

FAIR TESTS and MILEs

A Chromosome Fusion Extension

Much of the strength and confidence we have in good scientific hypotheses comes from these two measures.

Fair tests are observations or experiments that can be done for which the results could be consistent with the existing hypothesis, or inconsistent with it. Consistency strengthens the hypothesis, inconsistency weakens it.

MILEs (Multiple Independent Lines of Evidence) consist of two or more different kinds of evidence pointing to the same conclusion.

Study the two ape chromosomes that appear to have fused to form our #2 chromosome, based on the matching banding patterns. This alone provides a strong indication of the source of our #2 chromosome, but there are at least two **fair tests** that could be done that could confirm or weaken that hypothesis:

1. Is there evidence of the typical tandem repeats (found in telomeres) located in the region of expected fusion?

When you looked for this, what did you find? _____

2. Does the position of the centromere in our #2 chromosome coincide with its position in the corresponding ape chromosome?

When you look for this, what do you find? _____

There is at least one more test that you could look for that could confirm, or weaken, the argument for chromosome fusion. What is it?

3.

When we look for this, what do we find? _____

In a National Institutes of Health (NIH) News Release published 6 April 2005, announcing a study to be published in the 7 April 2005 issue of Nature, the following information was presented:

A detailed analysis of chromosomes 2 and 4 has detected the largest "gene deserts" known in the human genome and uncovered more evidence that human chromosome 2 arose from the fusion of two ancestral ape chromosomes, researchers supported by the National Human Genome Research Institute (NHGRI), part of the National Institutes of Health (NIH), reported today.

In the latest analysis, researchers searched the [human #2] chromosome's DNA sequence for the relics of the center (centromere) of the ape chromosome that was inactivated upon fusion with the other ape chromosome. They subsequently identified a 36,000 base pair stretch of DNA sequence that likely marks the precise location of the inactivated centromere. That tract is characterized by a type of DNA duplication, known as alpha satellite repeats, that is a hallmark of centromeres.

The complete News Release can be seen at: <http://www.genome.gov/13514624> (or it may be provided by your teacher).

Read the News Release, and answer the following questions:

How would you describe the effect of these **Fair Tests** on our original conclusion? _____
(strongly supportive, mildly supportive, not clear, slightly weakened, very weakened)

How many **MILEs** do we have favoring chromosome fusion as the source of our #2 chromosome? _____

What does all of this tell you about human ancestry, and how does it affect your confidence in this?

NIH NEWS RELEASE:

Scientists Analyze Chromosomes 2 and 4

NHGRI-Supported Researchers Discover Largest "Gene Deserts"; Find New Clues to Ancestral Chromosome Fusion Event

BETHESDA, Md., Wed., April 6, 2005 - A detailed analysis of chromosomes 2 and 4 has detected the largest "gene deserts" known in the human genome and uncovered more evidence that human chromosome 2 arose from the fusion of two ancestral ape chromosomes, researchers supported by the National Human Genome Research Institute (NHGRI), part of the National Institutes of Health (NIH), reported today.

In a study published in the April 7 issue of the journal *Nature*, a multi-institution team, led by Washington University School of Medicine in St Louis, described its analysis of the high quality, reference sequence of chromosomes 2 and 4. The sequencing work on the chromosomes was carried out as part of the Human Genome Project at Washington University; Broad Institute of MIT, Cambridge, Mass.; Stanford DNA Sequencing and Technology Development Center, Stanford, Calif.; Wellcome Trust Sanger Institute, Hinxton, England; National Yang-Ming University, Taipei, Taiwan; Genoscope, Evry, France; Baylor College of Medicine, Houston; University of Washington Multimegabase Sequencing Center, Seattle; U.S. Department of Energy (DOE) Joint Genome Institute, Walnut Creek, Calif.; and Roswell Park Cancer Institute, Buffalo, N.Y.

"This analysis is an impressive achievement that will deepen our understanding of the human genome and speed the discovery of genes related to human health and disease. In addition, these findings provide exciting new insights into the structure and evolution of mammalian genomes," said Francis S. Collins, M.D., Ph.D., director of NHGRI, which led the U.S. component of the Human Genome Project along with the DOE.

Chromosome 4 has long been of interest to the medical community because it holds the gene for Huntington's disease, polycystic kidney disease, a form of muscular dystrophy and a variety of other inherited disorders. Chromosome 2 is noteworthy for being the second largest human chromosome, trailing only chromosome 1 in size. It is also home to the gene with the longest known, protein-coding sequence - a 280,000 base pair gene that codes for a muscle protein, called titin, which is 33,000 amino acids long.

One of the central goals of the effort to analyze the human genome is the identification of all genes, which are generally defined as stretches of DNA that code for particular proteins. The new analysis confirmed the existence of 1,346 protein-coding genes on chromosome 2 and 796 protein-coding genes on chromosome 4.

As part of their examination of chromosome 4, the researchers found what are believed to be the largest "gene deserts" yet discovered in the human genome sequence. These regions of the genome are called gene deserts because they are devoid of any protein-coding genes. However, researchers suspect such regions are important to human biology because they have been conserved throughout the evolution of mammals and birds, and work is now underway to figure out their exact functions.

Humans have 23 pairs of chromosomes - one less pair than chimpanzees, gorillas, orangutans and other great apes. For more than two decades, researchers have thought human chromosome 2 was produced as the result of the fusion of two mid-sized ape chromosomes and a Seattle group located the fusion site in 2002.

In the latest analysis, researchers searched the chromosome's DNA sequence for the relics of the center (centromere) of the ape chromosome that was inactivated upon fusion with the other ape chromosome. They subsequently identified a 36,000 base pair stretch of DNA sequence that likely marks the precise location of the inactivated centromere. That tract is characterized by a type of DNA duplication, known as alpha satellite repeats, that is a hallmark of centromeres. In addition, the tract is flanked by an unusual abundance of another type of DNA duplication, called a segmental duplication.

"These data raise the possibility of a new tool for studying genome evolution. We may be able to find other chromosomes that have disappeared over the course of time by searching other mammals' DNA for similar patterns of duplication," said Richard K. Wilson, Ph.D., director of the Washington University School of Medicine's Genome Sequencing Center and senior author of the study.

In another intriguing finding, the researchers identified a messenger RNA (mRNA) transcript from a gene on chromosome 2 that possibly may produce a protein unique to humans and chimps. Scientists have tentative evidence that the gene may be used to make a protein in the brain and the testes. The team also identified "hypervariable" regions in which genes contain variations that may lead to the production of altered proteins unique to humans. The functions of the altered proteins are not known, and researchers emphasized that their findings still require "cautious evaluation."

In October 2004, the International Human Genome Sequencing Consortium published its scientific description of the finished human genome sequence in *Nature*. Detailed annotations and analyses have already been published for chromosomes 5, 6, 7, 9, 10, 13, 14, 16, 19, 20, 21, 22, X and Y. Publications describing the remaining chromosomes are forthcoming.

The sequence of chromosomes 2 and 4, as well as the rest of the human genome sequence, can be accessed through the following public databases: GenBank (www.ncbi.nih.gov/Genbank) at NIH's National Center for Biotechnology Information (NCBI); the UCSC Genome Browser (www.genome.ucsc.edu) at the University of California at Santa Cruz; the Ensembl Genome Browser (www.ensembl.org) at the Wellcome Trust Sanger Institute and the EMBL-European Bioinformatics Institute; the DNA Data Bank of Japan (www.ddbj.nig.ac.jp); and EMBL-Bank (www.ebi.ac.uk/embl/index.html) at EMBL's Nucleotide Sequence Database.

NHGRI is one of the 27 institutes and centers at NIH, an agency of the Department of Health and Human Services. The NHGRI Division of Extramural Research supports grants for research and for training and career development at sites nationwide. Additional information about NHGRI can be found at www.genome.gov.

For more information, contact:

Geoff Spencer, NHGRI
(301) 402-0911
spencerg@mail.nih.gov

Michael C. Purdy, Washington University
(314) 286-0122
purdym@wustl.edu