Evolutionary Anthropology 15:127–131 (2006)

Coding Fregoli's Illusion

KENNETH M. WEISS

It's more than entertaining to be able to change your appearance on a moment's notice. It's a life saver that connects us in unexpected ways to our single-celled ancestors.

Leopoldo Fregoli (1867–1936) was a famous Italian performer with a gift for mimicry and impersonation. He could change his voice or face so fast that he amazed audiences, and he was celebrated on the stage in many countries. Fregoli was the original "quick-change artist," who made his living by always being somebody different.

Fregoli's trick was entertaining partly because it was so unusual. But a similar trick is widely found in nature, where it has nothing to do with entertainment. Many organisms make their living in much the same way. Indeed, your own life would be very different if you yourself weren't very good at doing much the same thing.

I'm not referring to protective coloration, but to a genetic trick that is quite different. It's Fregoli's illusion at the molecular level, and there are many different genetic ways to code for it. The trick, a variety of mechanisms, some newly discovered, in which genomes are used, reveals unexpected genetic similarities between vertebrates like ourselves and single-celled organisms.

EVOLUTION'S BIGGEST RAT RACE

Predatory-prey dynamics are an endless evolutionary race among detection, pursuit, escape, and dinner. Animals, including primates, use vision, sound, teeth, locomotion, and other aspects of their biology to obtain prey and to avoid becoming prey. When predators and prey are comparable in size, tactics, and generation time, we might say that the rat race is a fair one (though the rat might disagree). But not all predators can be detected or avoided with such tools.

Humans don't prey on microscopic organisms (except fermenters and their like!) but they certainly prey on us, a fact that presents serious problems. We cannot cognitively detect microscopic predators. Moreover, the challenge they present seems like a race not even a rat could ever win. A microbial intruder has a huge advantage. Although its lifespan is short, it reproduces very rapidly. An individual microparasite—a single bacterium, for example—may live for only a few minutes, but its rapidly dividing lineage, the infection as an aggregate macropredator, should easily be able to use natural selection to evade any defense a lumbering beast like a vertebrate could throw against it. But this is not what has happened. Instead, in a kind of spy-versus-spy game, big organisms have evolved genetic ways to fight parasites on their own terms. We have immune defenses that work on the same generational scale as do the assailants, and attack them where they can be "seen," at the molecular level. This contest has a direct genetic basis. But how can our linear DNA sequence, the one sequence we inherited, be used to recognize an intruder, any intruder, that we can't know about in advance?

A MOST EXCLUSIVE CLUB: ONLY ONE MEMBER AT A TIME

New genes arise by duplication events that, over evolutionary time, produce tandemly arrayed, bead-like copies of an original gene. All the genes subsequently accumulate mutational diversity.

Our immune system includes many components, but relevant here are the antigen receptor (Ag) gene complexes. A lymphocyte, a white blood cell, produces Ag proteins that sit on its surface membrane, where they detect and bind to antigens—intruding foreign molecules—that fit the receptor's chemical shape in lock-and-key fashion. An Ag protein (immunoglobulin) is assembled from a main "heavy" chain gene complex and one of two alternative "light" chain complexes. These are in different places in the genome, but each complex contains a tandemly arrayed string of adjacent related genes produced by a long history of duplication events (Fig. 2A). Each complex contains Variable, Diversity, and Joining components.

Three important things happen to express an Ag protein. The white cell randomly selects one of its two light chain regions and, by a process called allelic exclusion or monoallelic expression, also randomly picks either the paternally or maternally inherited copy of the chosen light chain and its heavy chain complex. The unselected complexes are not used. Then, within each complex, only one of the respective V, D, and J elements is selected, again more or less randomly. These elements are joined to form the messenger RNA (mRNA) that codes for that cell's Ag protein. The unselected

Ken Weiss is Evan Pugh Professor of Anthropology and Genetics at Penn State University.

DOI 10.1002/evan.20114
Published online in Wiley InterScience (www.interscience.wiley.com).

Figure 1. Leopoldo Fregoli, master of illusion. Presumably, he looks like himself in this photo. Reprinted with permission from www.magiebourgogne.com article 113 by Sebastien Bazou.

parts of the active Ag complexes are physically cut from their chromosome and discarded, and the lymphocyte is then able to make only that one variety of Ag protein. That, in turn, determines what antigens the cell can recognize.

At any given time, our blood stream is circulating lymphocytes with millions of different Ag variants. Even though they are assembled without knowledge of specific needs, no matter what foreign molecules get into us, at least one variant Ag molecule will grab onto the parasite, recognizing it specifically. The successful lymphocyte rapidly divides to intensify the attack, initiating a cascade of responses that halts the invasion.

All cells, not just lymphocytes, need proteins and other molecules on their surfaces in order to deal with the outside world. They become the parasite's Achilles heel, so to speak, because its surface molecules present a target for defense by the host. But there's more to the story, because the parasite that comes in after you will quickly find that you are after it. In response, parasites have evolved ways of using genes that make them quick-change artists so they can evade your immunological reconnaissance.

They do this in many ways that share a common feature. The surfaces of many kinds of parasites, including the trypanosomes that cause leishmaniasis and sleeping sickness, are coated with antigens. The plasmodium parasite responsible for malaria nestles cozily inside a red blood cell for part of its life cycle. The parasite secretes a protein that ends up not on its own surface, but on that of the red cell. This allows the red cell to adhere to other cells in the body,[2] which is good for the parasite, but it pays a price. These surface proteins advertise "I'm here," so that your immune system detects and destroys the invading cells, thus interrupting the parasite's cozy visit, ending your bout with disease, and spelling doom for the parasite.

The parasite must balance the positive functions of its surface proteins against the risk of giving the host's immune system a point of attack. Near the parasite's telomeres, the sequences that cap the ends of chromosomes, are clusters of related genes (Fig. 2B). They arose by gene duplication, just as our Ag complexes did. At any given time, one gene from one

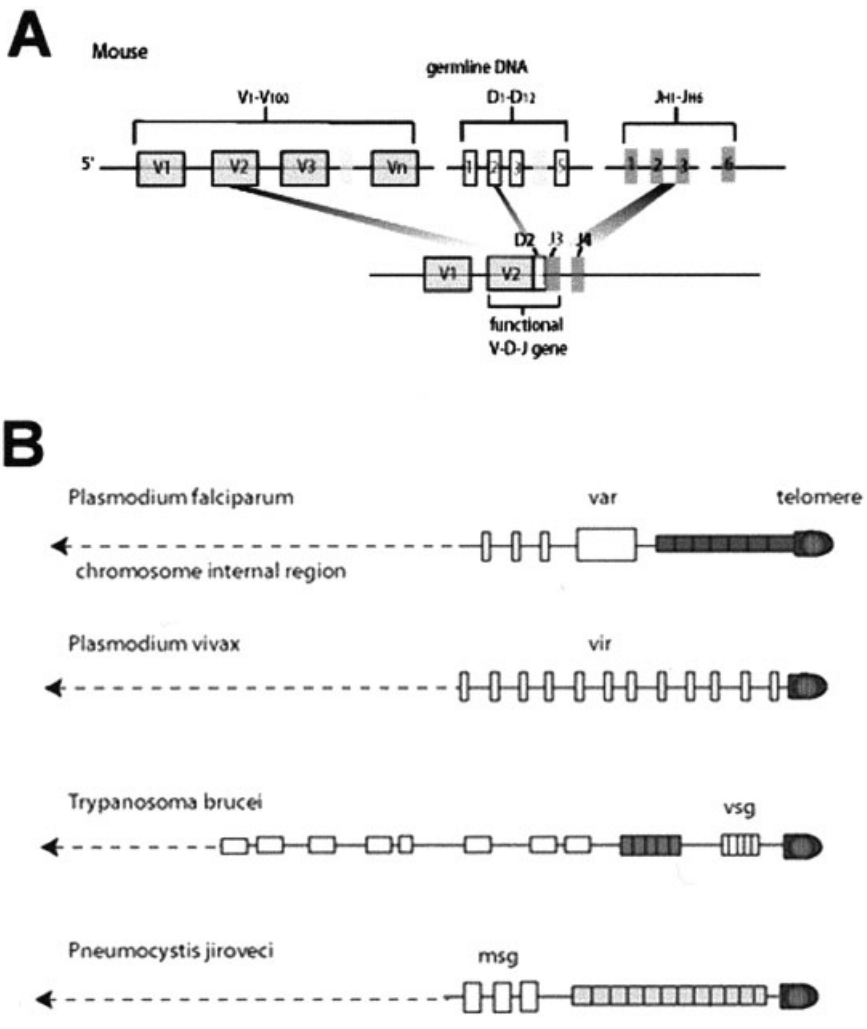

Figure 2. Weapons in the armory. A. Schematic of a mouse immune defense gene cluster complex (the human clusters are similar, and an additional part to the right is not shown). B. Examples of parasite telomeric clusters of antigenic surface protein genes (telomeres are the end units of chromosomes). By A. V. Buchanan; B after Scherf, Figueiredo, and Freitas-Junior.[1]

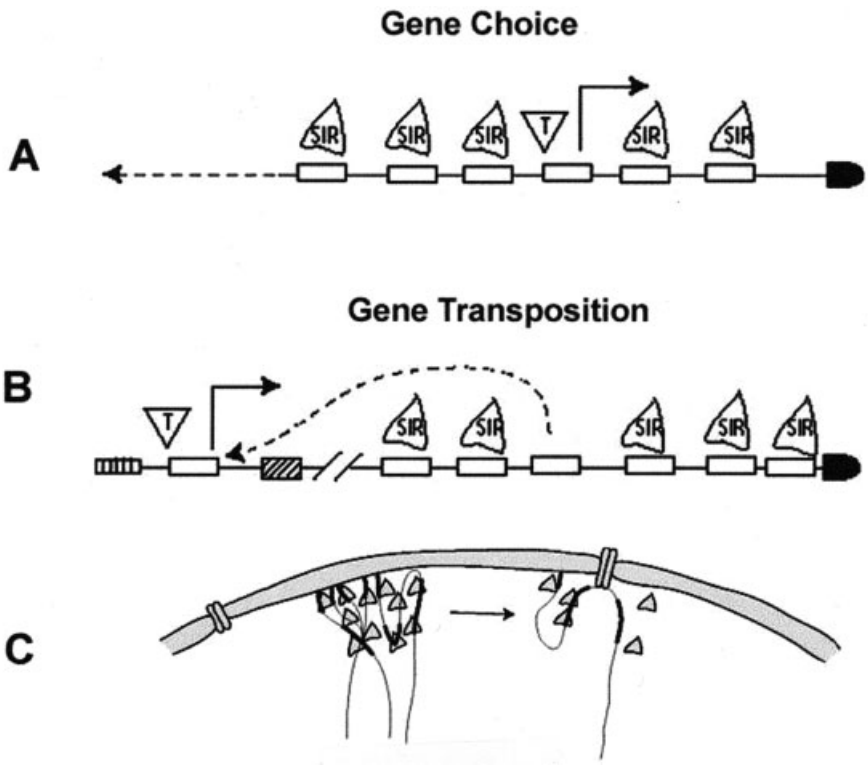

Figure 3. Parasite surface protein gene expression mechanisms. A. Repressor molecules (SIR protein figures) from a selected gene in a cluster are removed. The transcription protein (T-triangle) binds flanking sequences and transcribes mRNA from that gene (broken arrow). B. Repressor molecules leave a selected gene, but the gene is then physically copied to a separate expression cassette elsewhere in the genome. C. Telomeres cluster inside the nuclear membrane, where recombination generates variation among surface protein genes (thick bars); when one of the chromosomes moves to a nearby nuclear pore (arrow), the repressor proteins (triangles) fall off a lucky gene, which is transcribed. By AV Buchanan; part C modified from Ralph and Scherf.[3]

cluster is used to produce the parasite's surface protein. The other genes remain silent until, every so often, when the parasite divides, the newly made cell switches to a different surface gene. Just when your immune system has figured out what's there, the parasite's quick-change artistry makes the recognized protein disappear, and you have to start over. This is a challenge to vaccine development, which relies on dependable surface protein recognition. The switching is also, diabolically, just slow enough to allow you enough immune detection to keep from dying, which is good for you, but too fast for you to eliminate the infection, which is good for the parasite.

Telomeres are composed of many adjacent copies of chromosome-capping sequence elements.[4] Those elements are not related to immune defense. However, chromosomal regions with tandem arrays of very similar sequences are unstable and subject to frequent recombination, in which the similar elements are duplicated, deleted, or modified, generating highly variable and dynamic gene clusters. By huddling their surface protein genes near telomeres, parasites accumulate the high variability they need to stay one step ahead of your immune system. (Some parasite antigenic gene clusters are elsewhere in their genomes today, but may have originated near telomeres long ago.)

The way these genes are used varies among species (Fig. 3).[1,2,5,6] In plasmodia, trypanosomes, and giardias, once a specific cluster has been selected for use from one of its chromosomes, one lucky gene within the lucky cluster is expressed, which means that it is transcribed into mRNA, which is then translated into protein. Neighbors of that gene along the line are silenced. Fungi like *Pneumocystis jiroveci,* which cause pneumonia in immune-compromised patients such as those with HIV/AIDS, and some trypanosome surface genes, achieve the same end in a different way: The chosen gene is physically copied into a location elsewhere in the parasite's genome where the expression machinery is active, while the genes that remain behind are silent. About a hundred cell divisions later, a different gene is inserted into the expression cassette.

The mechanism for choosing which specific gene to use seems to involve the way chromosomes are physically packaged in the cluster regions. Members of a family of silent information regulator (SIR) or gene-silencing proteins help repress gene expression by covering DNA in a way that keeps the transcription mechanisms from making the contact they need to cause a gene to be transcribed. SIRs are known to be active near antigenic gene clusters.[7,8] In addition, in some parasite species, the telomeric regions of its different chromosomes aggregate inside its nuclear rather than its surface membrane (Fig. 3C). This aggregation fosters recombination between clusters on different chromosomes. And then, from time to time, just before cell division, a random one of these chromosomes leaves the cluster and moves to a nearby structure called a nuclear pore, an opening in the membrane that encloses the nucleus. That seems to be a location where the SIR and other repressor proteins are stripped away from one of the genes in the cluster at the end of that chromosome. With the repressors gone, the gene is now available to be expressed in the newly formed cell, switching from the gene the parent cell had been using.[3]

SOME HUMAN EQUIVALENTS

Gene-picking and monoallelic expression are not restricted to single-celled parasites, nor even to immune detection. The human genome uses similar mechanisms for other sorts of molecular recognition. The classical example is X-inactivation, by which the cells in a female embryo select one

of their two X chromosomes for expression.[9] The idea is that then cells in males, who only have one X, and cells in females, who have two but only use one, have the same dose levels of X-linked gene products. But inactivation isn't restricted to the X. Many genes on other chromosomes are chemically "imprinted," their nucleotides modified. This occurs differently in male and female gametes, so that only the copy inherited from one parent is used until the imprint is reset.

Sensory systems depend on allelic exclusion. We see color because of light-sensitive proteins coded by the opsin genes. Adjacent red and green opsin genes are on the X-chromosome; blue is on a different chromosome. Each cone cell in our retinas sends a color-specific signal to the brain because it uses either the red or green X-linked opsin on its active X or the blue gene, but not both. Our sense of smell works in a similar way. Our roughly 1,000 olfactory receptor genes (2,000 since we're diploid) code for odorant-detecting receptors on the surface of olfactory cells in the nose.[10] More than 20 different clusters of olfactory genes are scattered around our genome. Many are near telomeres and those that aren't there today may have started out there ancestrally. As each olfactory neuron forms, it selects only one of these two thousand genes, one cluster, one of the two chromosomes, and then one gene, the others remaining silent. Each olfactory cell presents only a single receptor to the outside world, so that when it detects an odorant molecule its signal to the brain is a specifically repeatable smell.

The fact that seeing and smelling depend on using only one surface receptor gene per cell may seem unrelated to the quick-change act of parasites, but the connection is actually rather striking. Just as in an expanding clone of an infection, where parasite cells change their clothes as they form, newly produced cells in your growing embryonic retina or olfactory epithelium pick the gene they'll use on their surfaces. They don't know they're part of larger organ system any more than parasites know they're part of an infection. In both cases, they're just individual cells trying to make a living, picking one gene from one copy of one open cluster of related genes.

AN UNSUSPECTED GENOMIC "BODY PLAN"

We were taught that our cells express both copies of all their genes. That's where the idea of dominance and recessiveness come from. X-inactivation was long regarded as a special case, but then there were Ag, imprinting, opsins, and olfaction. These are just the known indicators of what seems to be a grander plan. A substantial fraction of our genome contains known or suspected regions of monoallelic expression.[11–14] Indeed, the story is even more interesting than that: The human genome is littered with copies of a bit of DNA sequence called an L1 repeat element. L1s are more frequent in regions of monoallelic expression than in other regions. About 1,000 previously unsuspected candidate regions have been identified by this criterion.[14] L1s also seem to be involved in X-inactivation. In fact, surprisingly, only 75% of the X, an area dense with L1s, is inactivated. The cell uses both copies of the remaining 15%, while the final 10% have variable uses.[15] In other words, X-inactivation is only partial, in a replicable way, and may be just one instance of a gene-control mechanism that is far more widespread in our genomes than we thought.

There is also evidence that monoallelically expressed regions even on different chromosomes physically associate, possibly in a manner resembling the aggregation of surface protein genes in parasite telomeres shown in Figure 3C.[14,16,17] The known cases involve immune system genes and, once again, the association occurs when DNA is being replicated. This is interesting because it turns out that in every known case monoallelic expression is associated with asynchronous replication.[18] When a cell divides, some patterned sections of our genome are replicated first from one chromosome, maternal or paternal, before the other. The genes on the first-replicated copy are used; the corresponding genes on the later-replicating chromosome are silenced. As in X-inactivation, the choice is made early in embryonic development and is random between the chromosome copies. Once a "decision" is made, the same order of replication is maintained by that cell's clone of descendant cells in the embryo, just as in parasite gene-switching.[19]

The quick-change act for molecular recognition occurs across the deepest branches of the tree of life and may turn out to be part of a fundamental genomic "body plan." That plan goes something like this: A cell chooses a subset of its genes to use based on information from its environment. The chromosome regions containing those genes are opened for business; in diploids like primates, this sometimes also means only opening the region on the first of its two copies of the chromosome to be replicated. Then a specific gene within the opened region on the chosen chromosome is chosen for expression.

This is quite different from the way genome use is usually viewed. It implies that individual gene use is less independent and that genomes are more directly coordinated even between chromosomes than has been thought. (For geneticist readers, this means *trans* as well as *cis* regulation, new kinds of epistasis, and yet another source of variation among people with the "same" genotype.[20])

These facts can be seen to connect lives of single- and multi-celled organisms in a unifying way if we relax our concepts of cells and organisms. Diploid organisms like us and many parasites have extensive "random haploidy," our individual cells more like ancient ancestors represented by bacteria today, which have only single sets of chromosomes to work with. Across our genome, we are a cooperative mosaic of cells with individually haploid and diploid gene dose levels.

From its own point of view, each cell is an independent biological entity, a kind of organism, if you will. If one olfactory neuron dies, we, as aggregates of cells, might just lose a bit of our sense of smell, but if the cell death leads to failure to smell danger or dinner it could mean curtains for all our cells. It's similar for parasites. An infection is a kind of organism, a collection of mutationally differentiated cells descended from a single in-

Figure 4. Success through diversity. A poster of Fregoli's changes. Source: copyright from http://badigit.comune.bologna.it/bacer/1618.htm, reprinted with permission.

fecting cell. After all, what you are, too, is a collection of differentiated cells descended from a single cell, the fertilized egg. If one plasmodium fails to play the antigen expression game properly, it dies, but the infection persists. It's only if too many fail that the whole infection is put in jeopardy.

FROM ENTERTAINMENT TO SURVIVAL

Leopoldo Fregoli made his living through many changing identities (Fig. 4). Cells depend on a changing appearance as well. Through diverse mechanisms, they achieve the same task: using variation dynamically for an ever-changing surface identity.

One can view the diversity of quick-change mechanisms as an important general property of evolution at the gene level. Entirely unrelated genetic mechanisms that use the same logic—gene duplication, clusters of related genes, allelic exclusion, and gene switching—have evolved. They yield a very similar quick-change artistry for molecular recognition as their end result. An alternative view is that this may be evidence of a single ancient genomic "body plan," one that has been retained through evolution while all its details have changed.

Life is fleeting. We never know how (or if) we will be remembered. Fregoli himself is now forgotten, but his name lives on because his ability to appear to be different people on stage led to the naming of a psychiatric disorder after him. Those who suffer from Fregoli's delusion believe that different people are really the same person changing appearance or in disguise. We don't know the genetic code for Fregoli's delusion, but it is a breakdown in face-perception.

For pathogens, not being a quick-change artist leads to the ultimate bad fate. We're learning the codes pathogens use for their Fregoli illusion: being there, then not there. That talent forces us sluggish, slow-growing hosts to find the means to keep up and stay in the evolutionary race. And we have. Fregoli's fate was recognized bluntly on his tombstone, which is said to read "His last transformation." An organism with an identity that is too easy to read will quickly suffer its last transformation, too.

NOTES

I welcome comments on this column: kenweiss@psu.edu. I have a feedback and supplemental material page at http://www.anthro.psu.edu/weiss_lab/index.html. I thank Anne Buchanan, John Fleagle, and Sam Sholtis for critically reading this manuscript.

REFERENCES

1 Scherf A, Figueiredo LM, Freitas-Junior LH. 2001. Plasmodium telomeres: a pathogen's perspective. Curr Opinion Microbiol 4:409–414.

2 Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419:498–511.

3 Ralph SA, Scherf A. 2005. The epigenetic control of antigenic variation in *Plasmodium falciparum*. Curr Opinion Microbiol 8:434–440.

4 Weiss KM. 2004. Ponce de Leon and the telomere of youth. Evol Anthropol 13:82–88.

5 Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Bohme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC. 2005. The genome of the African trypanosome *Trypanosoma brucei*. Science 309:416–422.

6 Borst P, Ulbert S. 2001. Control of VSG gene expression sites. Mol Biochem Parasitol 114:17–27.

7 Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, Freitas-Junior LH, Scherf A, Crabb BS, Cowman AF. 2005. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. Cell 121:13–24.

8 Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, Montiel-Condado D, Ruvalcaba-Salazar OK, Rojas-Meza AP, Mancio-Silva L, Leal-Silvestre RJ, Gontijo AM, Shorte S, Scherf A. 2005. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell 121:25–36.

9 Weiss KM, Buchanan AV. 2004. Genetics and the logic of evolution. New York: Wiley-Liss.

10 Weiss KM. 2004. Perfume. Evol Anthropol 13: 205–210.

11 Ohlsson R, Tycko B, Sapienza C. 1998. Monoallelic expression: "there can only be one." Trends Genet 14:435–438.

12 Sano Y, Shimada T, Nakashima H, Nicholson RH, Eliason JF, Kocarek TA, Ko MS. 2001. Random monoallelic expression of three genes clustered within 60 kb of mouse t complex genomic DNA. Genome Res 11:1833–1841.

13 Singh N, Ebrahimi FA, Gimelbrant AA, Ensminger AW, Tackett MR, Qi P, Gribnau J, Chess A. 2003. Coordination of the random asynchronous replication of autosomal loci. Nat Genet 33:339–341.

14 Allen E, Horvath S, Tong F, Kraft P, Spiteri E, Riggs AD, Marahrens Y. 2003. High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. Proc Natl Acad Sci USA 100:9940–9945.

15 Carrel L, Willard HF. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400–440.

16 LaSalle JM, Lalande M. 1996. Homologous association of oppositely imprinted chromosomal domains. Science 272:725–728.

17 Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. 2005. Interchromosomal associations between alternatively expressed loci. Nature 435: 637–645.

18 Gimelbrant AA, Chess A. 2006. An epigenetic state associated with areas of gene duplication. Genome Res 16:723–729.

19 Mostoslavsky R, Singh N, Tenzen T, Goldmit M, Gabay C, Elizur S, Qi P, Reubinoff BE, Chess A, Cedar H, Bergman Y. 2001. Asynchronous replication and allelic exclusion in the immune system. Nature 414:221–225.

20 Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. 2002. Allelic variation in human gene expression. Science 297:1143.