

Available online at www.sciencedirect.com



Gene 373 (2006) 16-22



www.elsevier.com/locate/gene

### Ancient positive selection on CD155 as a possible cause for susceptibility to poliovirus infection in simians

Yoshiyuki Suzuki \*

Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, 1111 Yata, Mishima-shi, Shizuoka-ken 411-8540, Japan

Received 14 October 2005; received in revised form 16 December 2005; accepted 19 December 2005 Available online 24 February 2006 Received by S. Yokoyama

#### Abstract

Poliovirus is the etiological agent of poliomyelitis. From the observations that only simians are susceptible to poliovirus infection and that 37 amino acid sites (the poliovirus-binding associated [PBA] sites) in the domain D1 of CD155 are involved in the binding to poliovirus, it is considered that the susceptibility to poliovirus infection evolved through amino acid substitutions that occurred at the PBA sites on the ancestral branch of simians. Here it is shown that positive selection has operated on these substitutions by analyzing the nucleotide sequences encoding almost the entire region of D1 in humans, non-human hominoids (chimpanzees and gorillas), Old World monkeys (African green monkeys), New World monkeys (brown capuchins, squirrel monkeys, and marmosets), prosimians (ring-tailed lemurs), and non-primate mammals (rabbits). Positive selection is unlikely to have operated on the susceptibility to poliovirus infection, but possibly on the binding to another molecule. Elimination of susceptibility to poliovirus infection, but possibly on the binding to another molecule. Elimination of susceptibility to poliovirus infection. But possibly on the binding to another molecule. Elimination of susceptibility to poliovirus infection in simians may be difficult, because it also requires elimination of advantageous effects that have been exerted by CD155. © 2006 Elsevier B.V. All rights reserved.

Keywords: Poliovirus; Simian; CD155; Positive selection; Antagonistic pleiotropy

### 1. Introduction

Poliovirus, a member of the genus *Enterovirus* in the family *Picornaviridae*, is the etiological agent of poliomyelitis, which is an acute disease causing flaccid paralysis through selective destruction of motor neurons in the central nervous system (Landsteiner and Popper, 1909). Only humans are natural hosts for poliovirus. However, non-human hominoids and Old World monkeys are considered to be susceptible to poliovirus infection, because individuals or cell lines from these organisms may be infected by poliovirus under experimental conditions (Hsiung et al., 1964). New World monkeys are also known to be, at least

\* Tel.: +81 55 981 6847; fax: +81 55 981 6848.

E-mail address: yossuzuk@lab.nig.ac.jp.

partially, susceptible to poliovirus infection (Ida-Hosonuma et al., 2003). In contrast, prosimians and non-primate mammals do not appear to be susceptible.

Susceptibility to poliovirus infection is determined mainly by the interaction between the surface capsid proteins (VP1, VP2, and VP3) of poliovirus and the poliovirus receptor protein (CD155) on the host cell surface (Mendelsohn et al., 1989). CD155 (totally 417 amino acid sites) is a member of the immunoglobulin (Ig) superfamily, consisting of a signal peptide (amino acid positions 1–27), three extracellular Ig-like domains (D1 [positions 28–142], D2 [positions 143–242], and D3 [positions 243–330]), a transmembrane domain (positions 331– 355), and a cytoplasmic tail (positions 356–417). (Note that the numbers and positions of amino acid sites are those for human CD155.) D1 is known to bind to poliovirus (Koike et al., 1991; Selinka et al., 1991). From the analysis of the three-dimensional structure of CD155-poliovirus complex, 26 amino acid sites (positions 29, 30, 60, 61, 63, 75, 81–84, 86, 88–91, 93, 98, 102, 124, and 126–132) in D1 were found to be located at its interface (He et al., 2003). In addition, from mutagenesis analysis, 20 amino acid sites (positions 78, 80, 82-87, 92, 99-102, 117-119, 124, 126, 130, and 131) in D1 were identified as important for

*Abbreviations:* Ig, immunoglobulin; PBA, poliovirus-binding associated; INSD, International Nucleotide Sequence Database;  $r_{\rm S~(N)}$ , rate of synonymous (nonsynonymous) substitution;  $C_{\rm S~(N)}$ , number of synonymous (nonsynonymous) differences per sequence; *S* (*N*), number of synonymous (nonsynonymous) sites per sequence; *p*, probability; DNAM-1, DNAX accessory molecule-1; Tactile, T-cell activated increased late expression; Tage4, tumorassociated antigen E4; BRCA1, breast and ovarian cancer susceptibility gene 1; ANG, angiogenin; AP, antagonistic pleiotropy.

poliovirus binding (Bernhardt et al., 1994; Colston and Racaniello, 1994; Morrison et al., 1994; Harber et al., 1995; Liao and Racaniello, 1997). Therefore, 37 amino acid sites of D1 are involved in the binding to poliovirus, and are called the poliovirus-binding associated (PBA) sites in this paper.

CD155 has been identified not only in simians, which are susceptible to poliovirus infection, but also in prosimians and non-primate mammals, which are not susceptible (Ida-Hosonuma et al., 2003). Therefore, the difference in the susceptibility to poliovirus infection between these organisms appears to result from the difference in the amino acid sequence of CD155. To understand the evolutionary mechanisms of the susceptibility to poliovirus infection, it is interesting to examine natural selection that has operated on CD155. In general, natural selection can be detected by comparing the rates of synonymous  $(r_s)$  and nonsynonymous  $(r_N)$  substitutions for the protein-coding nucleotide sequences (Hughes and Nei, 1988). Positive and negative selection are inferred when  $r_{\rm S} < r_{\rm N}$  and  $r_{\rm S} > r_{\rm N}$ , respectively. Ida-Hosonuma et al. (2003) analyzed the nucleotide sequences encoding almost the entire region (positions 36-132) of D1 obtained from humans, non-human hominoids (chimpanzees and gorillas), Old World monkeys (African green monkeys), New World monkeys (brown capuchins, squirrel monkeys, and marmosets), prosimians (ring-tailed lemurs), and non-primate mammals (rabbits). They observed that  $r_{\rm S} > r_{\rm N}$  for all pairwise comparisons of these sequences, and concluded that positive selection had not operated on D1. However, the binding of CD155 to poliovirus is not mediated by the entire region of D1, but only by the PBA sites. In addition, according to the parsimony principle (Fitch, 1971), the ancestor of primates was not susceptible to poliovirus infection, but the susceptibility evolved on the ancestral branch of simians in the phylogenetic tree. Therefore, natural selection should be examined for the PBA sites on the ancestral branch of simians.

Here it is shown that ancient positive selection on CD155 is a possible cause for susceptibility to poliovirus infection in simians. The biological significance of the results is discussed.

### 2. Materials and methods

### 2.1. Sequence data

The nucleotide sequences encoding almost the entire region of D1, that were analyzed by Ida-Hosonuma et al. (2003), were used to examine natural selection that has operated at the PBA sites on the ancestral branch of simians. The accession numbers in the International Nucleotide Sequence Database (INSD) for the sequences of humans, chimpanzees, gorillas, African green monkeys (copies 1 and 2), brown capuchins, squirrel monkeys, marmosets, ring-tailed lemurs, and rabbits are M24407, AB086 255, AB086253, D12611, D12613, AB086124–AB086131, AB 086252, AB086254, AB086132–AB086137, and AB086138– AB086144, respectively. These sequences have been reported to be orthologous (Ida-Hosonuma et al., 2003). Note that African green monkeys possess two sequences because of the occurrence of gene duplication after speciation (Koike et al., 1992).

A multiple alignment of the nucleotide sequences was made using the computer program CLUSTAL W (version 1.81) (Thompson et al., 1994). The alignment contained a gap which was apparently accompanied by a frame shift. To eliminate the gap, a multiple alignment was also made for the amino acid sequences that were encoded by the nucleotide sequences (Fig. 1).

Position Humans Chimpanzees Gorillas African green monkeys African green monkeys Brown capuchins Squirrel monkeys Marmosets Ring-tailed lemurs Rabbits	1 2	333344444444455555 67890123456789012345 QVPGPLGDSVTLPCYLQVPN QVPGPLGDSVTLPCYLQVPN QVPGFLGDSVTLPCYLQVPG QVPGFLGDSVTLPCYLQVPG QVHGFLGDSVTLPCQLQLPS QVHGFLGDSVTLPCHLQLPS QVHGFLGDSVTLPCHLQLPS QVRGLLGGNVTLPCHLQLDPS WIQGFLGDTVTLPCYLPLEAD-	55556666 67890123 MEVTHVSC MEVTHVSC MEETHVSC MEVTHVSC TEVTHVSC TEVTHVSC TEVTHVSC VQVTQVTW MRVTQVTW ** *	666666 456789 LTWARE LTWARE LTWSRE LTWSRE LTWTRI LTWTRI LTWTRI MRRQPA MRQPA	7777777 0123456 IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGESGSMI IGGSSVI IGGRSVI IGGRSVI IGGRSVI IGGRSVI I	7777 5789 AVFH AVFH AVFH AVFH AVFH AVFH AVFH AVFH	88: 01: 02T 02T 02T 02T 02T 02T 01 01 01 01 01 01 01 01 01 01 01 01 01	8888888 2345676 QGPSYSI QGPSYSI QGPSYSI QGPNYSI QGPNYPI QGSNYPI QGSNYPI RGASVSI *** * *	899 901 SKR SKR PKR SER SER SER SER SER	999 223 21E 21E 21E 21E 21E 21E 21E 21E
Position Humans Chimpanzees Gorillas African green monkeys African green monkeys Brown capuchins Squirrel monkeys Marmosets Ring-tailed lemurs Rabbits	1 2	11111111111111 999999000000000011111 4567890123456789012345 FVAARLGAELRNASLRMFGLRV FVAARLGAELRNASLRMFGLRV FVAARLGTELRDASLRMFGLRV FVAARLGAELRDASLRMFGLRV FVAARLGAELRDASLRVFELRA FVAARLGAELRDASLRVFELRA FVAARLGAELRDASLRVFELRA FAARSRSELRDASLVVSQLHI FVAARPGAELWNASLAVSGLRA * * ####	11111111 11112222 67890123 EDEGNYTC EDEGNYTC EDEGNYTC EDEGNYTC EDEGNYTC EDEGNYTC EDEANYTC EDEANYTC ###	111111 222222 456789 2LFVTFF LFVTFF 2LFVTFF 2LFVMFF 2LFVTFFF 2LFVTFFF 2LFVTFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1111 2333 2012 2065 2065 2065 2065 2065 2065 2065 206					

Fig. 1. Multiple alignment of the amino acid sequences for almost the entire region (amino acid positions 36–132) of D1 obtained from humans, chimpanzees, gorillas, African green monkeys, brown capuchins, squirrel monkeys, marmosets, ring-tailed lemurs, and rabbits. African green monkeys 1 and 2 indicate copies 1 and 2 of CD155 in this species, respectively. The amino acid sites that were located at the interface of CD155–poliovirus complex in the analysis of the three-dimensional structure and those that were important for poliovirus binding in the mutagenesis analysis are marked with \* and <sup>#</sup>, respectively. The sites that were eliminated from the analysis of natural selection are dark-shaded. The PBA sites where amino acid substitutions occurred on the ancestral branch of simians are light-shaded.

The nucleotide and amino acid alignments were compared with each other, and only the codon sites in the former alignment that were consistently aligned with the latter alignment without gaps between the nucleotides within the codons were extracted. As a result, codon positions 54 and 55 were not extracted, and a total of 95 codon sites, consisting of 35 PBA sites and 60 non-PBA sites, was used for the analysis of natural selection.

### 2.2. Data analysis

The ancestral branch of simians in the phylogenetic tree could be identified only when the sequences analyzed were orthologous, as reported by Ida-Hosonuma et al. (2003). To confirm the orthologous relationships among these sequences, the phylogenetic tree was constructed by the neighbor-joining method (Saitou and Nei, 1987) using the *p*-distance with MEGA (version 3.1) (Kumar et al., 2004). Note that the topology of the phylogenetic tree constructed using the *p*-distance is generally known to be more reliable than that using other distances (Nei and Kumar, 2000). The reliabilities of the interior branches were assessed by the bootstrap method with 1000 resamplings (Felsenstein, 1985).

Natural selection that has operated at the PBA sites on the ancestral branch of simians was examined by the method of Zhang et al. (1998). For the PBA sites, the ancestral codon sequence was inferred for each interior node of the phylogenetic tree by the likelihood-based Bayesian method (Yang et al., 1995) using PAML (version 3.14) (Yang, 1997). The numbers of synonymous  $(C_{\rm S})$  and nonsynonymous  $(C_{\rm N})$ differences per sequence were counted for the ancestral branch of simians by comparing its flanking codon sequences. In addition, the average numbers of synonymous (S) and nonsynonymous (N) sites per sequence for the entire phylogenetic tree were estimated by the maximum likelihood method. The probability (p) of obtaining the observed or more biased values for  $C_{\rm S}$  and  $C_{\rm N}$ , under the assumption of selective neutrality, was computed by Fisher's exact test using a  $2 \times 2$ contingency table consisting of  $C_{\rm S}$  (in the first column of the first row),  $C_{\rm N}$  (in the first column of the second row),  $(S-C_{\rm S})$ 

(in the second column of the first row), and  $(N-C_N)$  (in the second column of the second row). Positive and negative selection were inferred when  $C_N/C_S > (N-C_N)/(S-C_S)$  and  $C_N/C_S < (N-C_N)/(S-C_S)$ , respectively, with p < 0.05.

### 3. Results

### 3.1. Phylogenetic relationships among CD155 sequences

The phylogenetic tree constructed for the nucleotide sequences encoding almost the entire region of D1 in humans, chimpanzees, gorillas, African green monkeys, brown capuchins, squirrel monkeys, marmosets, ring-tailed lemurs, and rabbits, is shown in Fig. 2. The topology of the phylogenetic tree was the same as that of the species tree (Fig. 3), except that chimpanzees clustered with gorillas instead of humans as well as brown capuchins clustered with marmosets instead of squirrel monkeys. However, bootstrap probabilities for the interior branches supporting the human–gorilla and brown capuchin–marmoset clusters were only 77% and 49%, respectively. In contrast, other branches were supported by high bootstrap probabilities (>90%).

Although the topology of the phylogenetic tree constructed using the *p*-distance is generally known to be more reliable than that using other distances as mentioned above, the *p*-distance may be biased if the pattern of nucleotide substitution is complicated. To take the pattern into account, the model of nucleotide substitution that was fittest to the data was inferred using MODELTEST (version 3.7) (Posada and Crandall, 1998). As a result, the model of Hasegawa et al. (1985) (HKY) with a gamma distribution (shape parameter [ $\alpha$ ] of 0.883) for the among-site rate variation (HKY+ $\Gamma$ ) was inferred as the fittest. However, the topology of the phylogenetic tree constructed using this model with PAUP\* (version 4.0b10) (Swofford, 1998) was the same as that using the *p*-distance (Fig. 2), suggesting that the *p*-distance was not biased.

The phylogenetic trees were also constructed for the PBA sites and non-PBA sites separately using the *p*-distance. The topology of the former tree was the same as that for the entire



Fig. 2. The phylogenetic tree constructed for the nucleotide sequences encoding almost the entire region of D1 (285 nucleotide sites) obtained from humans, chimpanzees, gorillas, African green monkeys, brown capuchins, squirrel monkeys, marmosets, ring-tailed lemurs, and rabbits. African green monkeys 1 and 2 indicate copies 1 and 2 of CD155 in this species, respectively. The bootstrap probabilities are indicated for interior branches. Note that the topology of the phylogenetic tree is the same as that of the species tree (Fig. 3), except that chimpanzees clustered with gorillas instead of humans as well as brown capuchins clustered with marmosets instead of squirrel monkeys. The scale bar indicates 0.05 substitutions per site.



Fig. 3. The species tree for humans, non-human hominoids (chimpanzees and gorillas), Old World monkeys (African green monkeys), New World monkeys (brown capuchins, squirrel monkeys, and marmosets), prosimians (ring-tailed lemurs), and non-primate mammals (rabbits). African green monkeys 1 and 2 indicate copies 1 and 2 of CD155 in this species, respectively. The *N/S* values at the PBA sites and non-PBA sites are indicated. The  $C_N/C_S$  values at the PBA sites and non-PBA sites are indicated above and below each branch, respectively. The total  $C_N/C_S$  values for the simian (broken line) and non-simian (dotted line) groups of branches are also indicated. \* indicates that  $C_N/C_S > (N-C_N)/(S-C_S)$  with p < 0.05 (positive selection), whereas <sup>#</sup>, <sup>##</sup>, and <sup>#####</sup> indicate that  $C_N/C_S < (N-C_N)/(S-C_S)$  with p < 0.05, p < 0.0005 (negative selection), respectively.

sequence (Fig. 2), but the topology of the latter tree was exactly the same as that of the species tree (Fig. 3). Kimura's (1980) (Kimura) and HKY+ $\Gamma$  ( $\alpha$ =0.909) models were inferred as the fittest to the PBA and non-PBA sites, respectively. The topology of the phylogenetic tree for the PBA sites constructed using Kimura model was the same as that using the *p*-distance (Fig. 2), whereas chimpanzees clustered with gorillas instead of humans in the phylogenetic tree for the non-PBA sites constructed using HKY+ $\Gamma$  model.

Overall, the interior branches supported by low bootstrap probabilities were either consistent or inconsistent with the species tree according to the regions of nucleotide sequences analyzed, whereas those supported by high bootstrap probabilities were generally consistent with the species tree. Although the possibility that multiple events of gene duplications and gene losses occurred for CD155 cannot be rejected, it appears to be reasonable to consider that the sequences analyzed were indeed orthologous as reported by Ida-Hosonuma et al. (2003), and the incorrect clusters were obtained due to statistical flucuations of the nucleotide substitution. Therefore, the topology of the species tree (Fig. 3) was used for the analysis of natural selection.

## 3.2. Natural selection at the PBA sites on the ancestral branch of simians

To examine natural selection that has operated at the PBA sites on the ancestral branch of simians, the ancestral codon sequence was inferred for each interior node of the phy-

logenetic tree. The average posterior probability of inferred codons for each ancestral sequence ranged from 87.6% to 100.0%, indicating that the inferred sequences were more or less reliable. The  $C_{\rm S}$  and  $C_{\rm N}$  values as well as the S and N values are summarized in Fig. 3. The S and N values were 33 and 72, respectively. For the ancestral branch of simians,  $C_{\rm S}$ and  $C_{\rm N}$  were 0 and 11, respectively. Since  $C_{\rm N}/C_{\rm S}$  was significantly greater than  $(N-C_N)/(S-C_S)$  (p=0.0159), positive selection was inferred. When natural selection was examined for other branches of the phylogenetic tree, neither positive nor negative selection was inferred for any branch, except for the chimpanzee and rabbit branches where negative selection was inferred (Fig. 3). Note, however, that for a given branch the failure to detect natural selection does not necessarily indicate that natural selection did not operate, because the occurrence of nucleotide substitution is influenced by statistical fluctuations. To obtain rough ideas about natural selection that has operated before and after the existence of the ancestral branch of simians, all other branches of the phylogenetic tree were classified into simian and non-simian groups, according to whether they belonged to the simian lineage or not in the phylogenetic tree (Fig. 3). For each group, the  $C_{\rm S}$  and  $C_{\rm N}$  values were summed for all branches and natural selection was examined in a way similar to the case for the analysis of individual branches. As a result, negative selection was inferred for both the simian and non-simian groups (Fig. 3). In particular, negative selection was inferred for the simian group even when the chimpanzee branch was eliminated.

# 3.3. Relative strength of natural selection at the PBA and non-PBA sites

Natural selection that has operated on the ancestral branch of simians as well as on other branches of the phylogenetic tree for the non-PBA sites was also examined in a way similar to the case for the PBA sites (Fig. 3). The average posterior probability of inferred codons for each ancestral sequence ranged from 87.1% to 100.0%. However, neither positive nor negative selection was inferred for any branch. No selection was inferred even when the branches were classified into simian and non-simian groups. These results suggest that natural selection operating at the non-PBA sites, if any, was weaker than that at the PBA sites. Interestingly, a similar observation was also obtained from the analysis of the duplicate copies (copies 1 and 2) of CD155 in African green monkeys. CD155 appears to have duplicated in African green monkeys, as mentioned above. In general, after gene duplication, the rate of amino acid substitution is often accelerated for one of two copies because of relaxation of functional constraint, and this copy eventually becomes a pseudogene (Wagner, 2002). In the case of CD155 in African green monkeys, the number (9) of amino acid substitutions that have accumulated in copy 2 is significantly greater than that (1) in copy 1 (p=0.0117), suggesting that the rate was accelerated for copy 2 (Fig. 3). Among the nine amino acid substitutions in copy 2, eight occurred at the PBA sites whereas only one at the non-PBA sites, suggesting that the extent of relaxation of functional constraint at the PBA sites was significantly greater than that at the non-PBA sites (p=0.00474). These results indicate that the functional constraint at the PBA sites was stronger than that at the non-PBA sites before duplication.

### 4. Discussion

# 4.1. Why were the PBA sites positively selected on the ancestral branch of simians?

Positive selection appears to have operated at the PBA sites on the ancestral branch of simians, where the susceptibility to poliovirus infection is considered to have evolved. However, it is unlikely that positive selection has operated on the susceptibility to poliovirus infection, because of the following two reasons. First, although poliovirus infection may not decrease the fitness of individuals to a large extent due to a relatively low rate (0.01–0.001) of progression to poliomyelitis, there is no evidence that it increases the fitness. Second, the emergence of poliovirus appears to be much later than the existence of the ancestral branch of simians, which is known to be 20–30 million years ago. Actually, the evolutionary distance measured in terms of the number of nucleotide substitutions between poliovirus and other viruses has been estimated to be smaller than 1 per site (Brown et al., 2003). In addition, the rate of nucleotide substitution for poliovirus has been estimated to be greater than  $1 \times 10^{-3}$  per site per year (Gavrilin et al., 2000). Therefore, the emergence of poliovirus should be only less than 1000 years ago.

The exact reason for the occurrence of positive selection is unclear partly because the physiological functions of PBA sites are unknown. However, since these sites are all involved in the binding of CD155 to other molecules (VP1, VP2, and VP3 of poliovirus), it is possible that the amino acid substitutions at these sites on the ancestral branch of simians also influenced the binding to another molecule. It has been reported that ectodomains of human CD155 bind to vitronectin (Lange et al., 2001), nectin-3 (Fabre et al., 2002), DNAM-1 (DNAX accessory molecule-1, or CD226), CD96 (Tactile [T-cell activated increased late expression]) (Bottino et al., 2003), and itself (dimerization) (Mueller and Wimmer, 2003). In particular, the binding of CD155 and nectin-2 (CD112) expressed on the tumor cells of colorectal carcinoma and malignant glioma to DNAM-1 and CD96 expressed on natural killer (NK) cells induces NK cell-mediated lysis of tumor cells (Bottino et al., 2003). This function appeared to be advantageous when it evolved. In mice, CD155 (Tage4 [tumorassociated antigen E4]) is known to bind to vitronectin (Lange et al., 2001) and nectin-3 (Ikeda et al., 2003; Mueller and Wimmer, 2003). It is therefore interesting to examine whether CD155 from non-simians binds to DNAM-1 and CD96, and whether the PBA sites are involved in the bindings. Note, however, that among 35 PBA sites, amino acid substitutions occurred only at 10 sites on the ancestral branch of simians (Fig. 1). Therefore, these sites may be more important than other sites for understanding the reason for the occurrence of positive selection. It is also interesting to note that the PBA sites on the simian group of branches are negatively selected. It was reported that 29 amino acid sites of the surface capsid proteins of poliovirus (positions 101, 107-109, 166-168, 214, 221-224, 226-228, 234, 280, 295-298, and 301 of VP1, positions 137-139 and 141 of VP2, and positions 58, 59, and 62 of VP3) were located at the CD155-poliovirus interface (He et al., 2003). (Note that the amino acid positions are those for serotype 1 poliovirus.) When the test of selective neutrality was conducted for each of 25 sites (excluding positions 108 and 221-223 of VP1 because of the existence of gaps and the absence of synonymous sites),  $r_{\rm S} > r_{\rm N}$  for 24 sites and negative selection was inferred for 18 sites, whereas  $r_{\rm S} < r_{\rm N}$  only for 1 site and positive selection was not inferred for any site (Suzuki, 2004). These observations indicate that the binding between CD155 and poliovirus is evolutionarily stable.

In contrast, natural selection, either positive or negative, operating at the non-PBA sites generally appears to be weaker than that operating at the PBA sites. Although the biological functions of non-PBA sites are unknown, it is implied that the non-PBA sites are involved in the physiological functions of CD155, such as the binding to other molecules, more indirectly than the PBA sites.

### 4.2. Association between positive selection and diseases

In this paper, it was shown that ancient positive selection on CD155 was a possible cause for susceptibility to poliovirus infection in simians. Interestingly, positive selection has often been found to operate on disease-associated genes. For

example, mutant alleles of  $\beta$ -globin genes that cause sickle-cell anemia are positively selected in regions endemic for malaria, because sickle-cells are resistant to malaria infection. Positive selection was also detected in the human-chimpanzee comparison of the breast and ovarian cancer susceptibility gene 1 (BRCA1) (Huttley et al., 2000). Angiogenin (ANG), a tumorgrowth promoter, appears to be positively selected in primates (Zhang and Rosenberg, 2002). Clark et al. (2003) examined 7645 sets of orthologous genes from humans, chimpanzees, and mice, and found that Mendelian disease genes were positively selected more often than other genes on the human branch. Nielsen et al. (2005) analyzed 8079 sets of orthologous genes from humans and chimpanzees, and found that cancer-related genes were positively selected more often than other genes. (Note, however, that no relationship was observed between positive selection and disease-associated genes in the analysis of 13,545 sets of orthologous genes from humans and chimpanzees [Chimpanzee Sequencing and Analysis Consortium, 2005], and that Mendelian disease genes were found to be negatively selected more often than other genes in the analysis of polymorphisms for 3277 human genes [Bustamante et al., 2005].) These observations have been interpreted as suggesting that diseases are caused by deleterious effects of genes that also exert advantageous effects. This property of disease-associated genes is called antagonistic pleiotropy (AP) (Medawar, 1946). Since advantageous and deleterious effects are exerted at different time points in the ontogeny of an individual, AP of these genes may be called AP in ontogeny.

The amino acid substitutions that occurred at the PBA sites on the ancestral branch of simians also appear to have exerted effects of AP. However, advantageous and deleterious effects did not both come about during the same time frame, with the latter effect occurring much later (<1000 years ago) than the former effect (20-30 million years ago). Interestingly, a similar pattern for the evolution of the effects of AP was observed for amino acid substitutions that occurred in the mitochondrial proteins of humans during the migration out of Africa to temperate and arctic areas (Ruiz-Pesini et al., 2004). These substitutions were adaptive to colder environments during migration, but later increased the predisposition to modern diseases. Since advantageous and deleterious effects are exerted at different time points in the phylogeny of organisms, AP of these genes may be called AP in phylogeny. Although the mechanisms for the evolution of diseases are different between APs in ontogeny and phylogeny, elimination of diseases in humans may be difficult in both cases, because it also requires elimination of advantageous effects that have been exerted by disease-associated genes.

#### Acknowledgements

The author thanks Naruya Saitou for suggestions on the topology of the species tree used in this study. I am also grateful to two anonymous reviewers for valuable comments on an earlier version of the manuscript. This work was supported by KAKENHI 17770007.

### References

- Bernhardt, G., Harber, J., Zibert, A., deCrombrugghe, M., Wimmer, E., 1994. The poliovirus receptor: identification of domains and amino acid residues critical for virus binding. Virology 203, 344–356.
- Bottino, C., et al., 2003. Identification of PVR (CD155) and Nectin-2 (CD112) as cell surface ligands for the human DNAM-1 (CD226) activating molecule. J. Exp. Med. 198, 557–567.
- Brown, B., Oberste, M.S., Maher, K., Pallansch, M.A., 2003. Complete genomic sequencing shows that polioviruses and members of human enterovirus species C are closely related in the noncapsid coding region. J. Virol. 77, 8973–8984.
- Bustamante, C.D., et al., 2005. Natural selection on protein-coding genes in the human genome. Nature 437, 1153–1157.
- Chimpanzee Sequencing and Analysis Consortium, 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437, 69–87.
- Clark, A.G., et al., 2003. Inferring nonneutral evolution from human-chimpmouse orthologous gene trios. Science 302, 1960–1963.
- Colston, E., Racaniello, V.R., 1994. Soluble receptor-resistant poliovirus mutants identify surface and internal capsid residues that control interaction with the cell receptor. EMBO J. 13, 5855–5862.
- Fabre, S., et al., 2002. Prominent role of the Ig-like V domain in transinteractions of nectins. J. Biol. Chem. 277, 27006–27013.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.
- Fitch, W.M., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. 20, 406–416.
- Gavrilin, G.V., Cherkasova, E.A., Lipskaya, G.Y., Kew, O.M., Agol, V., 2000. Evolution of circulating wild poliovirus and of vaccine-derived poliovirus in an immunodeficient patient: a unifying model. J. Virol. 74, 7381–7390.
- Harber, J., Bernhardt, G., Lu, H.-H., Sgro, J.-Y., Wimmer, E., 1995. Canyon rim residues, including antigenic determinants, modulate serotype-specific binding of polioviruses to mutants of the poliovirus receptor. Virology 214, 559–570.
- Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.
- He, Y., et al., 2003. Complexes of poliovirus serotypes with their common cellular receptor, CD155. J. Virol. 77, 4827–4835.
- Hsiung, G.-D., Black, F.L., Henderson, J.R., 1964. Susceptibility of primates to viruses in relation to taxonomic classification. In: Buettener-Janusch, J. (Ed.), Evolutionary and Genetic Biology of Primates, vol. 2. Academic Press, New York.
- Hughes, A.L., Nei, M., 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335, 167–170.
- Huttley, G.A., et al., 2000. Adaptive evolution of the tumour suppressor BRCA1 in humans and chimpanzees. Nat. Genet. 25, 410–413.
- Ida-Hosonuma, M., et al., 2003. Host range of poliovirus is restricted to simians because of a rapid sequence change of the poliovirus receptor gene during evolution. Arch. Virol. 148, 29–44.
- Ikeda, W., et al., 2003. Tage4/nectin-like molecule-5 heterophilically transinteracts with cell adhesion molecule nectin-3 and enhances cell migration. J. Biol. Chem. 278, 28167–28172.
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.
- Koike, S., Ise, I., Nomoto, A., 1991. Functional domains of the poliovirus receptor. Proc. Natl. Acad. Sci. U. S. A. 88, 4104–4108.
- Koike, S., Ise, I., Sato, Y., Yonekawa, H., Gotoh, O., Nomoto, A., 1992. A second gene for the African green monkey poliovirus receptor that has no putative N-glycosylation site in the functional N-terminal immunoglobulinlike domain. J. Virol. 66, 7059–7066.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.
- Landsteiner, K., Popper, E., 1909. Ubertragung der Poliomyelitis acuta auf Affen. Z. Immunitatstorsch. Orig. 2, 377–390.

- Lange, R., Peng, X., Wimmer, E., Lipp, M., Bernhardt, G., 2001. The poliovirus receptor CD155 mediates cell-to-matrix contacts by specifically binding to vitronectin. Virology 285, 218–227.
- Liao, S., Racaniello, V., 1997. Allele-specific adaptation of poliovirus VP1 B-C loop variants to mutant cell receptors. J. Virol. 71, 9770–9777.

Medawar, P.B., 1946. Old age and natural death. Mod 1, 30-56.

- Mendelsohn, C.L., Wimmer, E., Racaniello, V.R., 1989. Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell 56, 855–865.
- Morrison, M.E., He, Y.J., Wien, M.W., Hogle, J.M., Racaniello, V.R., 1994. Homolog-scanning mutagenesis reveals poliovirus receptor residues important for virus binding and replication. J. Virol. 68, 2578–2588.
- Mueller, S., Wimmer, E., 2003. Recruitment of nectin-3 to cell-cell junctions through trans-heterophilic interaction with CD155, a vitronectin and poliovirus receptor that localizes to alpha(v)beta3 integrin-containing membrane microdomains. J. Biol. Chem. 278, 31251–31260.
- Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, Oxford, New York.
- Nielsen, R., et al., 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol. 3, e170.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
- Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., Wallace, D.C., 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303, 223–226.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

- Selinka, H.C., Zibert, A., Wimmer, E., 1991. Poliovirus can enter and infect mammalian cells by way of an intercellular adhesion molecule 1 pathway. Proc. Natl. Acad. Sci. U. S. A. 88, 3598–3602.
- Suzuki, Y., 2004. Negative selection on neutralization epitopes of poliovirus surface proteins: implications for prediction of candidate epitopes for immunization. Gene 328, 127–133.
- Swofford, D.L., 1998. PAUP\*. Phylogenetic analysis using parsimony (\*and other methods), Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple-sequence alignment through sequence weighting, position-specific gap penalties, and weight-matrix choice. Nucleic Acids Res. 22, 4673–4680.
- Wagner, A., 2002. Asymmetric functional divergence of duplicate genes in yeast. Mol. Biol. Evol. 19, 1760–1768.
- Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.
- Yang, Z., Kumar, S., Nei, M., 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics 141, 1641–1650.
- Zhang, J., Rosenberg, H.F., 2002. Diversifying selection of the tumor-growth promoter angiogenin in primate evolution. Mol. Biol. Evol. 19, 438–445.
- Zhang, J., Rosenberg, H.F., Nei, M., 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc. Natl. Acad. Sci. U. S. A. 95, 7308–7313.