Genetic Differences Between Humans and Great Apes – Implications for the Evolution of Humans.

Ajit Varki

Departments of Medicine and Cellular & Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla CA, U.S.A.

Abstract. At the level of individual protein sequences, humans are 97-100% identical to the great apes, our closest evolutionary relatives. The evolution of humans (and of human intelligence) from a common ancestor with the chimpanzee and bonobo involved many steps, influenced by interactions amongst factors of genetic, developmental, ecological, microbial, climatic, behavioral, cultural and social origin. The genetic factors can be approached by direct comparisons of human and great ape genomes, genes and gene products, and by elucidating biochemical and biological consequences of any differences found. We have discovered multiple genetic and biochemical differences between humans and great apes, particularly with respect to a family of cell surface molecules called sialic acids, as well as in the metabolism of thyroid hormones. The hormone differences have potential consequences for human brain development. The differences in sialic acid biology have multiple implications for the human condition, ranging from susceptibility or resistance to microbial pathogens, effects on endogenous receptors in the immune system, and potential effects on placental signaling, expression of oncofetal antigens in cancers, consequences of dietary intake of animal foods, and development of the mammalian brain.

Why is an article about the genetic differences between humans and great apes being published in the proceedings of a symposium on Bioastronomy? The Search for Extraterrestrial Intelligence (SETI) by humans is itself an indication of the unique cognitive status of our species. While several other species show evidence of intelligence, no others appear to have asked questions about their own origins, let alone ask if other forms of intelligence might exist beyond this planet. Understanding how human intelligence evolved could potentially help in making predictions about how extraterrestrial intelligence might have evolved. There are many approaches towards explaining the phenomenon of humans. In our own work in this area, we are mindful of Dobzhansky's statement that "nothing in biology makes sense, except in the light of evolution" (Dobzhansky 1973).

Darwin and Huxley first postulated the close evolutionary relationships between humans and the great apes (Huxley 1863; Darwin 1871), and this prediction has now been amply confirmed by a large body of experimental evidence (Sibley & Ahlquist 1987; Ruvolo 1997; Goodman 1999; Satta et al. 2000; Gagneux & Varki 2001). Comparisons of DNA sequence data indicate that humans

shared a common ancestor with the orangutan about 13 million years ago, the gorilla about 8 million years ago and the chimpanzees (common chimpanzees and bonobos) about 6-7 million years ago. The average genomic sequence difference between humans and the chimp bonobo clade is only slightly higher than 1% when examining gene-rich regions (Goodman 1999; Satta et al. 2000; Varki 2000). At the level of protein sequences, the similarities are even more striking, averaging between 99-100%. This has lead some scientists to call for a great ape/primate genome project (McConkey & Goodman 1997; McConkey et al. 2000; McConkey & Varki 2000) to accelerate understanding of the similarities and differences. In fact the National Human Genome Research Institute of the NIH has recently placed the chimpanzee genome on its "high-priority list"(Cyranoski 2002).

To date there has been relatively little work done on genetic differences between humans and great apes. Apart from a few chromosomal differences of uncertain biological significance (Yunis & Prakash 1982; McConkey 1997; Johnson et al. 2001), there are only a few examples of human specific genetic differences that do not represent merely intra-species polymorphisms in humans (Gagneux & Varki 2001). This paper will briefly outline our own recent discoveries in this area, with some emphasis on their potential to explicate human intelligence.

Work in our laboratory is focused on the sugar chains that are found attached to the surfaces of all cells and most secreted proteins in nature. Systematic studies of the structure and function of this thick "glycocalyx" (or "icing on the cake") have only recently emerged into the mainstream of molecular biology, as a sub-discipline called "Glycobiology" (Rademacher et al. 1988; Varki 1999). Current information indicates that the production of these complex sugar chains is dictated by a series of genes that encode the enzymes capable of generating each of the different types of sugars and then attaching them to growing sugar chains, which in turn are typically attached to proteins or lipids.

Of all the sugars found on cell surfaces and secreted proteins, our work is focused on one specific class called the sialic acids (Schauer 1982; Angata & Varki 2002). Sialic acids are typically found attached to the outermost ends of sugar chains, and are hence in a position to mediate or modulate a variety of receptor protein recognition events. These receptor proteins may be either endogenous (made by the same organism that made the sialic acids) or exogenous (made by other organisms, typically microbes, that use the sialic acids as attachment sites to initiate invasion) (Schauer 1982; Angata & Varki 2002).

In addition to acting as ligands for receptor proteins, sialic acids also provide a net negative charge to the surfaces of cells and proteins. We reasoned that if there were significant differences in the number and/or distribution of sialic acids attached to proteins in humans versus the great apes, this would be easily detected in 2-dimensional gel electrophoresis. In this procedure, complex mixtures of proteins from biological sources are first separated by virtue of their net negative charge (iso-electric focusing) and then by their molecular weight. The resulting 2-dimensional pattern is a reproducible "fingerprint" which can be compared with other closely related fingerprints in order to detect specific differences. Such differences can be confirmed by analyzing samples of interest using

mass-spectrometric sequencing (this overall approach is known as "proteomics") (Page et al. 1999).

We first chose to study blood plasma. It is easy to obtain and human blood plasma has already been extensively studied by proteomic analyses. When our collaborators at Oxford Glycosciences compared proteomic maps of blood plasma from humans and the great apes, they noted very few differences, indicating that the patterns of protein expression and sialic acid addition in the liver and the B lymphocytes (the major sources of plasma proteins) had not changed very much over the last 13 million years since we shared a common ancestor with the orangutan (Gagneux et al. 2001). In total, only 8 differences were detected amongst all the humans and great apes, of which only two were human-specific. Of the latter, one (haptoglobin) was actually a polymorphism within the human population and hence of limited interest in this context. The other was an apparent reduction in the amount of a protein called transthyretin. This reduction in transthyretin content of blood was confirmed by analyzing multiple human and great ape samples (Gagneux et al. 2001).

Transthyretin contributes to the transport and distribution of thyroid hormones to tissues via the bloodstream. Indeed, we found differences in the levels of thyroid hormones between humans and chimpanzees in the blood plasma (Gagneux et al. 2001). These differences were relatively minor, befitting the relatively minor role played by transthyretin in binding thyroid hormones in the blood. However, in contrast to the situation in blood, transthyretin is known to be the major carrier of thyroid hormones in the cerebrospinal fluid (the fluid that bathes the brain). Thus, it was of great interest to find that humans also have much lower levels of cerebrospinal fluid transthyretin than chimpanzees (Gagneux et al. 2001). Unfortunately, the amounts of cerebrospinal fluid available would not allow direct analyses of thyroid hormone levels. Given the well-known role that thyroid hormones play in the development and function of the brain, it is appealing to speculate that this difference in thyroid hormone metabolism could have some role in explaining human-ape differences. An added factor of interest is that plasma transthyretin levels can modulate the levels of a specific binding protein for retinol, another bioactive molecule that can affects brain development. Further studies are needed to see if these findings are of relevance in explaining "humanness".

In parallel work, we have also found several specific differences between humans and great apes with respect to the biology of sialic acids. The first and most prominent difference has been the subject of several recent publications (Chou et al. 1998; Muchmore et al. 1998; Brinkman-Van der Linden et al. 2000; Angata et al. 2001; Hayakawa et al. 2001; Varki 2002; Chou et al. 2002). Sialic acids are really a large family of molecules, most of which are derived from various modifications of a common sialic acid called N-acetylneuraminic acid (Neu5Ac). The most common modification is the addition of a single oxygen atom to Neu5Ac, to produce N-glycolylneuraminic acid (Neu5Gc) (Schauer 1982; Angata & Varki 2002). In most mammalian cells and tissues (including those of great apes), one finds both Neu5Ac and Neu5Gc, in varying ratios. In striking contrast, human cells and tissues have little or no Neu5Gc, and an excess of Neu5Ac (accumulation of the precursor form). Thus, the surfaces of human

cells are markedly different from those of other mammals, including the great apes (Muchmore et al. 1998).

Comparative studies of human and great ape genomes revealed that humans are deficient in Neu5Gc due to a specific mutation of the gene encoding the enzyme responsible for generating it (Chou et al. 1998; Irie et al. 1998). This mutation is universal in modern day human populations, and was evidently caused by an *Alu* element (a member of the *Alu* family of repeating units that comprise a significant percentage of the so-called "junk DNA" in the human genome) (Hayakawa et al. 2001). By studying the molecular lesion in the human genome, and by extracting sialic acids from fossils of extinct hominids, we were recently able to date the onset of this mutation to $\sim 2.5 - 3$ million years ago (Chou et al. 2002). This indicates that the loss of Neu5Gc could not have contributed to bipedalism (standing upright, which evidently occurred > 5 million years ago) nor to the various changes that occurred, as a result, in the structure and function of the hands, legs, back and pelvis.

On the other hand, the mutation occurred when human ancestors still had brains the same size as those of chimpanzees, i.e., before the onset of human brain expansion $\sim 2.1 - 2.0$ million years ago. Thus, we are exploring the possibility that the loss of Neu5Gc expression might have contributed to the process of brain expansion. An intriguing clue is the high levels of expression of Neu5Gc in most tissues of most mammals, with one exception – the brain (Schauer 1982). Neu5Ac is the dominant sialic acid in mammalian brains, with Neu5Gc being present only in very small amounts. This results from down-regulation of expression of the enzyme responsible for making Neu5Gc in mammalian brains (Kawano et al. 1995). This remarkable and selective regulation has been conserved for tens of millions of years of mammalian evolution, for unknown reasons. Of course, since the gene in question is mutated in humans, we have no detectable Neu5Gc in our brains. To predict whether this could have contributed to human brain expansion, we first need to know what, if any, are the detrimental consequences of having too much Neu5Gc in a mammalian brain. Studies are currently underway to address this issue.

This human-specific genetic mutation has several other implications for the human condition, such as susceptibility or resistance to microbial pathogens, effects on endogenous receptors in the immune system, potential effects on placental signaling, expression of oncofetal antigens in cancers, and consequences of dietary intake of animal foods. These issues have been discussed elsewhere (Varki 2002). Of note, we have also encountered several additional differences in sialic acid biology between humans and great apes. We recently described a genomic point mutation in humans responsible for an amino acid substitution that abolishes sialic acid recognition by a human molecule (Siglec-L1) that otherwise bears the signatures of a typical sialic acid binding lectin in the immune system (Angata et al. 2001). Interestingly, the great ape counterpart of this molecule does not have the mutation and binds preferentially to Neu5Gc, the sialic acid that is missing in humans. "Chimpanizing" the human Siglec-L1 molecule (i.e., correcting the mutation back to the ancestral state) restores sialic acid binding, again with a preference for Neu5Gc (Angata et al. 2001). These data and other unpublished sialic acid-related differences lead us to suggest that multiple changes in sialic acid biology that occurred during human evolution are

related to one another, perhaps in a cascade initiated by one of them (e.g., the mutation eliminating Neu5Gc that we first encountered). The implications of these findings for the present human condition, especially human intelligence, are unknown.

Lest false expectations be raised, it is best to conclude by quoting from our own recent review of this subject: " ... the discovery of a major biochemical and structural difference between humans and great apes and the elucidation of its underlying genetic basis tend to raise hopes that this molecular change played a major role in the evolution of uniquely human characteristics. Indeed, this is the issue likely to be of particular interest to the readership of this journal. However, the fact that this happens to be the first such difference recognized should not be taken as any indication of its relative significance in human evolution. It is indeed possible that this genetic change came under strong positive selection pressure because it contributed towards the evolution of organs like the human brain. On the other hand, it may have simply aided in the ability of humans to evade certain pathogens of animal origin. A less likely possibility is that this gene inactivation was a fortuitous event that simply drifted to fixation in the immediate ancestors of modern humans, because of a population bottleneck caused by other unrelated factors. Further studies will eventually elucidate which of these possibilities is correct." (Varki 2002)

References

Angata, T., Varki, N. M., & Varki, A. 2001, J. Biol. Chem., 276, 40282

Angata, T., & Varki, A. 2002, Chem. Rev., 102, 439

Brinkman-Van der Linden, E. C. M., Sjoberg, E. R., Juneja, L. R., Crocker, P. R., Varki, N., & Varki, A. 2000, J. Biol. Chem., 275, 8633

Chou, H. H., Takematsu, H., Diaz, S., Iber, J., Nickerson, E., Wright, K. L., Muchmore, E. A., Nelson, D. L., Warren, S. T., & Varki, A. 1998, Proc. Natl. Acad. Sci.USA, 95, 11751

Chou H.H., Hayakawa, T., Diaz, S., Krings, M., Indriati, E., Leakey, M., Paabo, S., Satta, Y., Takahata, N., & Varki, A. 2002, Proc. Natl. Acad. Sci. USA, 99, 11736

Cyranoski, D. 2002, Nature, 418, 910

Darwin, C. 1871, The descent of man, and selection in relation to sex (New York: D. Appleton and company)

Dobzhansky, T. 1973, Amer. Biol. Teacher, 35, 125

Gagneux, P., Amess, B., Diaz, S., Moore, S., Patel, T., Dillmann, W., Parekh, R., & Varki, A. 2001, Am. J. Phys. Anthropol., 115, 99

Gagneux, P., & Varki, A. 2001, Mol. Phylogenet. Evol., 18, 2

Goodman, M. 1999, Am. J. Hum. Genet., 64, 31

Hayakawa, T., Satta, Y., Gagneux, P., Varki, A., & Takahata, N. 2001, Proc. Natl. Acad. Sci. USA, 98, 11399

Huxley, T.H. 1863. Evidence as to Man's place in Nature (London: Williams and Norgate)

Irie, A., Koyama, S., Kozutsumi, Y., Kawasaki, T., & Suzuki, A. 1998, J. Biol. Chem., 273, 15866

Johnson, M. E., Viggiano, L., Bailey, J. A., Abdul-Rauf, M., Goodwin, G., Rocchi, M., & Eichler, E. E. 2001, Nature, 413, 514

Kawano, T., Koyama, S., Takematsu, H., Kozutsumi, Y., Kawasaki, H., Kawashima, S., Kawasaki, T., & Suzuki, A. 1995, J. Biol. Chem., 270, 16458

McConkey, E. H. 1997, Cytogenet. Cell Genet., 76, 189

McConkey, E. H., & Goodman, M. 1997, Trends. Genet., 13, 350

McConkey, E. H., Fouts, R., Goodman, M., Nelson, D., Penny, D., Ruvolo, M., Sikela, J., Stewart, C. B., Varki, A., & Wise, S. 2000, Mol. Phylogenet. Evol., 15, 1

McConkey, E. H., & Varki, A. 2000, Science, 289, 1295

Muchmore, E. A., Diaz, S., & Varki, A. 1998, Am. J. Phys. Anthropol., 107, 187

Page, M. J., Amess, B., Rohlff, C., Stubberfield, C., & Parekh, R. 1999, DDT, 4, 55

Rademacher, T. W., Parekh, R. B., & Dwek, R. A. 1988, Annu. Rev. Biochem., 57, 785

Ruvolo, M. 1997, Mol. Biol. Evol., 14, 248

Satta, Y., Klein, J., & Takahata, N. 2000, Mol. Phylogenet. Evol., 14, 259

Schauer, R. 1982, Sialic Acids: Chemistry, Metabolism and Function, Cell Biology Monographs, Volume 10 (New York: Springer-Verlag)

Sibley, C. G., & Ahlquist, J. E. 1987, J. Mol. Evol., 26, 99

Varki, A. 1999, in Essentials of Glycobiology, ed. A. Varki, J.D. Esko, R. Cummings, H. H. Freeze, G. W. Hart, & J. Marth (Plainview, N.Y.: Cold Spring Harbor Laboratory Press), 1

Varki, A. 2000, Genome. Res., 10, 1065

Varki, A. 2002, Yearbook. Phys. Anthropol., 44, 54

Yunis, J. J., & Prakash, O. 1982, Science, 215, 1525