

# SCIENTIFIC REPORTS



OPEN

## Horizontal gene transfer events reshape the global landscape of arm race between viruses and homo sapiens

Received: 14 March 2016

Accepted: 09 May 2016

Published: 07 June 2016

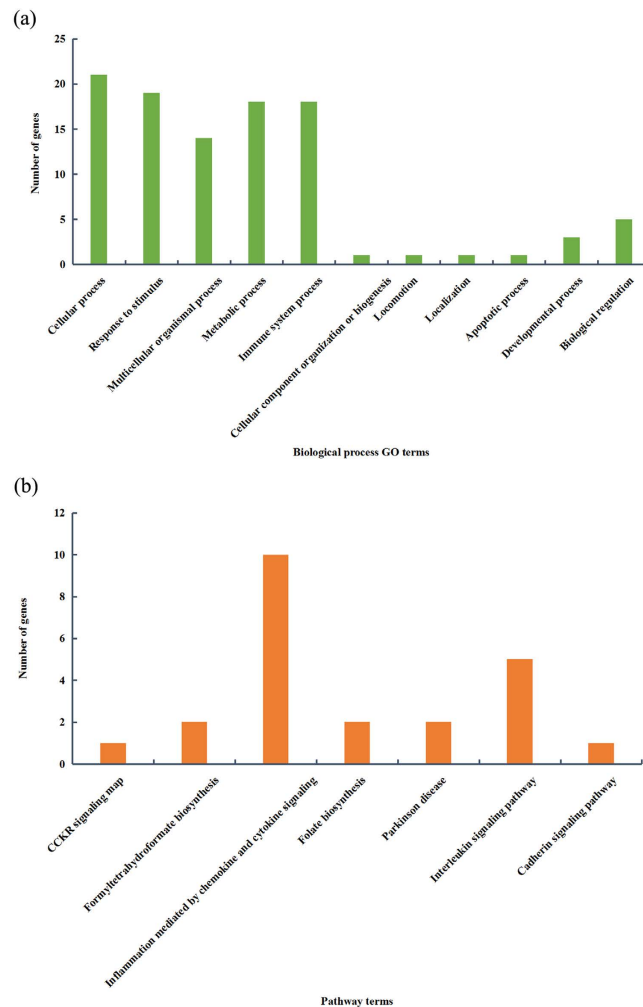
Dong-Sheng Chen<sup>1</sup>, Yi-Quan Wu<sup>2</sup>, Wei Zhang<sup>3</sup>, San-Jie Jiang<sup>4</sup> & Shan-Ze Chen<sup>5</sup>

Horizontal gene transfer (HGT) drives the evolution of recipient organism particularly if it provides a novel function which enhances the fitness or its adaption to the environment. Virus-host co-evolution is attractive for studying co-evolutionary processes, since viruses strictly replicate inside of the host cells and thus their evolution is inexorably tangled with host biology. HGT, as a mechanism of co-evolution between human and viruses, has been widely documented, however, the roles HGT play during the interaction between human and viruses are still in their infancy. In this study, we performed a comprehensive analysis on the genes horizontally transferred between viruses and their corresponding human hosts. Our study suggests that the HGT genes in human are predominantly enriched in immune related GO terms while viral HGT genes are tend to be encoded by viruses which promote the invasion of immune system of hosts. Based on our results, it gives us a hint about the evolution trajectory of HGT events. Overall, our study suggests that the HGT between human and viruses are highly relevant to immune interaction and probably reshaped the arm race between hosts and viruses.

Horizontal gene transfer has been well recognized as an important driving force of evolution, as it provides gene flow between evolutionary remote lineages which may facilitate the evolution of recipient species by providing novel functions<sup>1–9</sup>. Since the first report of HGT from bacteriophage to *corynebacterium diphtheriae* in 1951<sup>10</sup>, many other species varying from prokaryote to eukaryote have also been found to be the recipients of exogenous genes<sup>2,5,6,10</sup>. Most transferred genes are deleterious and will be wiped out by nature selection. In contrast, a small proportion of HGT events provide advantages against selection to hosts and thus can be retained in the species<sup>11,12</sup>. For instance, plant-parasitic nematodes gained cellulase genes and thus acquired cellulase activity resulting from HGT from several microbial donors, which significantly enhanced their parasitism and pathogenicity<sup>13</sup>. More interestingly, HGT was found to contribute to the plant parasitic mechanism of oomycetes, as some HGT genes are predicted to metabolize sugars, rutin, hemicellulose for feeding on plant tissues and some HGT genes are predicted to degrade structures components specific to plants which favors its attacking on plant cells<sup>14</sup>. HGT is common in animals as well, for example, a developmentally expressed functional gene in zebrafish was found to be laterally transferred from marine bacteria<sup>15</sup>.

Over the past hundreds of millions of years, viruses and hosts have been undergoing a continuous co-evolutionary process, involving both host defense and viral evasion mechanisms<sup>16,17</sup>. Viruses have evolved various strategies to evade host immune system. For instance, herpesviruses, a group of large DNA viruses, have developed to encode proteins that mimic cellular cytokines and cytokine receptors to modulate cytokine mediated signals during infection<sup>18–20</sup>. In contrast, host immune system is due to protect itself from invading pathogens, including viruses. HGT has been frequently reported between hosts and viruses and thus considered as an important contributor for host-virus co-evolution<sup>21–28</sup>.

<sup>1</sup>Department of Genetics, University of Cambridge, Cambridge, CB2 3EH, UK. <sup>2</sup>Max von Pettenkofer-Institute for Virology, Ludwig-Maximilians-University Munich, 80336 Munich, Germany. <sup>3</sup>Research unit gene vector, Helmholtz Zentrum, 81377 Munich, Germany. <sup>4</sup>Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK. <sup>5</sup>Department of Pathophysiology, West China School of Preclinical Sciences and Forensic Medicine, Sichuan University, 610041 Chengdu, Sichuan Province, China. Correspondence and requests for materials should be addressed to S.Z.C. (email: chenshanze2007@sina.com)



**Figure 1. The biological process and pathway classification of the 51 HGT genes encoded by human.** (a) Biological process classification (b) Pathway classification.

The systematic screen on 26 animal species suggest that there are hundreds of HGT genes in primates, flies and nematodes<sup>29</sup>. Alastair Crisp *et al.* identified hundreds of HGT genes in the genome of human beings which originally transferred from plants, fungus and bacterial<sup>29</sup>. It is of particular interest that they also predicted the HGT which happened between 51 human encoding genes and 27 viral encoding genes. In this study, we investigated the characteristics, evolution history and significance of genes transferred between human and viruses.

### Characterization of the 51 HGT Genes Encoded by Human

In order to study the functions of the 51 HGT genes encoded by human (Supplementary Table S1), the data set was submitted to PANTHER<sup>30,31</sup> for gene list classification analysis. Among the 51 HGT genes, some genes could not be mapped to any GO or pathways terms for lack of annotation while others might be related to multiple terms. Therefore the total number of genes in each classification analysis is not equal to 51 because of overlapping of gene list in different catalogues. Briefly, 27, 29, 5 and 17 genes were successfully mapped in 'molecular function', 'biological process', 'cellular process' and 'pathway' terms respectively (Supplementary Table S2). According to the 'molecular function classification' result, 27 out of 51 genes were found to have 'binding activity'(9), 'receptor activity'(17), 'catalytic activity'(18) and 'antioxidant activity'(1). 'Biological process classification' suggests that 29 out of 57 genes were mainly mapped to 'cellular process', 'localization', 'response to stimulus', 'metabolic process' and 'immune system process', with over 18 genes in each catalog (Fig. 1a). In 'cell components classification analysis', 5 out of 51 genes were mapped (3 genes encode 'membrane', 4 encode 'cell part' and 1 encodes 'organelle' and the other 1 encodes 'extracellular region'). Regarding 'pathway classification' analysis, 17 out of 51 genes were found to be involved in 7 pathways (Fig. 1b). Specifically, *YES1* from 'CCR3 signalling map pathway', *Q86XF0* and *P00374* from 'formyltetrahydroformate biosynthesis pathway', 10 genes (*CCR3*, *CCR1*, *CXCR1*, *CCR5*, *CXCR2*, *CXCR6*, *CCR6*, *CCR4*, *XCR1* and *CCR7*) from 'inflammation mediated by chemokine and cytokine signalling pathway', *DHFR1* and *DHFR* from 'folate biosynthesis', *YES1* and *FGR* from 'parkinson disease pathway', *CXCR2*, *CXCR1*, *IL17A*, *IL17F*, *IL10* from 'Interleukin signalling pathway', *YES1* from 'cadherin signalling pathway'. After classification of those HGT genes, we performed a GO enrichment analysis. In total 46 GO terms were found to be enriched with the p value smaller than 0.05 (Supplementary Table S3). Surprisingly, the GO terms

GO biological process	P value
Chemokine-Mediated Signaling Pathway (GO:0070098)	5.70E-18
Dendritic Cell Chemotaxis (GO:0002407)	1.43E-09
Dendritic Cell Migration (GO:0036336)	7.27E-09
Inflammatory Response (GO:0006954)	7.62E-09
Chemotaxis (GO:0006935)	2.05E-08
Taxis (GO:0042330)	2.05E-08
Cellular Response To Cytokine Stimulus (GO:0071345)	2.93E-08
Cytokine-Mediated Signaling Pathway (GO:0019221)	3.06E-08
Response To Cytokine (GO:0034097)	1.75E-07
Defense Response (GO:0006952)	8.64E-07

**Table 1. The top 10 GO enrichment of the 51 human encoded HGT genes.**

predominantly enriched in immune related terms. The most significant GO term is ‘chemokine-mediated signaling pathway’ with a p value 5.70E-18, followed by ‘dendritic cell chemotaxis’ with p value 1.43E-09 and then ‘dendritic cell migration’ (p value 7.27E-09). Actually, the top 10 GO terms are all immune related (Table 1). The disproportionate of enriched immune related GO terms strongly suggest that the HGT events between viruses and hosts have a clear link with immune response. When we checked the functional annotation of the 51 gene list, we found 27 genes were found to be directly related to immune response. Specifically, 19 genes encode chemokine receptors, 5 genes encode interleukin(interleukin 1 receptor antagonist, interleukin 1 receptor, type II, interleukin 10, interleukin 17A, interleukin 17F), and 3 genes encode cluster of differentiation (CD) molecules (CD200 molecule, CD59 molecule and CD69 molecule).

#### The 27 viral HGT genes are encoded by herpesviridae, poxviridae, retroviridae and bornaviridae.

Next, we investigated the origin of the HGT related genes derived from the viral genomes (Supplementary Table S1). The result shows that the 27 viral HGT related genes are found to be encoded by 14 virus species. Specifically, 5 out of 14 species belongs to herpesviridae (human herpesvirus type 5, 6, 8, equine herpesvirus 2, Saimiriine herpesvirus 2), 5 from poxviridae (deerpox virus, molluscum contagiosum virus, vaccinia virus, fowlpox virus, yaba monkey tumor virus), 3 from retroviridae (rous sarcoma virus, baboon endogenous virus, simian retrovirus), with the remaining one from bornaviridae (bornavirus). The common characteristics those viruses share is that they have the capability to establish latent infection, indicating that they are able to modify and evade the immune system of hosts<sup>32–36</sup>.

#### Orthologs of immune related HGT genes were mainly found in viruses capable of establishing latent infection.

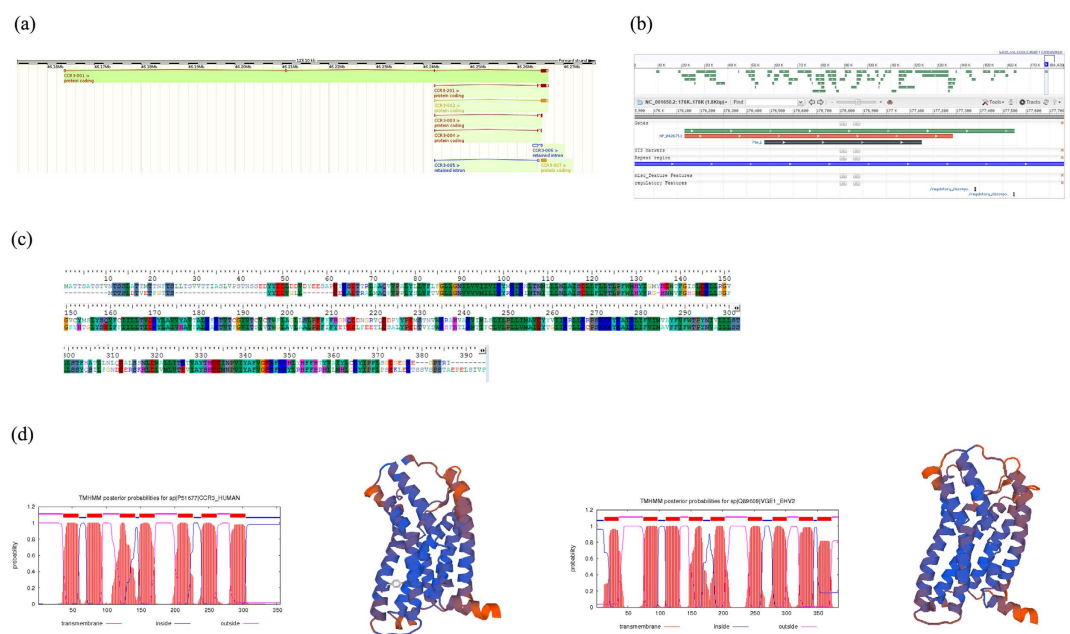
To test whether the 27 immune related HGT genes encoded by human have any homologs in other viruses species, we performed a blastp analysis using 27 viral HGT proteins against virus protein database. As the function as of a particular protein is largely determined by its structure, we usually infer the function of an unknown protein based on its similarity to the well-studied protein through homology structure modeling. However, the accuracy of modeling is highly dependent on the sequence identity between two proteins. Models tend to be reliable when sequence identity is above 50%, become less confident in the 30–50% range and show low-accuracy under 30% identity<sup>37</sup>. Therefore, we used two sequence identity cut-off values (30% and 50%) in the ortholog analysis. Our results indicated that there were 1019 records in viruses sharing similarity with the 27 immune related HGT genes encoded by human (Supplementary Table S4). When using a cut-off value of identity greater than 30% and removing the redundant hits, 151 unique viral proteins were left (Supplementary Table S5). We then traced the organisms of those proteins, those 151 viral proteins were mapped to 37 virus families (arteriviridae, asfarviridae, baculoviridae, betaflexiviridae, bromoviridae, bunyaviridae, caliciviridae, caulimoviridae, chrysoviridae, closteroviridae, coronaviridae, flaviviridae, hepadnaviridae, herpesviridae, iridoviridae, malacoherpesviridae, mimiviridae, myoviridae, narnaviridae, ophioviridae, orthomyxoviridae, paramyxoviridae, parvoviridae, phycodnaviridae, picornaviridae, plasmaviridae, podoviridae, poxviridae, reoviridae, retroviridae, rhabdoviridae, siphoviridae, togaviridae, tombusviridae, virgaviridae). Among those, 35 out of the 151 proteins belong to herpesviridae, 18 proteins belong to poxviridae and 11 proteins belong to retroviridae (Supplementary Table S4). In total around 42% (64 out of 151) viral proteins sharing more than 30% identity belong to herpesviridae, poxviridae and retroviridae. Viruses in those three families are well known for their ability of immune evasion and establishment of latent infection. When we increase the cut-off value of identity from 30% to 50%, this trend become even more obvious, only 12 unique viral proteins meet this criteria, and predominantly belongs to herpesviridae (10 out of 12). The remaining two proteins were encoded by retroviridae and bromoviridae respectively (Table 2).

#### Similarity between viral encoding G-protein coupled receptor E1 and human encoding chemokine receptors.

Of particular interest, we found out that 13 out of the 27 human immune related HGT proteins (chemokine (C-C motif) receptor (CCR) 1–7, CCR9, CCR-like 1 and CCR-like 2, C-X-C motif chemokine receptor 3, 4 and 6) harbor the best hit with the same viral protein, G-protein coupled receptor E1 (Q89609) encoded by equine herpesvirus 2 (EHV-2). The identity between those 13 host proteins and E1 vary from 33.5% to 55.8%. As the identity between CCR3 and E1 is the highest (55.8%), we performed a comparison analysis of protein sequence, transmembrane domains and 3D structure between CCR3 and E1. Overall, CCR3 and E1 exhibit very

Entry	Gene names	Taxonomic lineage (FAMILY)
Q8UYT3	ORF2b	Bromoviridae
Q00996	15	Herpesviridae
P04048	V-KIT	Retroviridae
Q89609	E1	Herpesviridae
Q69140	EBNA6 BERF3-BERF4	Herpesviridae
P68677	E7	Herpesviridae
P68678	E7	Herpesviridae
P0CAP9	BCRF1	Herpesviridae
P03180	BCRF1	Herpesviridae
P0C6Z6	BCRF1	Herpesviridae
P24916	13 KCLF2	Herpesviridae
O40633	13	Herpesviridae

**Table 2.** List of viral genes sharing similarity (identity  $\geq 50\%$ ) to 27 human immune related HGT genes.

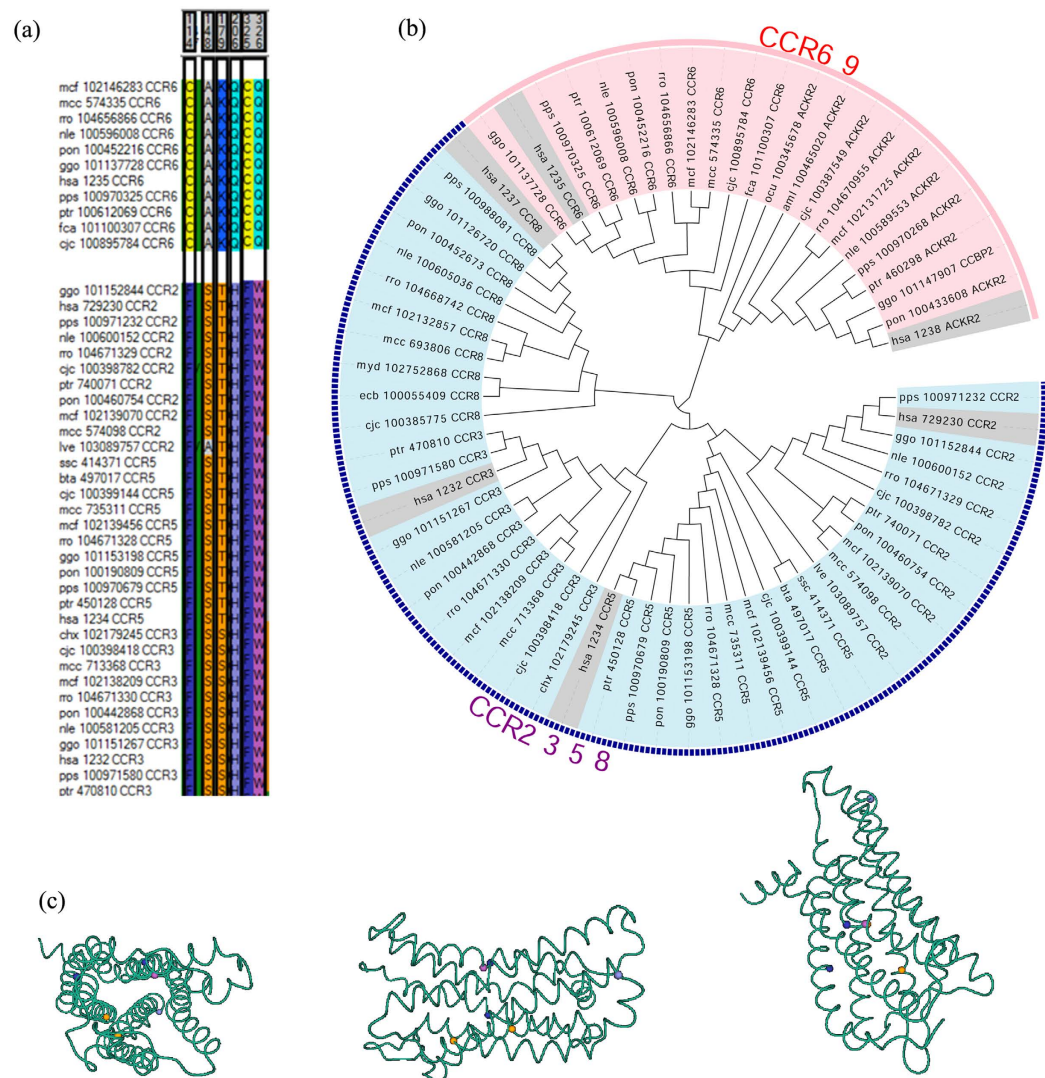


**Figure 2.** The similarity of protein sequence, secondary structure and 3D structure between human encoding CCR3 and viral protein E1. (a) Gene model of human CCR3 (b) Gene model of EHV E1 (c) Alignment of protein sequences between CCR3 and E1 (d) Transmembrane domains and predicted 3D structure of CCR3 and E1.

high degree of similarity. Due to alternative splicing, CCR3 possesses 6 isoforms CCR3-001 (2000bp, 355aa), CCR3-201 (1786bp, 376aa), CCR3-002 (1581bp, 355aa), CCR3-007 (1201bp, 355aa), CCR3-003 (400bp, 80aa) and CCR3-004 (212bp, 21aa). E1 is composed of 1149 nucleotides, which is subsequently translated into 383 aa, with 28 more aa than CCR3-001. TMHMM<sup>38</sup> transmembrane analysis suggests that CCR3 contains 7 transmembrane domains while E1 contains 9 domains. SWISS MODEL<sup>39</sup> predicted 3D structures indicate that the two proteins share almost identical structures except for some differences in the terminal regions (Fig. 2).

### Functional divergence after gene duplication contributes to the evolution of CCRs gene family in human.

Gene duplication serves as an important source of functional divergence, as duplicated genes may gain novel functions or lose original ones through modifying their existing regulatory network by accumulation of mutations in coding or non-coding regions<sup>40,41</sup>. There are 9 copies of C-C chemokine receptors in human genome which share distinct extent of similarity to each other, it is possible that those genes come from functional divergence after gene duplication. Using DIVERGE 3 developed by Gu's group<sup>42,43</sup>, we predicted the type I and type II functional divergence residuals among CCR paralogs. In type I functional divergence, the site is conserved in one group of paralogs while variable in the others. In contrast, type II functional divergence is the case that residuals are conserved in both paralog groups, though the property of amino acids differed between two clusters. Firstly, we constructed a maximum likelihood tree using 66 proteins sequences of CCR 2, 3, 5, 6, 8, 9 downloaded from KEGG. There are two distinct clades on the tree, indicating that all the 66 proteins are divided into two clusters.

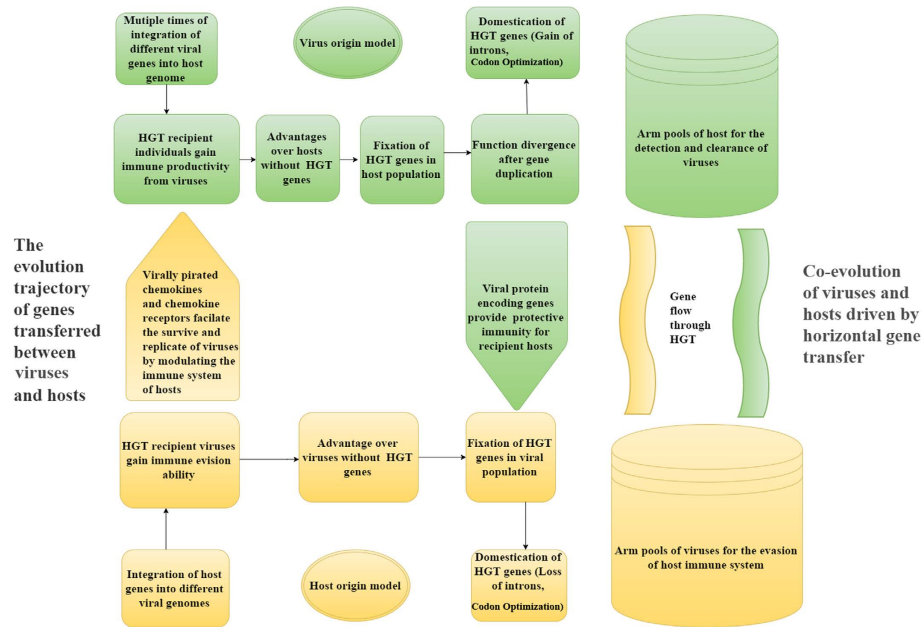


**Figure 3. Functional divergence of CCR3s paralogs.**

Cluster A includes CCR6 and CCR9 while cluster B is composed of CCR 2, 3, 5 and 8 (Fig. 3b). We performed the type I and type II functional divergence analysis, which revealed that there was no type I residuals, whereas 6 type II sites were identified at alignment positions 114, 148, 179, 206, 325 and 326 (Fig. 3a). We next studied the property of those amino acids in two clusters and found those residuals change from moderate to hydrophobic, hydrophobic to hydrophilic, positive charge to no charge and hydrophobic to moderate. We also mapped those functional divergence residuals to the 3D structure of CCR3, which supports the deduction that those 5 type II functional divergence residuals contribute to the functional difference between cluster 1 CCRs and cluster 2 CCRs (Fig. 3c).

## Discussion

The gene flow between viruses and hosts normally happens after the integration of viral genomic fragments or whole genomes into host chromosomes. This phenomenon occurs frequently in retroviruses such as human immunodeficiency virus (HIV) through the integration of double stranded DNA (dsDNA) which were reverse transcribed from viral RNA genome<sup>44,45</sup>. As for the dsDNA viruses, several members of herpesviruses such as human herpesvirus 6 (HHV-6) and marek's disease virus (MDV) have been found to be capable to integrate their genomes into host chromosomes under certain circumstances<sup>46–48</sup>. It is possible that herpesviruses integrate into the germ line chromosomes of the ancestor of homo sapiens, probably through non-homologous recombination or interaction with cellular retroelements<sup>49,50</sup>. If the immune related genes were horizontally transferred from viruses to host, the foreign genes were likely to be retained in the population for providing protective immunity for recipient hosts. Since then, the laterally transferred genes experienced functional divergence after gene duplication, which subsequently evolved into a family of CCRs with different functions. In our study, we predicted 6 amino acid residuals which might contribute to the functional divergence of CCR paralogs in human, however further research need to be done to validate the functional importance of those residuals. On the other hand, if those foreign genes were transferred from host to viruses, it would be a quite dissimilar story, viruses might acquire genomic fragments containing host origin genes when switching from latency to lytic infection.



**Figure 4. Horizontal gene transfer drives the co-evolution of hosts and viruses.**

In a recent report by Crisp *et al.*, the direction of gene transfer between viruses and humans remains an open question<sup>29</sup>. In our opinion, there are two possible models to explain the HGT direction, the ‘host origin model’ in which those HGT genes were laterally transferred from host to viruses, and the ‘virus origin model’ in which those HGT genes were transferred in an opposite direction. In theory, these two scenarios could happen with equal chance. In the former scenario, namely ‘host origin model’, the integration of viral structural gene into host genome could lead to the expression of viral antigen which would stimulate the immune system to protect itself against the infection of viruses. In other words, host might utilize the viral DNA as ‘DNA vaccine’ by acquiring and expressing viral protein encoding genes<sup>51,52</sup>. However, the HGT genes are unlikely to be fixed in host population unless they meet two criteria. Firstly, the HGT genes need to be integrated into the chromosome of sperms, otherwise the HGT will be difficult to be inherited by offsprings of the HGT recipient individuals. Secondly, the HGT genes should provide enhanced fitness to the environment compared to counterparts without the transfer of foreign genes. Once the viral origin HGT genes got spread and fixed in the population, they probably experienced domestication such as the optimizing of codons and the gain of introns. During millions of years, those HGT genes evolved into a family of proteins playing diverse roles in immune system. In contrast, in the second scenario (‘host origin model’), those immune related factors might be used by viruses to trick the immune response of hosts to achieve the goal of immune evasion. Viruses have been reported to hijack host immune response related signaling pathways by mimicking host immune modifiers, either cytokines or chemokines, in order to facilitate the replicate and survive of viruses inside host cells<sup>53–60</sup>. In our study, among the 51 human encoding HGT genes transferred between human and viruses, several immune related genes were found to share high similarity with viral proteins encoded by herpesviridae, poxviridae and retroviridae. As a variety of viruses have been shown to restrict the immune system of the host by protein imitating the host immune functional molecular<sup>61,62</sup>, it is highly possible that those viruses investigated in this study utilize a similar mechanism to tackle with the host immune system. Viruses carrying genes related to immune response might facilitate them to evade the immune system, which enhances their opportunity to survive the attack from host immune system and pass on their genetic elements to progeny viruses. Taken together, ‘host origin model’ and ‘virus origin model’ are not necessarily mutually exclusive. It could be possible that some genes were transferred from hosts to viruses while others from viruses to hosts. And it is the bi-directional gene flow of immune related genes through HGT that drives the co-evolution of viruses and hosts and altered the attack and defense strategies of players during the lasting war among the two sides (Fig. 4).

The comparison analysis suggests a high similarity of both secondary structure and 3D structure between viral protein E1 and host protein CCR3, implying they share similar biological functions as well. As CCR3 is well known for its importance in inducing the migration of immune cells into inflammatory sites<sup>63–65</sup>, it is reasonable to speculate that viral protein E1 is also related to immune response. Since the particular mechanism by which E1 manipulate host immune system is not yet fully understood, more work needs to be done to clarify the roles of E1 in the interaction between viruses and their hosts. Intriguingly, the 51 HGT genes in human are enriched in immune related GO terms, which suggests the HGT events are tend to happen among immune genes. In addition, the 27 viral HGT genes are encoded by viruses from herpesviridae, retroviridae and bornaviridae, which are capable to establish latent infection, suggesting HGT might contribute to the immune evasion of those viral species. In conclusion, our study suggests that the HGT between viruses and hosts might play a pivotal role in reshaping the arm race between viruses and hosts, which eventually drives their co-evolution.

## Methods

**Classification and GO enrichment analysis of gene lists.** The genes lists were analysed using PANTHER<sup>31</sup> classification tool with the default parameters. The GO enrichment analysis was performed using online tool of Gene Ontology Consortium<sup>66</sup>. Specifically, PANTHER overrepresentation test (release 20150430) with Bonferroni correction and GO ontology database (released 2015-08-06) were used to evaluate the enrichment of GO terms for the 51 human HGT genes, all the GO terms with P value greater than 0.05 were removed.

**Blast of 27 human encoding immune related genes.** The virus protein database were created using makeblast command from blast package, using the virus sequences downloaded from Uniprot<sup>67</sup>. Blastp programme was used to scan homologs for query protein sequences. Then 27 human immune related HGT protein sequences were blasted against the virus protein databases. The Blastp results were outputted as table format using the parameter outfmt 6. The list of Blastp hit records were extracted for further analysis. Unique viral protein Id were extracted by removing the redundant contents. To investigate the origin of those proteins, the protein IDs were submitted to UniProt using retrieve/ID mapping tool.

**Transmembrane domains prediction.** The transmembrane protein topology of E1 and CCR3 were predicted with TMHMM<sup>38</sup> while 3D structures were modelled using SWISS MODEL<sup>39</sup> with default parameters.

**Functional divergence analysis of CCRs genes.** The orthologs of human CCRs were downloaded from KEGG database and analysed using DIVERGE 3<sup>42,43</sup>.

## References

1. Yue, J., Hu, X., Sun, H., Yang, Y. & Huang, J. Widespread impact of horizontal gene transfer on plant colonization of land. *Nat. Commun.* **3**, 1152 (2012).
2. Pál, C., Papp, B. & Lercher, M. J. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37**, 1372–5 (2005).
3. Ochman, H., Lawrence, J. G. & Groisman, E. A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304 (2000).
4. Gogarten, J. P. & Townsend, J. P. Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **3**, 679–87 (2005).
5. Keeling, P. J. & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9**, 605–18 (2008).
6. Koonin, E. V., Makarova, K. S. & Aravind, L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**, 709–42 (2001).
7. Polz, M. F., Alm, E. J. & Hanage, W. P. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* **29**, 170–5 (2013).
8. Boto, L. Horizontal gene transfer in evolution: facts and challenges. *Proc. Biol. Sci.* **277**, 819–27 (2010).
9. Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* **16**, 472–482 (2015).
10. Freeman, V. J. Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *J. Bacteriol.* **61**, 675–88 (1951).
11. Klassen, J. L. Pathway evolution by horizontal transfer and positive selection is accommodated by relaxed negative selection upon upstream pathway genes in purple bacterial carotenoid biosynthesis. *J. Bacteriol.* **191**, 7500–8 (2009).
12. Novozhilov, A. S., Karev, G. P. & Koonin, E. V. Mathematical modeling of evolution of horizontally transferred genes. *Mol. Biol. Evol.* **22**, 1721–32 (2005).
13. Mayer, W. E., Schuster, L. N., Bartelmes, G., Dieterich, C. & Sommer, R. J. Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover. *BMC Evol. Biol.* **11**, 13 (2011).
14. Richards, T. A. *et al.* Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proc. Natl. Acad. Sci. USA* **108**, 15258–63 (2011).
15. Sun, B.-F. *et al.* Horizontal functional gene transfer from bacteria to fishes. *Sci. Rep.* **5**, 18676 (2015).
16. Kim, W. M. & Sigalov, A. B. Viral pathogenesis, modulation of immune receptor signaling and treatment. *Adv. Exp. Med. Biol.* **640**, 325–49 (2008).
17. Moelling, K., Matskevich, A. & Jung, J.-S. Relationship between retroviral replication and RNA interference machineries. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 365–8 (2006).
18. McFadden, G. & Murphy, P. M. Host-related immunomodulators encoded by poxviruses and herpesviruses. *Curr. Opin. Microbiol.* **3**, 371–8 (2000).
19. Alcamí, A. & Koszinowski, U. H. Viral mechanisms of immune evasion. *Immunol. Today* **21**, 447–55 (2000).
20. Tortorella, D., Gewurz, B. E., Furman, M. H., Schust, D. J. & Ploegh, H. L. Viral subversion of the immune system. *Annu. Rev. Immunol.* **18**, 861–926 (2000).
21. Monier, A. *et al.* Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. *Genome Res.* **19**, 1441–9 (2009).
22. Koonin, E. V. & Dolja, V. V. Evolution of complexity in the viral world: the dawn of a new vision. *Virus Res.* **117**, 1–4 (2006).
23. Katzourakis, A., Tristem, M., Pybus, O. G. & Gifford, R. J. Discovery and analysis of the first endogenous lentivirus. *Proc. Natl. Acad. Sci. USA* **104**, 6261–5 (2007).
24. Filé, J., Forterre, P. & Laurent, J. The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies. *Res. Microbiol.* **154**, 237–43 (2003).
25. Awadalla, P. The evolutionary genomics of pathogen recombination. *Nat. Rev. Genet.* **4**, 50–60 (2003).
26. Moreira, D. & López-García, P. Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* **7**, 306–11 (2009).
27. Moreira, D. & Brochier-Armanet, C. Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. *BMC Evol. Biol.* **8**, 12 (2008).
28. Bratke, K. A. & McLysaght, A. Identification of multiple independent horizontal gene transfers into poxviruses using a comparative genomics approach. *BMC Evol. Biol.* **8**, 67 (2008).
29. Crisp, A., Boschetti, C., Perry, M., Tunnacliffe, A. & Micklem, G. Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol.* **16**, 50 (2015).
30. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–66 (2013).
31. Mi, H., Poudel, S., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res.* **44**, D336–42 (2015).
32. Cooper, L. N. Theory of an immune system retrovirus. *Proc. Natl. Acad. Sci. USA* **83**, 9159–63 (1986).
33. Doehle, B. P. & Michael Gale, J. Innate Immune Evasion Strategies of HCV and HIV: Common Themes for Chronic Viral Infection. (2013).

34. Griffin, B. D., Verweij, M. C. & Wiertz, E. J. H. Herpesviruses and immunity: the art of evasion. *Vet. Microbiol.* **143**, 89–100 (2010).
35. Seet, B. T. *et al.* Poxviruses and immune evasion. *Annu. Rev. Immunol.* **21**, 377–423 (2003).
36. Buller, R. M. & Palumbo, G. J. Poxvirus pathogenesis. *Microbiol. Rev.* **55**, 80–122 (1991).
37. Baker, D. & Sali, A. Protein structure prediction and structural genomics. *Science* **294**, 93–6 (2001).
38. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–80 (2001).
39. Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252–8 (2014).
40. Conant, G. C. & Wolfe, K. H. Turning a hobby into a job: how duplicated genes find new functions. *Nat. Rev. Genet.* **9**, 938–50 (2008).
41. Hughes, A. L. The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.* **256**, 119–24 (1994).
42. Gu, X. & Vander Velden, K. DIVERGE: phylogeny-based analysis for functional-structural divergence of a protein family. *Bioinformatics* **18**, 500–1 (2002).
43. Gu, X. *et al.* An update of DIVERGE software for functional divergence analysis of protein family. *Mol. Biol. Evol.* **30**, 1713–9 (2013).
44. Cohn, L. B. *et al.* HIV-1 integration landscape during latent and active infection. *Cell* **160**, 420–32 (2015).
45. Craigie, R. & Bushman, F. D. HIV DNA integration. *Cold Spring Harb. Perspect. Med.* **2**, a006890 (2012).
46. Morissette, G. & Flamand, L. Herpesviruses and chromosomal integration. *J. Virol.* **84**, 12100–9 (2010).
47. Osterrieder, N., Wallaschek, N. & Kaufner, B. B. Herpesvirus Genome Integration into Telomeric Repeats of Host Cell Chromosomes. *Annu. Rev. Virol.* **1**, 215–235 (2014).
48. Kaufner, B. B., Jarosinski, K. W. & Osterrieder, N. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. *J. Exp. Med.* **208**, 605–15 (2011).
49. Katzourakis, A. & Gifford, R. J. Endogenous viral elements in animal genomes. *PLoS Genet.* **6**, e1001191 (2010).
50. Holmes, E. C. The evolution of endogenous viral elements. *Cell Host Microbe* **10**, 368–77 (2011).
51. Fu, T. M., Friedman, A., Ulmer, J. B., Liu, M. A. & Donnelly, J. J. Protective cellular immunity: cytotoxic T-lymphocyte responses against dominant and recessive epitopes of influenza virus nucleoprotein induced by DNA immunization. *J. Virol.* **71**, 2715–21 (1997).
52. Mancini, M. *et al.* DNA-mediated immunization in a transgenic mouse model of the hepatitis B surface antigen chronic carrier state. *Proc. Natl. Acad. Sci. USA* **93**, 12496–501 (1996).
53. Alami, A. Viral mimicry of cytokines, chemokines and their receptors. *Nat. Rev. Immunol.* **3**, 36–50 (2003).
54. Streblov, D. N. *et al.* The Human Cytomegalovirus Chemokine Receptor US28 Mediates Vascular Smooth Muscle Cell Migration. *Cell* **99**, 511–520 (1999).
55. Bais, C. *et al.* G-protein-coupled receptor of Kaposi's sarcoma-associated herpesvirus is a viral oncogene and angiogenesis activator. *Nature* **391**, 86–9 (1998).
56. Arvanitakis, L., Geras-Raaka, E., Varma, A., Gershengorn, M. C. & Cesarman, E. Human herpesvirus KSHV encodes a constitutively active G-protein-coupled receptor linked to cell proliferation. *Nature* **385**, 347–50 (1997).
57. Samson, M. *et al.* Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**, 722–5 (1996).
58. Berger, E. A., Murphy, P. M. & Farber, J. M. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu. Rev. Immunol.* **17**, 657–700 (1999).
59. Murphy, P. M. Viral exploitation and subversion of the immune system through chemokine mimicry. *Nat. Immunol.* **2**, 116–22 (2001).
60. Sodhi, A., Montaner, S. & Gutkind, J. S. Viral hijacking of G-protein-coupled-receptor signalling networks. *Nat. Rev. Mol. Cell Biol.* **5**, 998–1012 (2004).
61. Spencer, J. V. *et al.* Potent Immunosuppressive Activities of Cytomegalovirus- Encoded Interleukin-10. *J. Virol.* **76**, 1285–1292 (2002).
62. Eberhardt, M. K. *et al.* Host immune responses to a viral immune modulating protein: immunogenicity of viral interleukin-10 in rhesus cytomegalovirus-infected rhesus macaques. *PLoS One* **7**, e37931 (2012).
63. Spinetti, G. The chemokine receptor CCR8 mediates rescue from dexamethasone-induced apoptosis via an ERK-dependent pathway. *J. Leukoc. Biol.* **73**, 201–207 (2003).
64. Goya, I. *et al.* Identification of CCR8 as the Specific Receptor for the Human  $\beta$ -Chemokine I-309: Cloning and Molecular Characterization of Murine CCR8 as the Receptor for TCA-3. *J. Immunol.* **160**, 1975–1981 (1998).
65. Yabe, R. *et al.* CCR8 regulates contact hypersensitivity by restricting cutaneous dendritic cell migration to the draining lymph nodes. *Int. Immunol.* **27**, 169–181 (2014).
66. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–9 (2000).
67. The universal protein resource (UniProt). *Nucleic Acids Res.* **36**, D190–5 (2008).

## Acknowledgements

D.S.C., Y.Q.W., W.Z. and S.J.J. are funded by the scholarship from the China Scholarship Council.

## Author Contributions

S.Z.C. and D.S.C. designed the study. D.S.C., Y.Q.W., W.Z. and S.J.J. performed the data analysis. S.Z.C., D.S.C., Y.Q.W., W.Z. and S.J.J. wrote the manuscript. All authors approved the final version of this manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Chen, D.S. *et al.* Horizontal gene transfer events reshape the global landscape of arm race between viruses and homo sapiens. *Sci. Rep.* **6**, 26934; doi: 10.1038/srep26934 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>