1. Why did the authors write this paper?

   To propose the hypothesis that genes had been “laterally” transferred to humans.

   To compute the E-values for the BlastP matches to the proteins from the human proteome.

   To propose a plausible alternative to the hypothesis that genes had been “laterally” transferred to humans.

   To identify species that share more genes in common between humans and bacteria.

2. What is “lateral gene transfer”?

   When genetic material is passed from the genome of one organism to another through a process other than reproduction.

   When genes are transferred out of the nucleus and into the cell.

   When genetic material is transferred from RNA to proteins.

   When genetic material is passed between one human and another human.

3. Why is lateral gene transfer (LGT) from bacteria to humans unlikely?

   Bacterial genes are so different from human genes that a human could not survive with bacterial genes in his/her chromosomes.

   Because a bacterium would have to infect a germline cell, enter the nucleus of that cell, and insert some of its DNA into one of the host’s chromosomes, after which the mutation would then have to provide an evolutionary advantage to spread through the population.

   It is not unlikely; in fact, LGT has occurred and it is an ongoing process in the human population**.**

   Because bacteria never actually enter human cells during an infection.

4. What are homologs?

   Identical mutations that occurred over evolutionary time.

   Two copies of a gene in different organisms that share a common ancestor.

   Genes that perform similar biochemical functions but have different sequences.

   Genes that have evolved to have the same function but different sequences.

5. What was the main method used to rule out lateral gene transfers between humans and bacteria?

   If genes were found to have mutated between eukaryotic genomes and human genomes.

   If a homolog of a gene was found in prokaryotic genomes.

   If a homolog of a gene found in humans was also found in a species of nonvertebrate eukaryotes.

   If a homolog of a eukaryote was found in another eukaryote but not in humans.

6. Why would this method rule out lateral gene transfers?

   Genes that appear in nonvertebrate eukaryotic organisms and humans are genes that must have been shared in common well after humans and bacteria diverged, so these genes probably weren't transferred directly to humans.

   Lateral gene transfer is a more likely explanation only when organisms that are close together evolutionarily share homologous genetic material.

   Nonvertebrate eukaryotic organisms and bacteria are evolutionarily "closer" than invertebrate eukaryotic organisms and humans. If they share a homologous gene, then bacteria are likely to have passed genes directly to humans.

   Humans and bacteria are both likely to have shared an evolutionary history with nonvertebrate eukaryotic organisms, so genes are likely to be homologous across all three.

7. What are the biological, computational, and statistical parts of Figure 1?

   Biological: the argument that a Blast cutoff of 10^-10 should define homologs

   Computational: The identification of homologs by performing Blastp searches on known protein sets.

   Statistical: Observing and quantifying the trend in genes shared versus genome sample size.

   Biological: the argument that lateral transfer should be ruled out if there is a human/nonvertebrate eukaryote homologs.

   Computational: The identification of homologs by performing Blastp searches on known protein sets.

   Statistical: The calculation of the standard error for the sample size curves.

   Biological: the argument that lateral transfer should be ruled out if there is a human/nonvertebrate eukaryote homologs.

   Computational: The identification of homologs by performing Blastp searches on known protein sets.

   Statistical: Observing and quantifying the trend in genes shared versus genome sample size.

   Biological: the argument that gene should be ruled out if there is a human/nonvertebrate eukaryote homologs.

   Computational: Observing and quantifying the trend in genes shared versus genome sample size.

   Statistical: The identification of homologs by performing Blastp searches on known protein sets.

8. What are the biological, computational, and statistical parts of Figure 2?

Biological: the argument that lateral gene transfer is less common than standard gene flow through reproduction

Computational: The storage of data in a low redundancy protein database.

Statistical: The inference that humans cluster more closely (have smaller distances to) other eukaryotes than to bacteria.

Biological: the argument that lateral gene transfer is less common than standard gene flow through reproduction.

Computational: The calculation of statistical significance of the protein hits in the Blastp search.

Statistical: The statistical modeling of protein sequences via a Markov Model.

Biological: the argument that lateral gene transfer is less common than standard gene flow through reproduction

Computational: The identification of homologs of human HAS genes by iterative BlastP searches and application of the neighbor-joining algorithm to create the phylogenetic tree.

Statistical: The inference that humans cluster more closely (have smaller distances to) other eukaryotes than to bacteria.

Biological: the argument that proteins should have more similar sequences if they are evolutionarily closer.

Computational: The identification of homologs of human HAS genes by iterative BlastP searches and application of the neighbor-joining algorithm to create the phylogenetic tree.

Statistical: The inference that humans cluster more closely (have smaller distances to) other eukaryotes than to bacteria.

9. The analysis in this paper required multiple data sources. Which of the following data sources was not used in the paper?

    The complete set of noncoding RNA genes from the human genome.

    The complete set of genes from the fruit fly, nematode worm, yeast, and mustard weed genomes.

    The set of all known genes (at the time) from the malaria parasite, Plasmodium falciparum.

    The set of all known genes (at the time) from completed bacterial genomes.

10. In the end what is the conclusion of the paper?

    That increasing the number of sequenced genomes is likely to increase the number of potential lateral gene transfer events.

    That a more plausible explanation for the observation of homologous genes found in bacteria and humans but not in non-vertebrate eukaryotes is gene loss and low sample size.

    That the argument for lateral gene transfer is statistical because we must average over multiple possible transfer events.

    That genes are more likely to be laterally transferred from certain types of bacteria to humans.