely involves regulatory sequences in the untranslated portion of the mRNA.

(a) Who's a good girl?

(b) *Fgf4* retrogene genotypes in dogs of various breeds

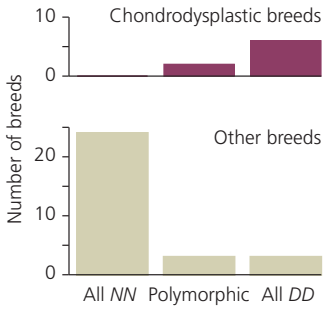

Figure 5.28 A new copy of the gene *fgf4* from retroposition (a) Like other Welsh corgis, Juno has chondrodysplasia—a condition characterized by short legs. (b) Across dog breeds, chondrodysplasia is strongly associated with the presence of a retroduplicated copy of the gene for fibroblast growth factor 4. Prepared using data from Parker and colleagues (2009).

All the chondroplastic breeds that Parker and colleagues tested carried the same duplicate gene at the same locus. It thus appears that the duplicate arose only once, before the development of today's breeds. It has subsequently been driven to high frequency in modern short-legged breeds by artificial selection.

Duplicated Genes and Evolution

Gene duplication is an important category of mutation for evolution. An increase in copy number may itself be adaptive. Duplication followed by divergence in function generates gene families.

Gene duplication accounts for a substantial fraction of the genomic variation among individuals, and thus the raw material for evolution. Recent estimates suggest that more of the genome is affected by copy number variation than by differences derived from point mutations (Mills et al. 2011). Variation among individuals in gene copy number can, by itself, serve as a substrate for adaptive evolution (Perry et al. 2007). And serial duplication followed by gene divergence has generated families of functionally related genes that, as in the case of olfactory receptor genes, can include hundreds of members (Young et al. 2008).

New Genes from Scratch

Although most new genes are born as duplicates of existing genes, some genes appear to be born from noncoding DNA. David Knowles and Aoife McLysaght (2009) found evidence for three examples in humans. We will briefly discuss one.

C22orf45 is a gene of unknown function unique to humans. It is transcribed in a variety of tissues, and its mRNA is known to be translated into protein.

Even though the gene is unique to humans, similar nucleotide sequences occur at the homologous locus in chimpanzees, gorillas, orangutans, gibbons, and macaques. The sequences in all of these nonhuman primates contain elements, including premature stop codons, that would substantially alter the encoded protein were the sequence transcribed and translated. One of the premature stop codons is shared by all five nonhuman species **(Figure 5.29)**.

Figure 5.29 A new human gene from noncoding DNA In other apes and in macaques, the sequence homologous to the human gene *C22orf45* contains a premature stop codon. From Knowles and McLysaght (2009).

The most parsimonious explanation for this pattern is that the sequence was noncoding in the last common ancestor of humans and the other five species, and became a coding gene in the lineage leading to modern humans.

Having discussed mutations that create new alleles and new genes, we now turn to the most drastic mutations: those that alter large portions of chromosomes or even the entire genome.

5.4 Chromosome Mutations

The mutations discussed thus far occur on the scale of a single base pair in DNA to segments containing tens of thousands of base pairs. These alterations pale in comparison to mutations that alter the gross morphology of chromosomes. Some of these mutations affect only gene order and organization; others produce duplications or deletions that affect the total amount of genetic material. They can also