DOWNLOAD PDF ADAPTIVE MOLECULAR EVOLUTION

Chapter 1 : PPT â€“ Adaptive Molecular Evolution PowerPoint presentation | free to download - id: NjI1M

It is thus plausible that the experimental evolution is more adaptive than natural evolution (in terms of the proportion of changes that are adaptive), and there are several measures of molecular evolution that can be observed to assess the degree of adaptive molecular evolution here.

Received Sep 21; Accepted Apr This article has been cited by other articles in PMC. Abstract Limestone Karst areas possess high levels of biodiversity and endemism. Primulina is a typical component of Karst endemic floras. The high species richness and wide distribution in various Karst microenvironments make the genus an idea model for studying speciation and local adaptation. Then, we used maximum-likelihood approaches to explore molecular evolution of PHYE in this Karst cave plant. The results showed that PHYE was dominated by purifying selection in both data sets, and two sites were identified as potentially under positive selection. These results suggest that potential positive selection in PHYE might have played an important role in the adaption of Primulina to heterogeneous light environments in Karst regions, and different species lineages might have been subjected to different selective pressures. Introduction Light is not only the source of energy, but also a very important environmental factor for plant growth and survival. As sessile organisms, plants have evolved sophisticated photosensory systems to respond appropriately to their light environments. Phytochromes are specialized photosensors that perceive and interpret light signals from the environment to regulate plant growth and development throughout the whole life cycle [1]. Recent studies have revealed that phytochromes play an important role in modulating both biotic and abiotic stress [2]. In angiosperms, the phytochrome apoprotein genes have been classified into four or five gene subfamilies based on sequence similarity to the five phytochrome genes of Arabidopsis: The crystal structure of Arabidopsis PHYB was resolved recently [7], and it provides a helpful scaffold for understanding the signaling and functional mechanism of plant phytochromes. Previous studies have demonstrated that the evolutionary adaptation of phytochromes is associated with polymorphisms in the phytochrome genes regulating ecologically important traits [8 â€” 12]. A phylogenetic analysis suggested that positive selection in PHYA has played a major role in the adaptive evolution of early angiosperms [13]. Molecular evolutionary analysis of the phytochrome genes in Sorghum [14] and Brassicaceae [15 , 16] have shown that the evolution of phytochromes is mainly constrained by purifying selection. More recently, population genetic studies of alpine plants have revealed positive selection in PHYE, suggesting its involvement in adaptation to local environments [16 , 17]. PHYE is broadly distributed in flowering plants, expressed throughout the course of development and present in various organs [18]. At cooler temperatures, PHYE plays a prominent role in regulation of germination [19] and flowering [20 , 21]. PHYB is a principal mediator that responses to R and FR [5 , 6], and evolves under constraints by purifying selection in Arabidopsis [15]. Therefore, PHYE should be a promising candidate gene for exploring local adaptation to different light intensity environments. Primulina Gesneriaceae , a typical cave plant, is a monophyletic genus comprising more than species of perennials that are widely distributed throughout the Karst regions of southern China and the adjacent countries of Southeast Asia. As sessile organisms, the heterogeneous light environments exert a selection pressure on Primulina to survival in Karst habitats. Identifying genes targeted by natural selection can greatly improve our knowledge of the role of adaptation in species evolution. Despite recent advances in our understanding of ecophysiological adaptation to the Karst environment [30 , 31], the molecular-genetic mechanisms by which Primulina adapts to heterogeneous light conditions have never been explored. Although phytochromes play an important role in plant life cycle, little is known about the composition and evolution of the phytochrome gene family in Primulina to date. In this study, we chose PHYE as the target to explore whether the phytochrome involved in local adaptation of Primulina to the diverse light environments in Karst areas. For this purpose, we obtained 10 full-length sequences and 74 partial sequences of PHYE from Primulina, which were sampled from a wide geographic range of the genus. We used molecular evolutionary approaches to test whether positive selection or selective constraints arisen on this gene. Materials and Methods Ethics Statement P. The leaf samples of this species were collected with permission from the

greenhouse of South China Botanical Garden. All other species are not recognized as the endangered or protected species at the moment, the leaf samples used in this study were collected from open areas, and the location is not privately owned or protected in any way, so no specific permits were required for the sampling. Plant materials and amplification of PHYE genes The plant materials used in this study were collected from fields throughout the geographic range of Primulina in China, as specified in S1 Table. These species are widely distributed across the phylogeny of the genus. One individual of each species was used, and a total of 74 Primulina species and two outgroups Didymocarpus hancei and Petrocodon dealbatus were included. The queried sequences were checked carefully by eyes and made certain that there were no any ambiguity characters, no frame-shift mutations or premature stop codons. The stop codons were excluded in the following analysis. Thus, this study mainly focused on the conserved core signaling domains of the phytochrome, i. This led to the two data sets analyzed in this study: The specific primers used in the amplification of the core signaling domain of the PHYE gene were designed according to the alignment of full length sequences using Primer Premier 5. Reaction conditions were as follows: Once purified, the PCR products were directly sequenced in both directions using the same primers as in amplification. As recombination can mislead phylogenetic and positive selection analyses [35], we used the genetic algorithm for recombination detection GARD [36] method implemented in the Datamonkey web-server www. Kishino-Hasegawa tests [38] were used to test statistical differences when potential breakpoints were detected. The phylogenetic relationships were reconstructed by MrBayes v3. For the Bayesian analysis of the partial sequences, Didymocarpus hancei and Petrocodon dealbatus were set as outgroups according to our previous phylogenetic analysis of the genus[30], and the Markov Chain Monte Carlo MCMC search was run 8,, generations and sampled every generations. For the Bayesian tree constructed for the full-length sequences, the tree was rooted at P. The MCMC was run for , generations and sampled every generations, and the first trees were discarded as burn-in.

Chapter 2 : Molecular evolution - Wikipedia

These results suggest adaptive convergent molecular evolution. Site is located in the receptor-binding region and is near antigenic site A, as defined by Wiley et al. (34). Thus, the observed AV mutation might modulate receptor affinity and contribute to immune escape (Figure 6), as observed in influenza A(H7N1) and A(H7N7) viruses.

Genome size[edit] Genome size is influenced by the amount of repetitive DNA as well as number of genes in an organism. Explanations for the so-called paradox are two-fold. First, repetitive genetic elements can comprise large portions of the genome for many organisms, thereby inflating DNA content of the haploid genome. Secondly, the number of genes is not necessarily indicative of the number of developmental stages or tissue types in an organism. An organism with few developmental stages or tissue types may have large numbers of genes that influence non-developmental phenotypes, inflating gene content relative to developmental gene families. Neutral explanations for genome size suggest that when population sizes are small, many mutations become nearly neutral. There is little evidence to suggest that genome size is under strong widespread selection in multicellular eukaryotes. Genome size, independent of gene content, correlates poorly with most physiological traits and many eukaryotes, including mammals, harbor very large amounts of repetitive DNA. However, birds likely have experienced strong selection for reduced genome size, in response to changing energetic needs for flight. Birds, unlike humans, produce nucleated red blood cells, and larger nuclei lead to lower levels of oxygen transport. Bird metabolism is far higher than that of mammals, due largely to flight, and oxygen needs are high. Hence, most birds have small, compact genomes with few repetitive elements. Indirect evidence suggests that non-avian theropod dinosaur ancestors of modern birds [6] also had reduced genome sizes, consistent with endothermy and high energetic needs for running speed. Many bacteria have also experienced selection for small genome size, as time of replication and energy consumption are so tightly correlated with fitness. Repetitive elements[edit] Transposable elements are self-replicating, selfish genetic elements which are capable of proliferating within host genomes. Many transposable elements are related to viruses, and share several proteins in common. The ant Myrmecia pilosula has only a single pair of chromosomes [7] whereas the Adders-tongue fern Ophioglossum reticulatum has up to chromosomes. Reduced linkage through creation of additional chromosomes should effectively increase the efficiency of selection. Changes in chromosome number can play a key role in speciation, as differing chromosome numbers can serve as a barrier to reproduction in hybrids. Human chromosome 2 was created from a fusion of two chimpanzee chromosomes and still contains central telomeres as well as a vestigial second centromere. Polyploidy, especially allopolyploidy, which occurs often in plants, can also result in reproductive incompatibilities with parental species. Some organisms, such as most bacteria, Drosophila, and Arabidopsis have particularly compact genomes with little repetitive content or non-coding DNA. Other organisms, like mammals or maize, have large amounts of repetitive DNA, long introns, and substantial spacing between different genes. The content and distribution of genes within the genome can influence the rate at which certain types of mutations occur and can influence the subsequent evolution of different species. Genes with longer introns are more likely to recombine due to increased physical distance over the coding sequence. As such, long introns may facilitate ectopic recombination , and result in higher rates of new gene formation. Organelles[edit] In addition to the nuclear genome, endosymbiont organelles contain their own genetic material typically as circular plasmids. Mitochondrial and chloroplast DNA varies across taxa, but membrane-bound proteins, especially electron transport chain constituents are most often encoded in the organelle. Chloroplasts and mitochondria are maternally inherited in most species, as the organelles must pass through the egg. In a rare departure, some species of mussels are known to inherit mitochondria from father to son. Origins of new genes[edit] New genes arise from several different genetic mechanisms including gene duplication, de novo origination, retrotransposition, chimeric gene formation, recruitment of non-coding sequence, and gene truncation. Gene duplication initially leads to redundancy. However, duplicated gene sequences can mutate to develop new functions or specialize so that the new gene performs a subset of the original ancestral functions. In addition to duplicating whole genes, sometimes only a domain or part of a

protein is duplicated so that the resulting gene is an elongated version of the parental gene. Retrogenes often insert into new genomic locations, and often develop new expression patterns and functions. Chimeric genes form when duplication, deletion, or incomplete retrotransposition combine portions of two different coding sequences to produce a novel gene sequence. Chimeras often cause regulatory changes and can shuffle protein domains to produce novel adaptive functions. Novel genes can also arise from previously non-coding DNA. De novo evolution of genes can also be simulated in the laboratory. More specifically, they selected sequences from a library that could complement a gene deletion in E. The deleted gene encodes ferric enterobactin esterase Fes , which releases iron from an iron chelator , enterobactin. While Fes is a amino acid protein, the newly selected gene was only amino acids in length and unrelated in sequence to Fes. The chemical kinetic elucidation of the detailed mechanism of replication [ref, ref] meant that this type of system was the first molecular evolution system that could be fully characterised on the basis of physical chemical kinetics, later allowing the first models of the genotype to phenotype map based on sequence dependent RNA folding and refolding to be produced [ref, ref]. Experiments with in vitro RNA quasi species included the characterisation of the error threshold for information in molecular evolution [ref], the discovery of de novo evolution [ref] leading to diverse replicating RNA species and the discovery of spatial travelling waves as ideal molecular evolution reactors [ref, ref]. Later experiments employed novel combinations of enzymes to elucidate novel aspects of interacting molecular evolution involving population dependent fitness, including work with artificially designed molecular predator prey and cooperative systems of multiple RNA and DNA [ref, ref]. Special evolution reactors were designed for these studies, starting with serial transfer machines, flow reactors such as cell-stat machines, capillary reactors, and microreactors including line flow reactors and gel slice reactors. These studies were accompanied by theoretical developments and simulations involving RNA folding and replication kinetics that elucidated the importance of the correlation structure between distance in sequence space and fitness changes [ref], including the role of neutral networks and structural ensembles in evolutionary optimisation. Molecular systematics and Phylogenetics Molecular systematics is the product of the traditional fields of systematics and molecular genetics. Molecular systematics has been made possible by the availability of techniques for DNA sequencing , which allow the determination of the exact sequence of nucleotides or bases in either DNA or RNA. At present it is still a long and expensive process to sequence the entire genome of an organism, and this has been done for only a few species. However, it is quite feasible to determine the sequence of a defined area of a particular chromosome. Typical molecular systematic analyses require the sequencing of around base pairs. The driving forces of evolution[edit] Main articles: Neutral theory of molecular evolution , Modern synthesis 20th century , and Mutationism Depending on the relative importance assigned to the various forces of evolution, three perspectives provide evolutionary explanations for molecular evolution. While acknowledging that many mutations are neutral, selectionists attribute changes in the frequencies of neutral alleles to linkage disequilibrium with other loci that are under selection, rather than to random genetic drift. The Neutral theory of molecular evolution proposes that most mutations in DNA are at locations not important to function or fitness. These neutral changes drift towards fixation within a population. Positive changes will be very rare, and so will not greatly contribute to DNA polymorphisms. The nearly neutral theory expanded the neutralist perspective, suggesting that several mutations are nearly neutral, which means both random drift and natural selection is relevant to their dynamics. He proposed that the variation in GC content was not the result of positive selection, but a consequence of the GC mutational pressure. It demonstrates how proteins evolve, keeping some regions conserved while others change dramatically. Evolution of proteins is studied by comparing the sequences and structures of proteins from many organisms representing distinct evolutionary clades. More specifically, homologous proteins that exist in two distinct species are called as orthologs. Whereas, homologous proteins encoded by the genome of a single species are called paralogs. The phylogenetic relationships of proteins are examined by multiple sequence comparisons. Phylogenetic trees of proteins can be established by the comparison of sequence identities among protoeins. Such phylogenetic trees have established that the sequence similarities among proteins reflect closely the evolutionary relationships among organisms. Through quantitative analysis and experimentation, scientists have strived to understand the rate and causes of protein evolution. Using the

amino acid sequences of hemoglobin and cytochrome c from multiple species, scientists were able to derive estimations of protein evolution rates. What they found was that the rates were not the same among proteins. Not all regions within a protein mutate at the same rate; functionally important areas mutate more slowly and amino acid substitutions involving similar amino acids occurs more often than dissimilar substitutions. Several species including humans, fruit flies, and mice have similar levels of protein polymorphism. Amino acid sequences and nucleic acid sequences do not mutate at the same rate. Due to the degenerate nature of DNA, bases can change without affecting the amino acid sequence. For example, there are six codons that code for leucine. Thus, despite the difference in mutation rates, it is essential to incorporate nucleic acid evolution into the discussion of protein evolution. At the end of the s, two groups of scientistsâ€"Kimura and King and Jukes â€"independently proposed that a majority of the evolutionary changes observed in proteins were neutral. These discordances can be categorized as two types: Neutral evolution possibly could explain the incongruences in some cases. Other journals dedicated to molecular evolution include Journal of Molecular Evolution and Molecular Phylogenetics and Evolution. Research in molecular evolution is also published in journals of genetics , molecular biology , genomics , systematics , and evolutionary biology.

Chapter 3 : Adaptive Molecular Evolution of PHYE in Primulina, a Karst Cave Plant

Evidence for adaptive evolution has been found previously among HIV sequences from intra- and interpatient studies (4, 29, 30, 38, 40). Early studies involved the pairwise comparison of synonymous (silent, d S) and nonsynonymous (amino acid changing, d N) substitutions between protein-coding DNA sequences.

Advanced Search Abstract Cichlid fish inhabit a diverse range of environments that vary in the spectral content of light available for vision. These differences should result in adaptive selective pressure on the genes involved in visual sensitivity, the opsin genes. This study examines the evidence for differential adaptive molecular evolution in East African cichlid opsin genes due to gross differences in environmental light conditions. First, we characterize the selective regime experienced by cichlid opsin genes using a likelihood ratio test format, comparing likelihood models with different constraints on the relative rates of amino acid substitution, across sites. Second, we compare turbid and clear lineages to determine if there is evidence of differences in relative rates of substitution. Third, we present evidence of functional diversification and its relationship to the photic environment among cichlid opsin genes. We report statistical evidence of positive selection in all cichlid opsin genes, except short wavelengthâ€”sensitive 1 and short wavelengthâ€”sensitive 2b. In all genes predicted to be under positive selection, except short wavelengthâ€”sensitive 2a, we find differences in selective pressure between turbid and clear lineages. Potential spectral tuning sites are variable among all cichlid opsin genes; however, patterns of substitution consistent with photic environmentâ€—driven evolution of opsin genes are observed only for short wavelengthâ€—sensitive 1 opsin genes. This study identifies a number of promising candidate-tuning sites for future study by site-directed mutagenesis. This work also begins to demonstrate the molecular evolutionary dynamics of cichlid visual sensitivity and its relationship to the photic environment. Some of the most dramatic examples have been found in fish rod photoreceptors. Deep-sea fish rod spectral sensitivities are shortwave shifted, relative to shallow-dwelling species, to match the ambient spectra of the deep-sea environment Partridge, Archer, and Lythgoe , Crescitelli Adaptation to depth has also been observed in the rods of freshwater teleosts. Muntz compared a shallow and deeper living pair of closely related cichlid species of the genus Lethrinops. Again, the deepwater species had shortwave-shifted rod sensitivity. There are also several examples of differences in cone spectral sensitivity associated with disparities in photic environment. Both deep-dwelling Lake Baikal cottoids and coelacanths show a marked shortwave shift in cone spectral sensitivities Bowmaker et al. Lutjanid fishes, of the Great Barrier Reef, demonstrate the interaction between water clarity and cone spectral sensitivity, with fish in clearer habitats having shortwave-shifted visual sensitivities Lythgoe et al. Visual pigments determine spectral sensitivity and are spectrally distinct photosensory molecules in the outer segments of retinal photoreceptor cells. Visual pigments are composed of a vitamin Aâ€—derived chromophore bound to an opsin protein. Photoisomerization of the chromophore initiates the transductional cascade culminating in a neural response. Interactions between the chromophore and the amino acid residues of the opsin protein determine the absorbance properties of a visual pigment Sakmar, Franke, and Khorana ; Zhukovsky and Oprian ; Nathans a , b ; Sakmar et al. Spectral sensitivity of cichlid fishes can be tuned using four nonexclusive mechanisms. First, the lenses of some cichlids contain inert short wavelengthâ€—absorbing carotenoid pigments Thorpe, Douglas, and Truscott The presence of ocular pigments seems to be independent of photic environment, with nonpigmented species occurring in fish of both turbid and clear habitats Thorpe, Douglas, and Truscott Second, Carleton and Kocher have shown that cichlids use differential cone opsin expression to modulate visual sensitivity. Cichlids have six opsin genes five cone opsins and one rod opsin: Individual species express varying subsets of the five cone opsin genes. Third, chromophore usage can vary among cichlids vitamin A1 or A2 derived. Visual pigments based on a vitamin A2â€—derived chromophore have long wavelengthâ€—shifted absorbance maxima, relative to those based on a vitamin A1â€—derived chromophore Partridge and Cummings Fish that inhabit turbid environments more commonly use A2 or A1-A2 mixtures Bowmaker , although chromophore thermal stability may also shape usage Partridge and Cummings Finally, amino acid substitutions in the opsin protein can alter visual sensitivity summarized in Yokoyama and

Takahashi and Ebrey The effects of individual substitutions are highly variable, ranging from 0 to 75 nm. Further, the effects of individual substitutions depend upon the amino acid background of the opsin protein. There is mounting evidence for molecular adaptation to photic environment, via amino acid substitutions in opsin proteins, in East African cichlids. Recently, Sugawara, Terai, and Okada found evidence of functional divergence in Tanganyikan cichlid Rh1 opsin genes. They found an AS substitution in several Tanganyikan lineages bovine rhodopsin numbering will be used exclusively in this paper. Interestingly, all three species with the AS substitution are deepwater species, further supporting the relationship of spectral sensitivity to depth. Further, Terai et al. The present study looks for evidence of adaptive molecular evolution among closely related cichlid species that inhabit dramatically different photic environments. We then test for differences in relative rates of evolution between turbid and clear-water lineages. Finally, we identify amino acid substitutions that are likely to be involved in the adaptive-functional differentiation of cichlid opsins. Cichlids endemic to clear lakes Tanganyika and Malawi are contrasted against cichlids endemic to more turbid environments in Lake Victoria and the Nile River. Due to geologic and climatic conditions, Lake Tanganyika and Lake Malawi are among the clearest freshwater systems in the world Muntz and provide a stark contrast to the generally more turbid environments of Lake Victoria Seehausen, van Alphen, and Witte and the Nile River. Because turbidity directly limits the transmission of the shortest wavelengths of the visible spectrum, the spectrum of ambient light is shifted toward the long-wavelength region. Thus, the spectral breadth of light available for vision is restricted in turbid habitats. PAML has been used to detect positive selection among fertilization proteins Civetta ; Swanson, Nielson, and Yang , lysozymes Yang ; Yang and Nielsen , tumor suppressors Yang and Nielsen , dopamine receptors Ding et al. Opsin-coding sequences were obtained for 17 East African cichlid species table 1. Cone opsin gene sequences from five species and rod opsin gene sequences from two of those species were obtained from GenBank Carleton, Harosi, and Kocher ; Carleton and Kocher ; K. All other rod opsin sequences and the cone opsin sequences from the remaining 12 species are new additions to the sequence database. Retinal tissue was used to extract opsin messenger RNA, whenever possible. Genomic DNA was extracted as well and used to determine opsin-coding sequences when necessary. MJ Research, Waltham, Mass. Carleton and Kocher Sequences were aligned using Sequencher 4. Whenever possible the complete opsin-coding sequence was used. The regions used in the analysis always included all transmembrane TM and inter-TM regions, which control the spectral absorbance of visual pigments.

Chapter 4 : Adaptive evolution in the human genome - Wikipedia

Highlight and copy the desired format. Emerging Infectious Diseases, 24 10 , Abstract The substantial increase in prevalence and emergence of antigenically divergent or highly pathogenic influenza A H7N9 viruses during â€”17 raises concerns about the epizootic potential of these viruses. We investigated the evolution and adaptation of H7N9 viruses by analyzing available data and newly generated virus sequences isolated in Guangdong Province, China, during â€” Phylogenetic analyses showed that circulating H7N9 viruses belong to distinct lineages with differing spatial distributions. Hemagglutination inhibition assays performed on serum samples from patients infected with these viruses identified 3 antigenic clusters for 16 strains of different virus lineages. We used ancestral sequence reconstruction to identify parallel amino acid changes on multiple separate lineages. We inferred that mutations in hemagglutinin occur primarily at sites involved in receptor recognition or antigenicity. Our results indicate that highly pathogenic strains likely emerged from viruses circulating in eastern Guangdong Province during March and are associated with a high rate of adaptive molecular evolution. Since its first detection in March , avian influenza A H7N9 virus has caused 1, human infections that, as of November 30, , had resulted in deaths. Recurrent waves of human cases have been reported in 27 provinces in China, indicating sustained transmission of H7N9 viruses 1. Moreover, since its emergence, H7N9 virus has reassorted with influenza A H9N2 viruses that co-circulate in China, resulting in an increasingly diverse array of virus genomes 2 â€” 4. The fifth influenza epidemic wave â€”17 was marked by a notable increase in the number of human cases during September â€”May , making it the largest outbreak of influenza A H7N9 since Moreover, geographic distribution of human cases suggests that H7N9 virus is now more widespread and that residences of patients have shifted gradually from urban to semiurban and rural areas 1 , 5 â€” 7. These epidemiologic observations have raised public health concerns. Previous molecular surveillance studies suggested that H7N9 virus lineages originate in 2 densely populated areas, the Yangtze River Delta region in eastern China and the Pearl River Delta region in central Guangdong Province 8. Preliminary epidemiologic data suggested that most human infections in the current fifth epidemic wave were caused by viruses from the Yangtze River Delta region 5 previously named lineage C viruses 3. These viruses, in contrast to viruses from the Pearl River Delta region previously named lineage B viruses 3 , appear to exhibit reduced cross-reactivity with existing candidate vaccine virus strains 9. Furthermore, a subset of lineage C isolates has also acquired a highly pathogenic HP phenotype 5 , 10 , These observations suggest that the increased epidemic activity of H7N9 viruses in China might be driven, at least in part, by ongoing virus evolution and adaptation. Decreased cross-reactivity and increased pathogenicity of some H7N9 viruses was discovered only recently 9 , and the genetic diversity and evolution of the current fifth epidemic wave of these viruses are not yet well understood. Information necessary to clarify this situation includes geographic distribution of currently circulating H7N9 virus lineages, origin and genetic composition of newly emerged HP H7N9 viruses, and evolutionary and structural characterization of mutations associated with the fifth epidemic wave of these viruses. We report 47 hemaglutinnin HA and 43 neuraminidase NA gene sequences of human-derived and poultry-derived H7N9 viruses that were isolated during â€” in Guangdong Province, China. We conducted structural and evolutionary analyses of these strains and characterized the evolution and emergence of currently circulating H7N9 viruses in China. Written consent was obtained from patients or their guardian s when samples were collected. Patients were informed about the study before providing written consent, and data were anonymized for analysis. Sample Collection Samples from persons with suspected cases of influenza A H7N9 were initially tested for avian influenza A virus in provincial clinics in Guangdong Province. Specimens with positive results were subsequently analyzed 12 , For poultry-related samples, we obtained samples from locations where poultry were housed and processed e. Respiratory specimens were collected from persons with suspected cases of influenza A H7N9 by the Ministry of Health of China. EPIâ€”77, â€”6, â€”, , , â€”42, â€”60, and â€” These new H7N9 sequences were combined with all

available H7N9 gene sequences whose sampling dates and locations were known. Two gene sequence datasets were generated: We constructed multiple sequence alignments by using ClustalW 14 and edited these sequences manually by using AliView We computed 4 independent Markov Chain Monte Carlo runs of 1. We computed maximum clade credibility trees for each dataset by using TreeAnnotator Because sporadic human cases detected in Malaysia and Taiwan were believed to have originated in China, we used available epidemiologic information to assign their location to the most likely source in China. Hong Kong and central Guangdong Province were treated as a single location because of their proximity to each other. To estimate directionality of virus lineage movement, we used asymmetric continuous-time Markov chain phylogeographic model 19 and a Bayesian stochastic search variable selection procedure We estimated maximum posterior probability amino acid sequences for each internal node by using BEAST with a Jonesâ€"Taylorâ€"Thornton amino acid substitution model 20 , gamma-distributed among-site rate heterogeneity 21 , and a strict molecular clock model. To infer amino acid substitutions along the trunk branches of the H7N9 phylogeny, we mapped amino acid changes onto internal branches by using a Java script available on request. We performed residue mapping onto the H7 and N9 structures by using PyMol These methods included single-likelihood ancestor counting 27 , fixed effects likelihood 27 , mixed effects model of evolution 28 , and the fast unconstrained Bayesian approximation approach We used a consensus of H7N9 first-wave sequences as an outgroup to estimate derived and ancestral mutational site frequencies in each subsequent wave. We calculated the number of adaptive substitutions from the number of synonymous and nonsynonymous sites in each category and assessed statistical uncertainty by using a bootstrap approach 1, replicates 30 , Serologic Analysis We obtained serum samples from 4 patients with influenza A H7N9 2â€"3 weeks after clinical symptoms were observed. We performed hemagglutination inhibition assays by using different lineages of H7N9 viruses as antigens online Technical Appendix. We calculated serum titer for each H7N9 strain as the highest reciprocal serum dilution providing complete hemagglutination inhibition. Regression of root-tip divergence estimated from hemagglutinin gene of influenza A H7N9 viruses, China. Arrow indicates the time of the most recent common ancestor of the epidemic lineage. Figure 2 Figure 2. Genetic evolution and spatial spread of epidemic lineage of influenza A H7N9 viruses, China, â€" Bayesian maximum clade credibility tree of the hemagglutinin gene is shown. Black bars to the right of During â€", the influenza A H7N9 virus epidemic lineage was geographically structured and classified into 3 major lineages, A, B, and C, in accordance with the lineage naming scheme used in a previous study 3. H7N9 virus has evolved in a clock-like manner i. Figure 3 Figure 3. Geographic location and lineage classification of influenza A H7N9 human viruses, China. Values in parentheses indicate number of sequenced viruses from each region. Pie charts indicate approximate percentages of each virus Different H7N9 virus lineages are associated with different epidemiologic patterns Figures 2 , 3. In addition, lineage B viruses isolated from the fourth and fifth influenza waves were almost exclusively restricted to central rather than eastern Guangdong Province Figures 2 , 3. In contrast, viruses in eastern China, composed of 2 lineages A and C have been exported to and become dominant in multiple regions as the epidemic has progressed 3. These findings indicate a comparatively broader geographic dissemination Figure 3 ; Technical Appendix Figure 1. Figure 4 Figure 4. Reconstruction of amino acid changes along trunk of lineage B phylogenies of influenza A H7N9 viruses, China. Maximum clade credibility tree of hemagglutinin gene sequences from lineage B is shown. The new isolates from eastern Guangdong Province, combined with isolates from eastern China 1 , 5 , suggest that recent H7N9 virus activity is driven primarily by lineage C viruses Figure 2. For lineage C, we observed 2 clades Figure 4. The larger of these clades C1 circulates mainly in central and eastern China, and the smaller clade C2 is found predominantly in eastern Guangdong Province. Clade C2 also includes recently identified HP viruses Figures 1 , 2 , 4. To investigate these HP viruses, we undertook retrospective screening of poultry-related samples collected in Guangdong Province during January â€"February and identified 7 HP influenza virus isolates that belong to the HP cluster within C2 Figure 2. Our analyses indicated that the HP clade likely emerged from clade C2 viruses circulating in eastern Guangdong Province in Reconstruction of amino acid changes along trunk of lineage C phylogenies of influenza A H7N9 viruses, China. Maximum clade credibility tree of hemagglutinin gene sequences from lineage C is shown. We then investigated whether

the increasing prevalence of lineage C viruses might be associated with virus adaptation. We combined ancestral sequence reconstruction of lineage B and C HA gene sequences Figures 4 , 5 by mapping residues that have undergone changes onto the crystal structure of the trimeric hemagglutinin. Our analysis identified several notable amino acid substitutions that occurred along the internal branches of lineage C viruses Figure 4. Figure 6 Figure 6. Crystal structure of the homotrimeric H7 hemagglutinin bound to a human receptor analog We found by evolutionary analysis that several HA sites acquired amino acid mutations independently in different phylogenetic clades. The observation of parallel amino acid changes along those particular lineages Technical Appendix Tables 2, 3 that have persisted until the fifth influenza epidemic wave i. One subclade of lineage B viruses appears to have acquired mutations AV and SN within the last 12 months Figure 5. Therefore, we suggest that this subclade should be closely monitored in the future. This internal branch represents a period of approximately 1 year Figure 4. Although all of these changes appeared in residues with partial or full solvent accessibility, mutations KE, LQ, and IV are particularly noteworthy because they have arisen at or near known antigenic, receptor-binding, and proteolytic cleavage sites, respectively Figure 6. We also investigated whether amino acid changes in the HA gene during emergence of influenza A H7N9 virus have been driven by adaptive evolution similar to that observed for seasonal human influenza We found evidence for adaptive evolution in HA genes of B and C virus lineages. We estimated that lineage B adapted at a rate of 0. These results indicate molecular adaptation across the whole H7N9 lineage and suggest that adaptation is faster in the 2 C clades than in the A and B lineages. Previous estimates of rates of adaptive substitution were 1. In this context, the rate of adaptive evolution observed for lineage C here is notable and raises concern for ongoing evolution of these viruses. Antigenic Properties We collected serum samples from 4 patients infected with H7N9 virus during and Table. For patients 3 and 4, the corresponding virus strains were isolated and sequenced. Phylogenetic analysis indicated that patient 3 was infected with clade C1 virus and that patient 4 was infected with HP virus. Hemagglutination inhibition results suggested the presence of 3 antigenic clusters among the 16 H7N9 virus strains selected. Serum samples from patients 1, 2, and 3 showed similar patterns, reacting robustly to clade C1 viruses and moderately to clade C2 and lineage B viruses but poorly to HP viruses. Discussion Our results show that H7N9 viruses of lineage C, which were prevalent in the recent fifth influenza epidemic wave in China, comprise 2 geographically distinct clades C1 and C2 that have undergone adaptive evolution. Our ancestral state reconstruction analysis provides crucial evidence that 2 successful lineages of H7N9 viruses lineages B and C have experienced multiple parallel amino acid changes Figures 4 , 5 , suggesting the possible action of convergent molecular evolution.

Chapter 5 : CiteSeerX â€" Citation Query Adaptive Molecular Evolution

This study examines the evidence for adaptive molecular evolution in cichlid opsin genes in response to gross differences in environmental light conditions. One might be tempted to conclude that the cichlid opsin gene with the least nucleotide divergence (SWS2a) has experienced the least divergent selection, relative to other opsin genes.

Advanced Search Background and Aims Limestone karst areas possess high floral diversity and endemism. The genus Primulina, which contributes to the unique calcicole flora, has high species richness and exhibit specific soil-based habitat associations that are mainly distributed on calcareous karst soils. The adaptive molecular evolutionary mechanism of the genus to karst calcium-rich environments is still not well understood. Methods Specific amplification and sequencing primers were designed and used to amplify the full-length coding sequences of TPC1 from cDNA of 76 Primulina species. The sequence alignment without recombination and the corresponding reconstructed phylogeny tree were used in molecular evolutionary analyses at the nucleic acid level and amino acid level, respectively. Finally, the identified sites under positive selection were labelled on the predicted secondary structure of TPC1. No signal of substitution saturation was detected in the sequences, while significant recombination breakpoints were detected. The molecular evolutionary analyses showed that TPC1 was dominated by purifying selection and the selective pressures were not significantly different among species lineages. However, significant signals of positive selection were detected at both TPC1 codon level and amino acid level, and five sites under positive selective pressure were identified by at least three different methods. Different species lineages suffered similar selective pressure associated with calcium in karst environments, and episodic diversifying selection at a few sites may play a major role in the molecular evolution of Primulina TPC1. As plants are sessile organisms during most of their life cycles, edaphic factors, such as physical and chemical characteristics, are thought to confer strong selective pressures on their local adaptation to different environments, but relatively little is known about which genes are involved in such adaptation Turner et al. Limestone karst topography is mainly composed of calcium carbonate sedimentary rock and has unique geological conditions and natural environments. Karst soils, which were formed from carbonate bedrock, are usually characterized by higher concentrations of calcium Ca and magnesium Mg , higher pH and lower water storage capacities than the surrounding non-karst soils Hao et al. Thus, the unique physical and chemical profiles of karst soils generate a range of microclimatic and edaphic habitats that are centres of high diversity and endemism. Plants on these limestone rock outcrops are usually adapted to severe summer drought, extreme temperature fluctuations and unusual soil chemistry because of proximity to bedrock. The underlying genetic basis of edaphic adaptation for karst endemics, however, remains poorly understood. Calcium is an essential macronutrient element in plants; it plays a crucial role in regulating nearly all aspects of plant growth and development Hepler, Calcium may traverse the roots either through the symplast pathway or through the apoplast pathway. In plant cells, calcium is mainly stored in the cell wall and vacuoles, and the endoplasmic reticulum is also necessary in regulating the balance of the intracellular calcium ion concentration. These channels are predominantly involved in signal transduction and play important roles in various crucial cellular responses, such as hormone activity, flowering, pathogen defence and salt stress Jammes et al. The slow-activating vacuolar SV channel is a voltage-dependent calcium channel with slow activation kinetics. The SV channel mainly localizes in vacuolar membranes and seems to be widespread in terrestrial plants, including ferns and liverworts Pottosin and Schonknecht, However, except in tobacco Nicotiana tabacum and mosses Physcomitrella patens , where there are two and nine copies of TPCs, respectively Verret et al. All known plant TPCs have similar transmembrane topology, comprising two fused domains, each with six transmembrane helices and one highly conserved pore loop P-loop. Previous studies mainly focused on characterizing the functions of TPC1 in model species Kadota et al. The genus Primulina is a typical limestone karst or karst cave plant. This genus is a monophyletic group comprising more than species widely distributed throughout the limestone karst areas in southern China. However, almost all Primulina species exhibit specific soil-based habitat associations, with the majority of species occurring in calcareous soils that originated from limestone bedrock Hao et al. For this

purpose, we obtained full-length coding sequences from 76 Primulina species, covering the main distribution areas and soil habitat types in China, and then used molecular evolutionary methods to test whether positive selection or selective constraints have arisen in the gene. As far as we know, this is the first study of the molecular evolution of plant TPC1 and the first attempt to explore the adaptive molecular evolutionary mechanism of Primulina to a karst high-calcium soil habitat. The closely related species Hemiboea henryi was used as the outgroup in this study. These 76 Primulina species represent nearly the whole of the distribution range of this genus and the typical soil environments in China Supplementary Data Fig. Specific information on the plant materials used in this study is given in Supplementary Data Table S1. TPC1 amplification and sequencing Transcriptome resources of 11 Primulina species were constructed in previous work Ai et al. Seven full-length sequences and four partial-length sequences were obtained from Primulina species. The length of the seven full-length sequences that were obtained varied from to bp according to the alignment, and we therefore designed another two pairs of internal sequencing primers to ensure the convenience of sequencing: The sequencing results were assembled using SeqMan v7. Sequence alignment The assembled sequences were firstly translated into amino acid sequences using MEGA v5. After checked carefully by eye, the amino acid alignment was converted into the corresponding codon-based nucleotide multiple sequence alignment using the program PAL2NAL http: Test for substitution saturation Substitution saturation will decrease phylogenetic information contained in sequences, and phylogeny constructed by sequences that have experienced full substitution saturation will not reflect phylogenetic relationships Xia et al. The entropy-based index of substitution saturation Iss was used together with plotting transition and transversion rates against a corrected genetic distance to evaluate whether the sequences showed signs of saturation. Test for recombination It is well known that recombination can affect phylogenetic reconstruction Posada and Crandall, and positive selection analysis Shriner et al. Phylogenetic reconstruction MrModeltest v2. The final phylogenetic tree was displayed with TreeView v1. Selection analysis at nucleic acid level Codon-based multiple sequence alignment of PrimulinaTPC1 and the Bayes phylogenetic tree were used in subsequent molecular evolution analysis.