A case of horizontal gene transfer from *Wolbachia* to *Aedes albopictus* C6/36 cell line

Qing Hou,1,† Ji He1,2,† Jing Yu,1 Yuting Ye,1 Dan Zhou,1 Yan Sun,1 Donghui Zhang,1 Lei Ma,1 Bo Shen,1 and Changliang Zhu1,†

1Department of Pathogen Biology; Nanjing Medical University; Nanjing, Jiangsu People’s Republic of China; 2State Key Laboratory of Surveillance and Detection for Medical Vectors; Xiamen Entry–Exit Inspection and Quarantine Bureau; Xiamen, Fujian People’s Republic of China; †These authors contributed equally to this work.

**Introduction**

Horizontal gene transfer (HGT) is the exchange of genetic elements between phylogenetically distant and reproductively isolated species or organisms.1 HGT was traditionally thought to be unlikely between prokaryotes and eukaryotes.2 However, recent studies have revealed that microbial endosymbiotic DNA fragments can integrate into unrelated eukaryotic organisms.3-5 Although, this process is still considered extremely rare, endosymbiotic gene transfer is believed to be possible because of the intimate symbiotic relationship.1

*Wolbachia* is a very important and common bacterial endosymbiont,6 infecting around 66% of arthropods worldwide including mosquitoes, *Drosophila*, and beetles.7,8 *Wolbachia* is able to invade rapidly and spread widely among the host arthropod population, and can manipulate its host’s reproduction and also prevent its host from infecting humans with RNA viruses and other pathogens.9-11 This behavior of *Wolbachia*, along with the increasing resistance of many arthropods to insecticides, has led to considerable research aimed at using *Wolbachia* as a biological control for insect-borne diseases.12,13 Research efforts have particularly focused on mosquito-borne diseases including malaria, dengue fever, West Nile fever, lymphatic filariasis, which become a significant concern to global human health and cause the mortality and morbidity of hundreds of thousands of people annually.14-16 A large range of insect hosts can be naturally infected with *Wolbachia*, including four mosquito species, *Culex pipiens*,17 *Culex quinquefasciatus*, *Aedes fluviatilis*,18 and *Aedes albopictus*.19 such an important vector of dengue, arguably the most important arboviral diseases of humans, leading to tens of thousands of deaths globally each year.

Recent studies have revealed that genetic fragments, ranging in size from single genes to even entire genome, have horizontally transferred from *Wolbachia* to their insect hosts as illustrated below. Genes originating from *Wolbachia* were identified and located on the X chromosome of the adzuki bean beetle, *Callosobruchus chinensis*.20 Hotopp et al.,3 meanwhile, found nearly the entire *Wolbachia* genome in *Drosophila ananassae* Hawaii, while much smaller fragments were found in the other three insects and four nematode species. However, these horizontally transferred genes were all found during tetracycline treatment and their transcription levels were far lower than those of the control gene (act5C) in their recipient, suggesting they could not be expressed or were potentially nonfunctional.3 Therefore, the mechanism and process of evolution of *Wolbachia* mediated HGT remains poorly understood. A natural gene-exchange model is required to more fully understand the processes of HGT evolution.

C6/36 cell line was derived from *Ae. albopictus* which can be naturally infected with both *Wolbachia* wAlbA and wAlbB strains.19 However, it is accepted that *Ae. albopictus* C6/36 cell lines lack *Wolbachia* endosymbionts (based on professor Scott O’Neill’s personal communication). In this study, we report the discovery of a gene WP0273(C6/36) within *Wolbachia* or were potentially nonfunctional.3 Therefore, the mechanism and process of evolution of *Wolbachia* mediated HGT remains poorly understood. A natural gene-exchange model is required to more fully understand the processes of HGT evolution.

C6/36 cell line was derived from *Ae. albopictus* which can be naturally infected with both *Wolbachia* wAlbA and wAlbB strains.19 However, it is accepted that *Ae. albopictus* C6/36 cell lines lack *Wolbachia* endosymbionts (based on professor Scott O’Neill’s personal communication). In this study, we report the discovery of a gene WP0273(C6/36) within *Wolbachia*-uninfected C6/36 cells of *Ae. albopictus* and it is highly similar to the gene originating from *Wolbachia* of *Cx. quinquefasciatus* Pel wPip strain (the putative transcriptional regulator, WP0273). We
further revealed the high transcription level of the horizontally transferred gene WP0273(C6/36) and demonstrated that WP0273(C6/36) encodes a protein in the host cell. Taken together, these findings strongly suggest that this represents a natural HGT event from Wolbachia to Ae. albopictus C6/36 cell line, which is involved in a particular functional role.

**Results**

The expression of Wolbachia gene in the *Aedes albopictus* C6/36 cell line

Protein-protein interactions (PPI) are important for the majority of biological functions. We have found that ribosomal protein L39 (RPL39) of *Culex pipiens pallens* is a deltamethrin resistance-associated protein,\(^{21,22}\) to explore whether there are some proteins interacting with RPL39 and participate the mechanism of deltamethrin resistance in *Cx. Pipiens pallens*, we fused the protein RPL39 with tandem affinity peptides, expressed it in *Ae. albopictus* C6/36 cells and then co-purified the binding proteins through two affinity steps. Surprisingly, our results showed for the first time, a domain of 97 amino acids that is the component of the putative transcriptional regulator of Wolbachia endosymbiont of *Culex quinquefasciatus* Pel (YP 001975078.1) was 87% identical by mass spectrometry (Fig. 1A). Additional, we demonstrated that WP0273(C6/36) protein binds to RPL39 protein in a His pull-down assay (Fig. 1B).

Specific PCR detection of Wolbachia WP0273(C6/36) gene

To confirm the absence of Wolbachia in *Ae. albopictus* C6/36 cell lines, PCR detection of Wolbachia was conducted using the wsp gene\(^{23,24}\) with the Wolbachia infected *Ae. albopictus* strains as a positive control, we amplified the fragment of expected size in the DNA samples of *Ae. albopictus* but did not in the C6/36 cells. The results indicate that Wolbachia is not present in the C6/36 cell line, and the presence of WP0237(C6/36) in this cell line probably results from an HGT event.

Specific amplification of the WP0237 gene in C6/36 cells was performed and a DNA segment of 905bp size was detected, which was still amplified in the tetracycline treated C6/36 cells after three generations. DNA sequencing and BLASTN analysis revealed that the target gene WP0273(C6/36) (GenBank accession number KF283997) shared the highest similarity (99% identity) with Wolbachia of *Cx. pipiens pallens* wPip strain transcriptional regulator WP0273 (GenBank AM999887.1) according to the NCBI database (Fig. 2).

**Phylogenetic analysis**

To identify the phylogenetic position of WP0273(C6/36), we used the nuclear acid sequence of WP0273, as well as transcriptional regulators of *Wolbachia* of mosquito and *Drosophila* hosts, and choose the transcriptional regulator of *Wolbachia* of *Brugia malayi* as the outgroup sequence, to construct phylogenetic tree using MEGA5.1.\(^{25}\) The earliest branching sequences were all from *Wolbachia* strains infecting in mosquito and *Drosophila*. The phylogenetic distances in these trees suggested that WP0273(C6/36) was derived from endosymbiont *Wolbachia* (Fig. 3). This indicates that the WP0273(C6/36) gene found in the mosquito genome cannot be explained by sequence conservation, and was rather inserted into the mosquito genome through HGT.

**Estimating the transcriptional level of the horizontally transferred gene WP0273(C6/36)**

Quantitative PCR analysis was conducted to assess the transcriptional level of WP0273(C6/36) relative to a control gene, act5C, which has been shown to be highly transcribed in the *Drosophila* genome and thus offers an effective comparison for assess whether horizontally transferred genes are likely to be functional.\(^{3,20,26}\) Our results showed that the transcriptional level of WP0273(C6/36) was 1.7 x 10^4 times lower than that of act5C of C6/36 cells (Fig. 4). This value is significantly higher than the transcriptional level of most horizontally transferred genes from Wolbachia to insect hosts recorded in previous studies (estimated to be 10^4–10^7 lower than the control). This result indicates clearly that WP0273(C6/36) was transcribed.

**Demonstrating the protein expression of WP0273(C6/36) in the host cells**

We constructed the prokaryotic expression vector via recombination of a pET-32a plasmid and the open reading frame of WP0273(C6/36), and expressed the protein for polyclonal antibody preparation. Western immunoblotting was conducted to demonstrate the expression of protein WP0273(C6/36) with non-blood-sucking female and Wolbachia-free *Anopheles sinensis* mosquitoes as a negative control. The WP0273(C6/36) gene translated into the protein in the environment of mosquito host cells. Conversely, the WP0273 gene did not express in the *Anopheles sinensis* (Fig. 5).

**Analysis of protein sequence and prediction of protein function**
We constructed the secondary structure and a 3D model of the WP0273(C6/36) protein. The detailed template information of all alignment coverage suggested that the WP0273(C6/36) protein was very likely a transcriptional regulator or a DNA-binding protein. Subsequently, motif scan results revealed that the whole sequence contained two helix-turn-helix motifs, which were composed of two α-helix motifs joined by a short strand of amino acids and a major structure capable of binding DNA and which can regulate gene expression. We also found a structure highly similar to the Ankyrin repeat domain (ANK), which could mediate protein–protein interaction and has previously been found in some arthropod-infesting Wolbachia.27
The horizontally transferred gene WP0273(C6/36) has met all the needs described above, suggesting WP0273(C6/36) has obtained a new function. However, what the new capacity of WP0273(C6/36) is and which role it plays in the host cell remains unknown. Further studies are now required to determine whether WP0273(C6/36) acts to regulate transcription or engenders other effects.

Based on our results examining protein structure and motif analysis, we hypothesize that WP0273(C6/36) probably regulates gene expression in C6/36 cells via DNA-binding sites. We will aim to test this hypothesis in future work using Chromatin Immunoprecipitation (Chip) and EMSA-electrophoretic mobility shift assays. Ankyrin repeats consist of 33 amino-acid sequences and the first ANK-containing proteins to be characterized were the yeast cell cycle regulator Swi6/Cdc10 and the Drosophila cell signaling Notch protein. ANK abounds in arthropod-infecting Wolbachia (for example, 54 in wPip, 23 in wMel, and 34 in wAna). Ank may play an interesting role in the relationship between host and Wolbachia because ANK is one of the most common protein-protein interaction motifs in nature, involved in many physiological processes including cell signaling, apoptosis, and cell cycle control. Previous results have showed that the protein WP0273(C6/36) could interact with the protein RPL39, indicating it is worthy of further study as there are any other features of WP0273(C6/36).

HGT is a crucial driving force in bacterial evolution, with a remarkable impact on pathogenicity and antibiotic resistance of human associated microbes. Transfer of the vanA gene from an E. faecium isolate of animal origin to an E. faecium isolate of human origin can occur in the intestines of humans. "Pathogenicity islands" horizontally acquired were major contributors to the virulence of many pathogenic bacteria. Successful adaptive HGT is traditionally considered beneficial either to hosts or to the transferred genes. However, the mechanisms by which Wolbachia is able to manipulate its insect host (to affect reproductive characters, disrupt pathogen infection or shorten hosts lifespan) is still poorly understood. One hypothesis posits that some Wolbachia strains interfere with a range of human pathogens in a manner correlated with the innate immune response in the insect.

Here, we propose that interaction of Wolbachia-host and the regulation of insect behavior by Wolbachia are mediated by the
effect of horizontally transferred genes that can obtain a new function in the recipient.

It is unknown whether there are other cases of HGT between Wolbachia and Ae. albopictus C6/36 cell line. Elucidating this will require a large-scale exhaustive search using PCR and specific quantitative PCR detection in the C6/36 cell line. Should additional transcriptionally active genes be detected, then the Ae. albopictus C6/36 cell line will offer an ideal model to study HGT mechanisms and evolutionary processes. It will nevertheless be necessary to identify the biological function of any further transferred genes identified, including WP0273(C6/36).

The results of this study provide strong evidences that WP0273(C6/36) discovered in Ae. albopictus C6/36 cells is an event of HGT and it appears to be functional. Importantly, our results highlight a novel mechanism of Wolbachia-host interaction and establish a basis for the further study of HGT between endosymbionts and eukaryotes.

### Materials and Methods

**Cell culture and mosquito strains**

*Anopheles sinensis* and *Aedes albopictus* C6/36 cells were obtained from the China Center for Type Culture Collection (Wuhan, China). Cells were maintained in DMEM/High Glucose media (Hyclone) in a 5% CO₂ humidified incubator at 28 °C. *Aedes albopictus* was provided by Disease Prevention and Control Center of Nanjing Military Region. Non-blood-sucking female *Anopheles sinensis* were supplied by Jiangsu Institute of Parasitic Diseases (Wuxi, China).

**Tandem affinity purification and mass spectrometry**

C6/36 cells were stably transfected with the pIB/V5-GS vector (pIB/V5-GS vector was introduced with ORF of RPL39 and GS tag sequence in-frame fusion which was inserted between EcoRV and NotI sites). Cells were then lysed and the purified eluted proteins separated on a SDS-PAGE gel and stained with silver. The differentially expressed protein was then analyzed by mass spectrometry performed on a time-of-flight Ultraflex II mass spectrometer (Biflex).

**Plasmid Construction**

Rosetta (DE3) strain of *Escherichia coli* was used for the protein expression. pGEX-6P-1 GST expression vector was used to express GST-RPL39 fusion protein, the ORF of RPL39 was inserted between BamHI and EcoRI sites. pET32a vector was used to express His-WP0273(C6/36) fusion protein, the ORF of WP0273(C6/36) was inserted between EcoRI and NotI sites.

### Table 1. Primers used for detection of Wolbachia and WP0273(C6/36)

| Primer | Primer sequence (5’-3’) | Product size (bp) | Reference |
|--------|-------------------------|-------------------|-----------|
| wsp    | F-TGGTCCAAAAAGGAAGAAAAC | 590-632           | 23, 24    |
| WP0273(C6/36) | F-ATGTTTTGTCTGTAAGGATATT | | |

**In vitro His pull-down assay**

A total of 1 ug purified His or His-WP0273(C6/36) proteins was applied to His resins and incubated for 1 hr at 4 °C. The resins were washed and mixed with purified GST-RPL39 protein and were incubated for 4 hr at 4 °C. After washing, bound proteins were eluted and subjected to SDS-PAGE, followed by western blotting.

**DNA extraction, RNA extraction, and cDNA synthesis**

Total mosquito C6/36 cells DNA was extracted using TAKARA MiniBEST Universal Genomic DNA Extraction Kit Ver4.0 (TAKARA) according to the manufacturer's protocol. Total RNA from mosquitoes and cells was extracted using TriZol Reagent (Invitrogen) according to the manufacturer's protocol. The cDNA was reverse transcribed from the RNA using the SuperScript® VILO™ cDNA Synthesis kit (Invitrogen) according to the manufacturer's instructions.

**PCR detection of Wolbachia genes**

PCR detection of *Wolbachia* genes was conducted using the primers listed in Table 1, PCR reactions were performed using TAKARA’s Ex TaqVersion 2.0 DNA Polymerase (TAKARA). PCR reactions were composed as follows: Total DNA 200 ng, 25 ul Premix Ex Taq, 1 ul of each 20 uM primer, and ddH₂O was added to bring up the total volume to 50 ul. The amplification of *wsp* gene was performed as follows: 94 °C for 5 min, 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 60 s, followed by a final extension at 72 °C for 7 min. The amplification of WP0273(C6/36) gene was performed as follows: 94 °C for 5 min, 35 cycles of 94 °C for 30 s, 53 °C for 30 s, and 72 °C for 60 s, followed by a final extension at 72 °C for 7 min. The products were separated by 1% agarose gel electrophoresis and purified using TAKARA MiniBEST Agarose Gel DNA Extraction Kit Ver3.0. The purified PCR products were sequenced by the Beijing Genomics Institute (Shanghai, China).

**Quantitative PCR analysis**

Quantitative PCR was performed using the ABI PRISM 7300 (Applied Biosystems) with LightCycler FastStart DNA Master SYBR Green I (Roche) as the detection dye, according to the manufacturer's instructions. The primers used were listed in Table 2.

**Table 2. Primers used for quantitative PCR**

| Primer | Primer sequence (5’-3’) | Product size (bp) | Reference |
|--------|-------------------------|-------------------|-----------|
| WP0273(C6/36) | F-GAGGCGAAGTAAAGTACAAA | 453 | |
| act5C | F-ATCGTACGAACTCCGAGTG | 188 | |
| β-actin | F-CCACCATGACCCAGGAATC | 186 | |

*Wolbachia* Primers used for detection of *Wolbachia* and WP0273(C6/36) (Applied Biosystems) with LightCycler FastStart DNA Master SYBR Green I (Roche) as the detection dye, according to the manufacturer's instructions. The primers used were listed in Table 2.
Table 2. β-actin was used as the internal control. The relative gene expression level was calculated from the threshold cycle (Ct) value of each reaction.

Protein extraction from cells and mosquitoes

The culture media was removed from the cell culture plate, washed twice with PBS, then 300 µL RIPA buffer, and 3 µL PMSF was added. Cells were harvested and then lysed. Collected cells were placed in a centrifuge tube on ice for 30 min and then in a water bath at 100 °C with loading buffer for 5 min. This solution was centrifuged at 12,000 rpm for 5 min and the supernatant then transferred into a new tube and stored at −20 °C.

Five non-blood-sucking female Anopheles sinensis mosquitoes were placed in a centrifuge tube to which was added 300 µL RIPA buffer and 3 µL PMSF. The mosquitoes were ground completely on ice and left to stand for 30 min. The supernatant was transferred into a new tube, placed in a water bath at 100 °C with loading buffer for 5 min, then centrifuged at 12,000 rpm for 5 min, and stored at −20 °C. These protein extracts were used to perform Western immunoblotting.

Phylogenetic Analysis

Phylogenetic analyses were conducted by the maximum-likelihood (ML), neighbor-joining (NJ), and minimum-evolution (ME) method by using MEGA 5.1.

Analysis and prediction of protein sequence and structure

The 3D structure of the protein was constructed using Phyre2 server (protein homology/analogy recognition engine). Protein function was predicted using the Phyre2 server SMART software (http://smart.embl-heidelberg.de/) and Myhits Motif Scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Professor Giulio Superti-Furga (Research Center for Molecular Medicine of the Austrian Academy of Sciences) for sharing of Protein G and Streptavidin-binding peptide. This work was supported by the National Institutes of Health of US (NIH) (Grant no. R01AI075746-05), the National Natural Science Foundation of China (Grant No. 30901244, 30972564, 81171900, 81000751, and 8110279), the National S and T Major Program (Grant No. 2012ZX10004-219, 2008ZX10004-010, 2012ZZ10004-220, and 2008ZX10004-011), Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No.2013234120007), National Science Foundation of Jiangsu Province (Grant No.8110279), and Priority Academic Program Development of Jiangsu Higher Education Institutions.
24. Zhou W, Rousset F, O’Neil S. Phylogeny and PCR-based classification of Wolbachia strains using wsp gene sequences. Proc Biol Sci 1998; 265:509-15; PMID:9569669; http://dx.doi.org/10.1098/rspb.1998.0324
25. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011; 28:2731-9; PMID:19574458; http://dx.doi.org/10.1093/molbev/msr121
26. Klasson L, Kambris Z, Cook PE, Walker T, Sinkins SP. Horizontal gene transfer between Wolbachia and the mosquito Aedes aegypti. BMC Genomics 2009; 10:33; PMID:19154594; http://dx.doi.org/10.1186/1471-2164-10-33
27. Report of the World Health Organization Technical Consultation on Prevention and Control of Iron Deficiency in Infants and Young Children in Malaria-Endemic Areas, Lyon, France, 12-14 June 2006. Food Nutr Bull 2007; 28(Suppl):S489-631; PMID:18368738
28. Woolfit M, Iturbe-Ormaetxe I, McGraw EA, O’Neill SL. An ancient horizontal gene transfer between mosquito and the endosymbiotic bacterium Wolbachia pipientis. Mol Biol Evol 2009; 26:367-74; PMID:18988686; http://dx.doi.org/10.1093/molbev/msn253
29. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature 2000; 405:299-304; PMID:10830951; http://dx.doi.org/10.1038/35012500
30. Foord R, Taylor IA, Sedgwick SG, Smerdon SJ. X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of the ankyrin fold and has implications for Swi6 function. Nat Struct Biol 1995; 6:157-65; PMID:10044928; http://dx.doi.org/10.1038/35845
31. Lubman OY, Kopan R, Waksman G, Korolev S. The crystal structure of a partial mouse Notch-1 ankyrin domain: repeats 4 through 7 preserve an ankyrin fold. Protein Sci 2005; 14:1274-81; PMID:15802643; http://dx.doi.org/10.1111/j.0965-5433.2005.02496.x
32. Breden L, Nasmyth K. Similarity between cell-cycle genes of budding yeast and fusion yeast and the Notch gene of Drosophila. Nature 1987; 329:651-4; PMID:2821408; http://dx.doi.org/10.1038/329651a0
33. Serhus LR, Casper-Lindley C, Landmann F, Sullivan W. The genetics and cell biology of Wolbachia-host interactions. Annu Rev Genet 2008; 42:683-707; PMID:18713031; http://dx.doi.org/10.1146/annurev.genet.41.110306.130354
34. Smillie CS, Smith MB, Friedman J, Cordero OX, David LA, Alm EJ. Ecology drives a global network of gene exchange connecting the human microbiome. Nature 2011; 480:241-4; PMID:22037308; http://dx.doi.org/10.1038/nature10571
35. Lester CH, Fimodt-Møller N, Sørensen TL, Monnet DL, Hammerum AM. In vivo transfer of the vanA resistance gene from an Enterococcus faecium isolate of animal origin to an E. faecium isolate of human origin in the intestines of human volunteers. Antimicrob Agents Chemother 2006; 50:596-9; PMID:16436715; http://dx.doi.org/10.1128/AAC.50.2.596-599.2006
36. Hacker J, Blum-Oehler G, Mühldorfer I, Tschäpe H. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol Microbiol 1997; 23:1089-97; PMID:9106201; http://dx.doi.org/10.1046/j.1365-2958.1997.1597132.x
37. Ritter A, Blum G, Emödy L, Kerenyi M, Böck A, Neuhierl B, Scheutz F, Hacker J. tRNA genes and pathogenicity islands: influence on virulence and metabolic properties of uropathogenic Escherichia coli. Mol Microbiol 1995; 17:109-21; PMID:7476197; http://dx.doi.org/10.1111/j.1365-2958.1995.mmi_17010109.x
38. Osborne SE, Leong YS, O’Neill SL, Johnson KN. Variation in antiviral protection mediated by different Wolbachia strains in Drosophila simulans. PLoS Pathog 2009; 5:e1000656; PMID:19911047; http://dx.doi.org/10.1371/journal.ppat.1000656
39. Hedges LM, Brownlie JC, O’Neill SL, Johnson KN. Wolbachia and virus protection in insects. Science 2008; 322:702; PMID:18974344; http://dx.doi.org/10.1126/science.1162418
40. Bian G, Xu Y, Lu P, Xie Y, Xi Z. The endosymbiotic bacterium Wolbachia induces resistance to dengue virus in Aedes aegypti. PLoS Pathog 2010; 6:e1000833; PMID:20368968; http://dx.doi.org/10.1371/journal.ppat.1000833
41. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 2009; 4:363-71; PMID:19247286; http://dx.doi.org/10.1038/nprot.2009.2
42. Schulz J, Copley RR, Doerks T, Ponting CP, Bork P. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res 2000; 28:251-4; PMID:10952234; http://dx.doi.org/10.1093/nar/28.1.231
43. Pagni M, Ioannidis V, Cerutti L, Zahn-Zabal M, Jongeneel CV, Hau J, Marrin O, Kuznetsov D, Falquet L. MyHits: improvements to an interactive resource for analyzing protein sequences. Nucleic Acids Res 2007; 35(Web Server issue):W433-7; PMID:17545200; http://dx.doi.org/10.1093/nar/gkm352