CBS 530 Assignment No 2

SHUBHRA GUPTA
shubhg@asu.edu
993755974

Review of the papers:
Construction and Analysis of a Human-Chimpanzee Comparative Clone Map and
Intra- and Interspecific Variation in Primate Gene Expression Patterns

Abstract

In this report the author gives her own view on the recent articles of Fujiyama et al[1] and Enard et al[2]. The author shall try to cover the questions that the work of this paper addressed, a critique of the appropriateness of the methodology, the primary results that relate to the question and the discussion of whether the data answer the question.

1 Introduction

As human genome project is going on most of the scientists and researchers are working on it and they are looking for very minutely genetic things. The recently released human genome sequences provide reference data to conduct comparative genomic research on primates. As in the first paper Fujiyama et al[1] showing by experiment 98.7% alikeness between humans and chimpanzee based on human genome sequences. They have presented a first-generation human-chimpanzee comparative genome map. Human genome sequences and comparative genomics has become a powerful approach to extract genetic information from large stretches of nucleotide sequences. Genomic information is also the most valuable resource for understanding the genetic differences between species and a basis for deciphering how genome information is produced into phenotypes. Although human and their closest evolutionary relatives, the chimpanzee, are 98.7% identical in their genomic DNA sequences, but they are differ in morphological, behavioral and cognitive aspect. Why these differences are occurring, at the same time Enard et al[2] is searching for these difference which altered gene expression of mammals. In addition, humans and chimpanzee differ in several other traits, that are of medical interest such as susceptibility of AIDs, epithelial neoplasms, malaria and Alzheimer's disease. Although it was pointed out 25 years ago that many of these differences may be due to quantitative differences in gene expression, rather than structural changes in gene products.
The objective of first and second paper is to determine genomic(what genes are expressing) similarity and genomic differences at transcriptome(what the mRNA expression) and proteome (gene expression at protein level).

2 Questions Related To Work

What type of similarity and differences between humans and other mammalians at genetic level? Why they have compared humans with chimpanzee, and both with other rodent species? Why the X and Y chromosome coverage is so lower as compared with the other chromosomes? What is qualitative and quantitative effect of mRNA and protein levels on all the species? How they compared leukocytes blood, liver and brain of humans, chimpanzee, orangutan, and macaques using micro arrays? what is protein expression patterns which is used for humans and chimpanzee? What procedure they used to made BESs and to construct a human-chimpanzee comparative map? What is percentage of coverage in human genome and chromosomes? How they studied the apparent acceleration which is likely due to nucleotide sequence differences between the apes and humans and also in rodent.

3 Results Related To Question

Because chimpanzees are our closest relatives, the differences between us less than any other species. Thus comparison between humans and chimpanzees are most efficient and effective approach to understand what makes us human.

In the experiment they produced total BESs $114,421$ from which $77,461$ Numbers of BESs, in which $49,160$ BESs formed paired ends and $28,301$ BESs formed singleton, having an alignment longer than 50bp with $\geq$ 90% identity. The remaining were not mapped to the human genome were categorized into 3 different classes (1) either repeats (2) low and similarity hits or (3) not found in current data base. The BESs mapped with high confidence were used to calculate the difference between the chimpanzee and human genomes at nucleotide level. Although most of the BESs have higher identity, they found the existence of many low-identity BESs in the genome.

Low coverage of chromosome X can be explained by the haploid status of the chromosome in the chimpanzee BAC libraries and the Y chromosome coverage is so much lower (4.8%) they have given one possibility and another hypothesis. For complete result they said we will have to wait until sequencing of the human Y chromosome is finished. They also need an independent approach, to construction and analysis Y chromosome of chimpanzee.

They found that 48.6% of the whole human genome was covered by the chimpanzee BAC and reason behind this low coverage of chromosome is that they used very stringent condition for the calculation that is BAC clones were incorporated, when they had 2 sequenced end in the same NT counting. The coverage for all paired and X,Y chromosomes. Coverage for chromosome $14, 20, 21, 22$ was substantially higher.

They took for brain sample 7 humans 4 chimpanzee and 2 macaque, for liver 6 humans, 5 chimpanzee and 4 macaque, and for blood each of 10 using the Affymetrix array. They studied mRNA expression levels change evolutionary for blood and liver human expression patterns are more similar to those of chimpanzee than to macaque, furthermore rate of change on the lineages

for blood are equal the chimpanzee and humans, but 1.3fold difference in liver expression pattern in the chimpanzee brain is more similar to the macaque than to humans. However in rodent this acceleration is of similar magnitude in brain and liver.

For quantitative and qualitative effect 2 types of differences were found (1) shifts in the migration position of protein, which represent a shift in size or change of the protein that is changes in amino acid sequence. (2) differences in quantity of protein without a shift in migration position which represent differences in amount of protein and relative amount of qualitative protein differences in mammalians are similar and in rodent the quantitative protein difference s are similar to the qualitative differences. Whereas in human and chimpanzee brain quantitative difference is 6 times more than qualitative differences.

4 Methodology and Results

They present the construction and analysis of a first generation human chimpanzee comparative genomic map based on the alignment of chimpanzee bacterial artificial chromosome(BAC) end sequences(BESs) to human genomic sequences which is obtained from the public database.

They analyzed the chimpanzee-human relationship of 21 chromosome by combining the BES mapping information with a sequence tagged site(STS).To identity the boundaries of possible genome rearrangements they went through candidate clones containing chromosomal breakpoints and they found PTB-053J22 BESs,which they deducted through the study,contains one of the breakpoints corresponding to the human chromosomal inversion.

Their results shows that the large number of quantitative changes in gene expression can be detected between closely related mammals. The underlying reasons for such expression differences have various form like, duplication and deletion of genes, promoter changes, changes in level of transcription factor and changes in cellular composition of tissue.

To test mapping procedure they took 15 chimpanzee BAC clones mapped to human chromosome(1−8) by the BES alignment procedure subjected to FISH analysis, among them 13 clones showed single locus signals and 2 clones showed similar signals at 2 loci on the human and chimpanzee chromosomes.

They studied transcriptome and proteme to understand, the evolution of the mammalian in different tissue of humans, chimpanzee, orangutan and macaque and for comparative purpose they performed similar studies in rodent species. They compared mRNA level in brain and liver of 3 male humans, 3 male chimpanzee and 1 male orangutan using Affymetrix array were performed for each individual and analysis independently and for all possible pairwise comparison among the 6 humans, 6 chimpanzee and 2 orangutan they found that for the brain and liver samples, the distance measured in between any two duplicate from the same individual is less than the distance measured in between individuals and result shows that the variation in gene expression measured between individuals within the species is substantial relative to the difference between

humans and chimpanzee. The amount of gene expression differences shared among all humans is larger than those shared among all chimpanzee and orangutan is further removed from human and chimpanzee.

Does observations made among the primate species are typical in mammals. They studied 3 mouse species Mus spretus(SPR), M.caroli(CAR), M musculus(MUS) they did the same procedure as in primates. They found when the more distinctly related CAR is into account,all MUS and SPR individual share gene expression pattern that separate them from the other species as in human and chimpanzee. When they compared differences between species it found MUS and SPR 2.1 fold in brain and 2.3 fold in liver.

They tested the human draft sequences to some position for chromosome which they were retrieve from public portion of the celera data. They observed no substantial difference between the mapping results obtained through the public database and celera. They found acceleration is larger in brain than in liver in human, raising the possibility that gene expression pattern may have changed more in the brain than in liver.

5 Future Work and Critique

They told that the exact position of these clones will become clearer on the progress of the human genome sequencing and/or sequences of the chimpanzee clones in future.

The region around deletions or mutation sites remains to be primary target of further investigation for future.

They believed that the whole genome chimpanzee/human comparative map built by the BES alignment procedure is accurate and useful for future studied. However, the quality of the map and its usefulness should improve the finishing of the human genome sequence.

A challenge for the future is to investigate the relative contributions of these factors to the expression differences.

In the first paper their calculation comes from a much larger random comparison, and for most of the results they are waiting for future results, but whatever they got are good.

In second paper whatever results they got are good.

How accurate these results are? Can we believe on the accuracy of softwares which they have used for experiment?

References

[1] A Fujiyama, H. Watanabe, A. Toyoda, T.D. Taylor, H.-S. park, M.-L.Yaspo, H. Lehrach, Z. chen, G. Fu, N. Saitou, K. Osoegawa, P.J.D. Jong, Y. Suto, M. Hattori, Y. Sakaki. *Construction and Analysis of a Human-Chimpanze Comparative Clone Map* Science **295** (January 2002) pp 131–134.

[2] W Enard, P. Khaitovich, J. Klose, S. Zöller, F. Heissing, P. Giavalisco, K. Nieselt-Struwe, E. Muchmore, A. Varki, R. Ravid, G.M. Doxiadis, R.E. Bontrop, S. Pääbo, *Intra- and Inter-specific Variation in Primate gene Expression Patterns* Science **296** (April 2002) pp 340–343.