

OPEN ACCESS

Full open access to this and thousands of other papers at <http://www.la-press.com>.

Phylogenetic Analysis of Kindlins Suggests Subfunctionalization of an Ancestral Unduplicated Kindlin into Three Paralogs in Vertebrates

Ammad Aslam Khan¹, Axel Janke², Takashi Shimokawa¹ and Hongquan Zhang^{1,3}

¹Karolinska Institutet, Center for Biosciences, Department of Biosciences and Nutrition, Novum, 141 83 Huddinge, Sweden.

²LOEWE—Biodiversity and Climate Research Center—BiK-F, Senckenberganlage 25, D-60325 Frankfurt am Main, Germany.

³Key laboratory of Carcinogenesis and Translational Research, the Ministry of Education and Peking University Institute of Systems Biomedicine and School of Basic Medical Sciences, Laboratory of Molecular Cell Biology and Tumor Biology, Beijing 100191, P.R. China. Corresponding author email: hammad.khan@ki.se; hongquan.zhang@bjmu.edu.cn

Abstract: Kindlin proteins represent a newly discovered family of evolutionarily conserved FERM domain-containing proteins. This family includes three highly conserved proteins: Kindlin-1, Kindlin-2 and Kindlin-3. All three Kindlin proteins are associated with focal adhesions and are involved in integrin activation. The FERM domain of each Kindlin is bipartite and plays a key role in integrin activation. We herein explore for the first time the evolutionary history of these proteins. The phylogeny of the Kindlins suggests a single ancestral Kindlin protein present in even the earliest metazoan ie, hydra. This protein then underwent duplication events in insects and also experienced genome duplication in vertebrates, leading to the Kindlin family. A comparative study of the Kindlin paralogs showed that Kindlin-2 is the slowest evolving protein among the three family members. The analysis of synonymous and non-synonymous substitutions in orthologous Kindlin sequences in different species showed that all three Kindlins have been evolving under the influence of purifying selection. The expression pattern of Kindlins along with phylogenetic studies supports the subfunctionalization model of gene duplication.

Keywords: kindlins, cell adhesion molecules, phylogenetics, natural selection, molecular evolution, integrin interacting proteins

Evolutionary Bioinformatics 2011:7 7–19

doi: [10.4137/EBO.S6179](https://doi.org/10.4137/EBO.S6179)

This article is available from <http://www.la-press.com>.

© the author(s), publisher and licensee Libertas Academica Ltd.

This is an open access article. Unrestricted non-commercial use is permitted provided the original work is properly cited.



Introduction

The Kindlins represent a class of focal adhesion proteins implicated in integrin activation. They comprise three evolutionarily conserved members, Kindlin-1 (FERMT1, C20orf42, URP1; chromosome 20p12.3), Kindlin-2 (FERMT2, MIG2, PLEKHC1, UNC112; chromosome 14q22.1) and Kindlin-3 (FERMT3, UNC-112 related protein 2; chromosome 11q13.1), that share considerable sequence and structural similarities.¹ The Kindlins have a bipartite FERM (four point one protein, ezrin, radixin, moesin) domain interrupted by a pleckstrin homology domain and can bind directly to various classes of integrins as well as participating in inside-out integrin activation.² The F3 sub domain of the FERM domains of all three Kindlins contains a crucial phosphotyrosine binding (PTB) fold resembling that of Talin.^{3,4} The relationship between Kindlins and integrin signaling was explored by initial studies on *C.elegans* embryos harboring homozygous mutations in UNC-112, the worm ortholog of KIND1. These embryos develop a paralyzed, arrested elongation at two-fold (PAT) phenotype due to failure of organization of PAT3/integrin in body muscle wall.⁵ Functionally, Kindlins primarily mediate protein-protein interactions and lack catalytic domains.⁶ In recent years they have emerged as a key class of adaptor molecules involved in integrin activation.⁷

Loss-of-function mutations in Kindlin-1 and Kindlin-3 cause Kindler syndrome and leukocyte adhesion deficiency-III syndrome, respectively. Kindler syndrome was in fact the first human genetic disorder clinically associated with Kindlins. It is caused by mutation of Kindlin-1 and is characterized by skin blistering, severe periodontitis and poikiloderma.^{8,9} No human disease has yet been associated with Kindlin-2 gene pathology, however Kindlin-2 knockout mice die in early embryonic stage indicating the essential role it plays in development.

The expression patterns of all three Kindlins are quite distinct. For instance, Kindlin-1 is predominantly expressed in the epidermis and only weakly expressed in the dermis, while Kindlin-3 expression is restricted exclusively to hematopoietic tissues, where it is the dominant form of Kindlins expressed. On the other hand, Kindlin-2 is ubiquitously expressed in most parts of the body.¹⁰ These differential expression patterns may in part explain the distinctive phenotypes

that result from the loss of different Kindlins. For instance, as noted above, the Kindlin-2 homolog in *C. elegans*, UNC-112 is essential for embryonic development.¹¹ Also, loss of Kindlin-2 in mice results in pre-implantation embryonic lethality and knockdown of Kindlin-2 in zebrafish reveals a strong relationship between cardiac development and the function of Kindlin-2.¹² Consistent with the restricted expression of Kindlin-3 in hematopoietic tissues, mice lacking Kindlin-3 show severe osteoporosis, hemorrhage and defects in the erythrocyte membrane skeleton, and die within one week after birth.^{13,14}

Kindlin-1 shares 62% sequence similarity with Kindlin-2 and 49% with Kindlin-3. Until now, the evolutionary aspects of Kindlin structure and function have not been addressed. In this study, we explored the evolution and divergence of these proteins in vertebrates and invertebrates. We studied the natural forces shaping the evolution of the Kindlin family of proteins by comparing different evolutionary trends in different vertebrate clades. A phylogenetic analysis of three Kindlin family members illustrates the phylogenetic history of Kindlin paralogs and also documents the duplication events leading to the formation of these paralogs from one single ancestral Kindlin. We show that the original Kindlin arose at least as early as simple metazoans, such as hydra. We also explored the effect of functional constraints on the evolution of these three paralogs in vertebrates and found evidence that purifying selection is a major force shaping the evolution of Kindlins.

Material Method

Relative levels of Kindlin-1, Kindlin-2 and Kindlin-3 transcripts were determined by real-time RT-PCR using SYBR Green. Human tissue cDNA panels (BD Biosciences) were used as a template. Triplicate samples of each PCR mixture, each containing 4.7 µl of POWER SYBR Green PCR master mixture (Applied Biosystems), 0.3 µl of a 10 pmol/µl of primer mixture, 0.3 µl of cDNA, and water to a total volume of 10 µl were transferred into a 96-well plate on an ABI 7500 Fast Real Time PCR System (Applied Biosystems). The samples were initially incubated at 95°C for 3 min, followed by 45 cycles with 95°C for 15 s, 60°C for 60 s. Dissociation curves were generated after each PCR run to ensure that a single, specific product was amplified. The results were analyzed with the comparative Cycle



threshold (Ct) method. For normalization, we used the expression level of β -actin (ACTB). The PCR primers are shown in Table 1.

In order to explore the evolutionary history of Kindlins, sequences of the complete transcripts and the corresponding protein sequences of all three Kindlins (Kindlin-1, Kindlin-2 and Kindlin-3) from different species were extracted from NCBI (<http://www.ncbi.nlm.nih.gov>) and ENSEMBLE (<http://www.ensembl.org>) genome browsers (Table 2). After alignment using CLUSTALW program,¹⁵ all positions containing gaps and missing data were eliminated, because of the possible ambiguity of the alignments. This stringent approach reduced the risk of misinterpretations. The evolutionary history was inferred by using the maximum likelihood method (ML) as implemented in the TREEFINDER (TF) program package.¹⁶ The TF support values indicate the reliability of internal branches. The analyses of amino acid sequences were performed using the WAG2000 model¹⁷ applying eight classes of rate heterogeneity among sites (8 Γ). The ML analysis involved 26 amino acid sequences. A Neighbor Joining analysis¹⁸ on amino acid sequences was done by the MEGA 4 program.¹⁹ The reliability of the NJ tree was estimated by the bootstrap method, based on 1000 pseudo replicates.

A variety of methods were employed to explore the functional constraints shaping the evolution of Kindlins. Pairwise comparison of the number of synonymous nucleotide substitutions per synonymous (dS) site and non-synonymous nucleotide substitutions per non-synonymous site (dN) was carried out by using Nei-Gojobori method.²⁰ In addition to pairwise methods, the dN/dS ratio in different branches of the maximum-likelihood tree was estimated using the codon-based genetic algorithm implemented in

the GA-BRANCH program available at the Datamonkey server (<http://www.datamonkey.org/help/GABranch.php>). This approach assigns each branch to an incrementally estimated class of dN/dS ratios without requiring a specification of the branches *a priori* and is both less parameterized and more user friendly than a fully local model.²¹

Evolutionary distance between all possible pairs of Kindlin paralogs was estimated by Tajima's relative rate test.²² Each pair of paralogs was compared with amphioxus protein sequences taken as an out-group. Mega 4 software was used for evolutionary analysis.¹⁹

Results and Discussion

Expression pattern of kindlins

Realtime PCR analyses of the expression pattern of Kindlins revealed distinct patterns of Kindlin expression. As expected, Kindlin-2 was expressed almost ubiquitously in all six of the tissues tested, while Kindlin-1 showed a one hundred-fold lower expression in each of these tissues with the exception of human kidneys where the expression level is low but significant. Kindlin-3 showed detectable expression in leukocytes only where it is expressed moderately (Fig. 1). These results are very much in agreement with existing data on the expression patterns of Kindlin proteins, which also show the ubiquitous nature of Kindlin-2 expression and the tissue specific expression of both Kindlin-1 and Kindlin-3.¹⁰

Phylogeny

Phylogenetic analysis was carried out on the amino acid sequences of Kindlin proteins, with the phylogenetic tree rooted by orthologous genes from invertebrate species (Fig. 2). The phylogenetic tree was calculated based on the maximum likelihood method. Notably, an NJ analysis on the same data predicted the same topology (Fig. 3). The resulting phylogenetic tree is very well supported for most branches except for the deep divergences of *Branchisotoma floridae* and *Strongylocentrotus purpuratus* and one leading to the grouping of *Xenopus laevis* Kindlin-2 with the mammals. Otherwise the vertebrate relationships for each Kindlin ortholog are resolved in a manner concomitant with previous phylogenomic analyses,²³ indicating a reliable evolutionary reconstruction among the Kindlin families.

Table 1. Primers used for Realtime PCR.

Sequence name	Forward primer	Reverse primer
Kindlin-1	CATGCTGTCATC CACTGACTTTAC	TCAATCCTGAC CGCCGGTCAA
Kindlin-2	CCATGGCTCTG GACGGGATAAGG	TCACACCCAAC CACTGGTAAG
Kindlin-3	GAGACCCACCTG CAGCCCCCAG	AAACACCCGC AGTCCCATGAC
β -actin	CAAGGCCAACCG CGAGAAGATGAC	GCCAGAGGCGT ACAGGGATAGCACA

**Table 2.** Names and IDs of peptides and transcripts of Kindlin genes.

Sequence name	Peptide ID	Transcript ID
Human (<i>Homo sapiens</i>)		
Kindlin-1	AAN75822.1	AF443278_1
Kindlin-2	NP_006823.1	NM_006832.2
Kindlin-3	NP_848537.1	NM_178443.2
Mouse (<i>Mus musculus</i>)		
Kindlin-1	NP_932146.2	NM_198029.2
Kindlin-2	NP_666166.2	NM_146054.2
Kindlin-3	NP_722490.1	NM_153795.1
Rat (<i>Rattus norvegicus</i>)		
Kindlin-1	NP_001099985.1	NM_001106515.1
Kindlin-2	NP_001011915.1	NM_001011915.1
Kindlin-3	NP_001121015.1	NM_001127543.1
Frog (<i>Xenopus leavis</i>)		
Kindlin-1	NP_001079432.1	NM_001085963.1
Kindlin-2	NP_001086955.1	NM_001093486.1
Kindlin-3	NP_001004882.1	NM_001004882.1
Fugu (<i>Takifugu rubripes</i>)		
Kindlin-1	ENSTRUP00000018258	ENSTRUT00000018334
Kindlin-2	ENSTRUP00000036787	ENSTRUT00000036919
Kindlin-3	ENSTRUP00000035295	ENSTRUT00000035423
Zebrafish (<i>Danio rerio</i>)		
kindlin-1	ENSARP00000069034	ENSART00000074546
kindlin-2	XP_685536.2	XM_680444.3
Kindlin-3	NP_957198.1	NM_200904.1
Amphioxus (<i>Branchiostoma floridae</i>)		
BRAFLDRAFT_285279	XP_002590896	XM_002590850.1
Acorn Worm (<i>S. Kowalevskii</i>)		
fermitin 2-like	XP_002741351.1	XM_002741305.1
Sea Urchin (<i>Strongylocentrotus purpuratus</i>)		
Similar to Plekhc1-prov protein	XP_784927.2	XM_779834.2
Fruitfly (<i>Drosophila melanogaster</i>)		
Fermitin-1	NP_728936.1	NM_168060.1
Fermitin-2	NP_648947.1	NM_140690.2
Ciano Intestilis		
Novel protein	ENSCINP0000002380	ENSCINT0000002380
<i>C. elegans</i>		
UNCoordinated family member (unc 112)	NP_506628.1	NM_074227.5
Hydra (<i>Hydra magnipapillata</i>)		
Novel protein	XP_002158978.1	XM_002158942.1

The resulting phylogenetic tree suggests an interesting evolutionary history for Kindlin family proteins. The phylogeny exhibits a topology of the form (A)(BC) ie, Kindlin-1 and Kindlin-3 form a cluster while Kindlin-2 forms an out-group. It appears that in invertebrates before the divergences of arthropods, a single ancestral Kindlin had its ancient origin in hydra. This gene then underwent a lineage specific duplication event in insects, giving rise to two Kindlin paralogs. Although no study exists to date on the roles of Kindlins in insects, it is plausible

to assume that because of the diversity found in insects, the duplicated copies were maintained in response to selection pressures impacting this highly diverse group. One of the two Kindlins was lost in other higher phyla before the origin of vertebrates ie, echinoderms (*Strongylocentrotus purpuratus*) urochordates (*Ciona intestinalis*), cephalochordates (*Amphioxus*) and hemichordates (*Saccoglossus kowalevskii*). However, the remaining single Kindlin gene copy underwent two duplication events in vertebrates that may have occurred together with two

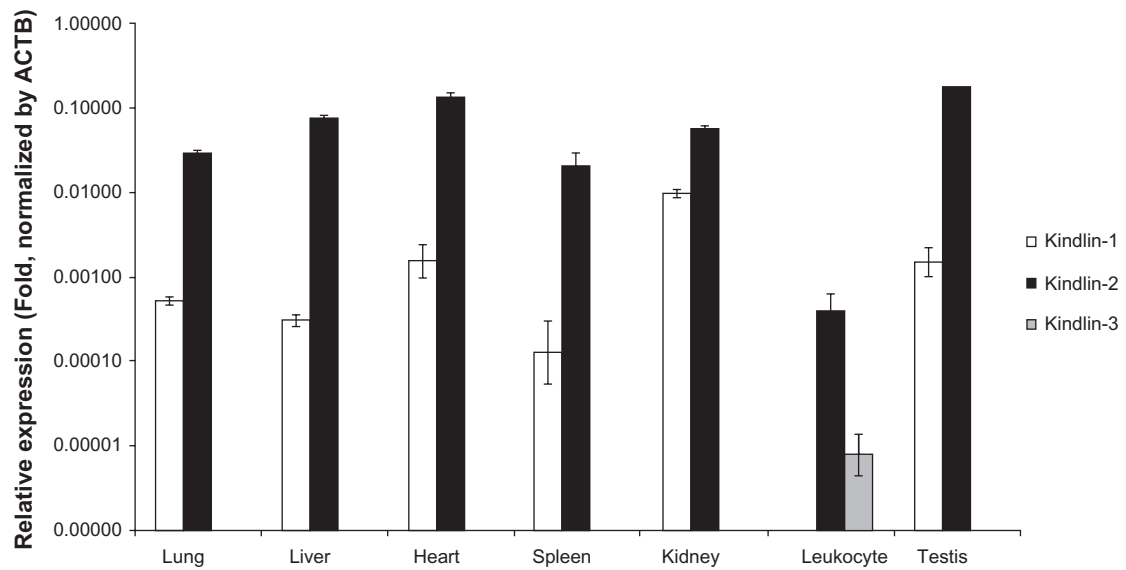


Figure 1. The expression profile of all three Kindlin paralogs was quantified by real-time RT-PCR using SYBR Green. Data are presented as the relative expression of Kindlins (fold change) normalized by a housekeeping gene, β -actin. **Note:** The error bars indicate the standard deviation.

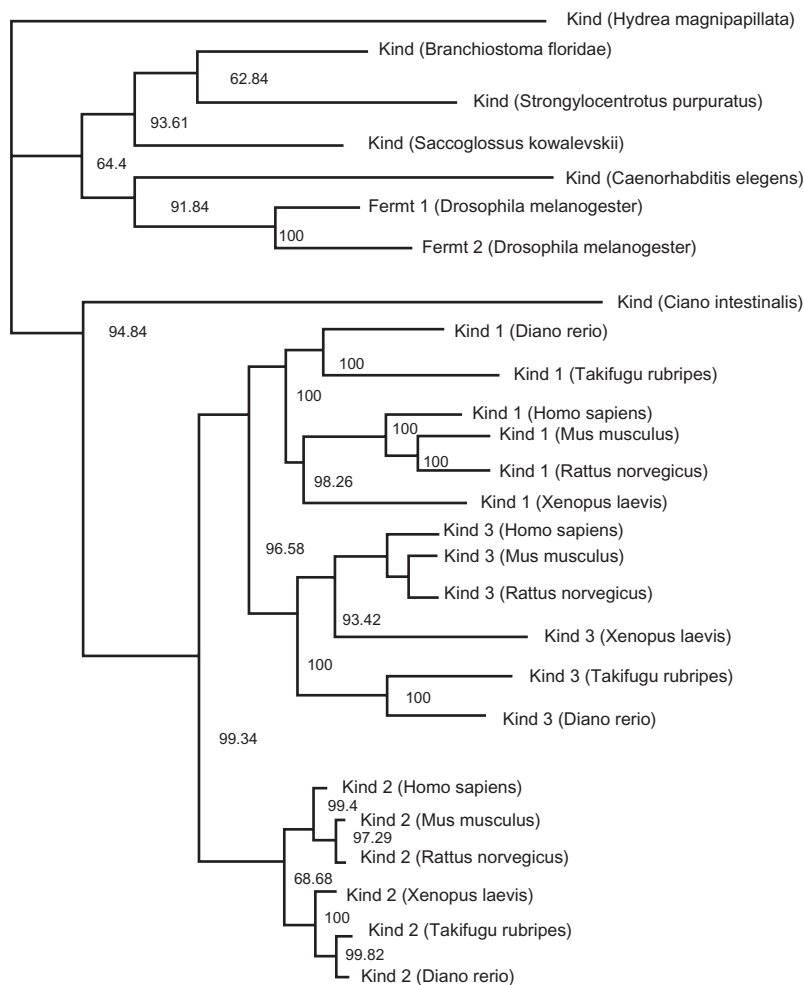


Figure 2. The evolutionary history of the Kindlin protein family inferred by using the Maximum likelihood method.

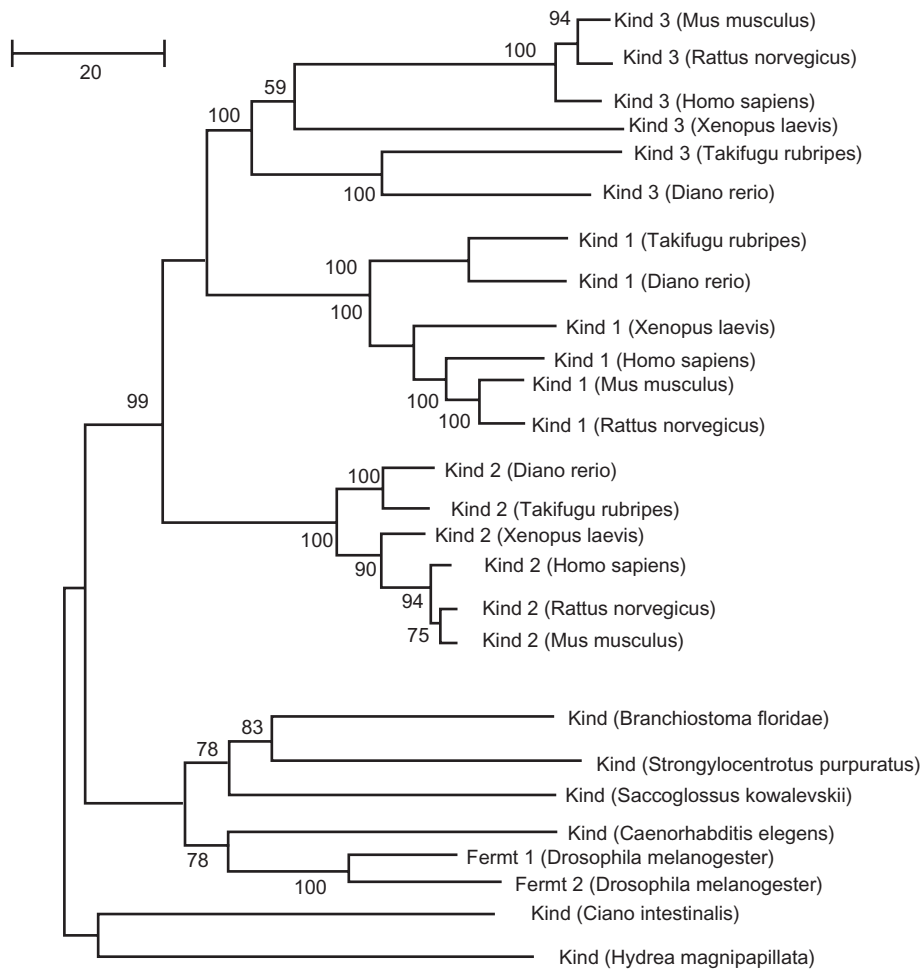


Figure 3. The evolutionary history of the Kindlin protein family inferred by using the neighbor joining method.

rounds of genome duplication. Regardless of the exact mechanism, these events gave rise to three Kindlin paralogs: Kindlin-1, Kindlin-2 and Kindlin-3. More data is still required to determine whether these Kindlin gene duplication events in the fish were due to whole genome duplications or individual segmental gene duplication.

Whatever was the cause of this duplication, it is evident from our analyses that three vertebrate Kindlin paralogs originated after duplication events in the fish genome and were maintained in all subsequent vertebrate forms, probably due to the selection pressures that have promoted the diverse and complicated morphological and physiological properties of vertebrates.^{24–27} It is important to notice that branch lengths of Kindlin-2 are quite short when compared to both Kindlin-1 and Kindlin-3. Very short branch lengths for Kindlin-2 on the phylogenetic tree indicate that this gene has experienced a

slower evolutionary rate than its paralogs. This result is congruent with existing experimental studies on Kindlin genes.^{9,10,13,14} For instance, Kindlin-2 is a ubiquitously expressed gene playing its structural and functional roles in broad array of tissues.¹⁰ It has been indicated in various studies that ubiquitously expressed genes tend to evolve slowly compared to those with tissue specific expression.²⁸ As noted previously, Kindlin-1 and Kindlin-3 are expressed in specific tissues (epithelial tissues and the hematopoietic system, respectively) (Fig. 1), and it is therefore commensurate that the evolutionary rate of these paralogs is much higher than that of Kindlin-2. In support of this are Kindlin knock out studies for all three Kindlins which show that Kindlin-2 knockout mice die during early embryogenesis.²⁹ In contrast to the milder phenotypes of Kindlin-1 and Kindlin-3, these studies support the idea that Kindlin-2 is under tighter functional constraint than Kindlin 1 and 3.

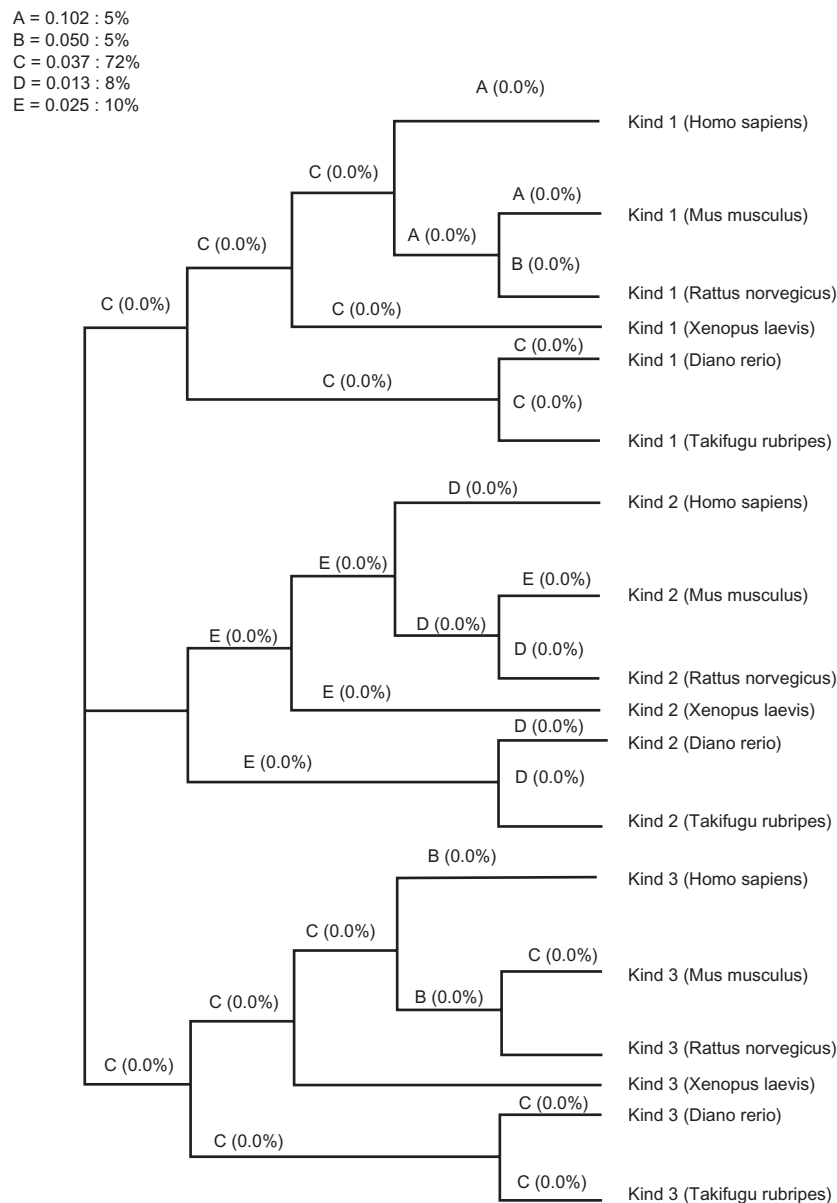


Figure 4. Lineage-specific analysis of selective pressure in vertebrate Kindlins. A cladogram is shown with maximum-likelihood estimates of lineage-specific d_N/d_S during vertebrate Kindlin evolution. Percentages for branch classes in the legend reflect the proportion of total tree length (measured in expected substitutions per site per unit time) evolving under the corresponding value of dN/dS .

Note: While letters A, B, C, D and E represents each branch class.

Estimation of the selective forces shaping the evolution of Kindlin proteins

Comparison between non-synonymous and synonymous substitutions in orthologous transcript sequences can reveal the selective pressure that shapes the evolution of these genes. The ratio between non-synonymous substitution per non-synonymous site to synonymous substitution per synonymous site d_N/d_S or omega (ω) indicates whether the evolution of genes is due to adaptive selection or due to neutral evolution. A value of ω more than 1 suggests that the gene

is under positive selection. A value close to 1 suggests that a gene is under neutral selection and is experiencing neutral evolution. However, a value of less than 1 indicates that a gene is under the influence of negative or purifying selection. The Nei-Gojobori method we employed to estimate the d_N/d_S ratio clearly showed ω values below one, (ie, $d_N/d_S < 1$) for all the three Kindlin paralogs in all of the vertebrate species compared for this analysis (Table 3). Interestingly, the ω value for Kindlin-2 was much lower (at least ten fold less) than ω values for either of the other two

**Table 3.** Estimation of dn/ds values for Kindlin orthologs.

	Human	Mouse	Rattus	Xenopus	Diano
Kindlin-1					
Mouse	0.06/0.46				
Rattus	0.05/0.47	0.01/0.21			
Xenopus	0.12/0.81	0.14/0.82	0.14/0.80		
Diano	0.18/0.85	0.19/0.80	0.19/0.80	0.18/0.82	
Fugu	0.200/0.836	0.2109/0.818	0.211/0.7879	0.2044/0.8378	0.143/ 0.72478
Kindlin-2					
Mouse	0.01/0.80				
Rattus	0.01/0.78	0.004/0.51			
Xenopus	0.03/0.70	0.03/0.72	0.03/0.80		
Diano	0.05/0.90	0.05/0.73	0.04/0.75	0.05/0.83	
Fugu	0.05/0.87	0.06/0.78	0.06/0.7707	0.06/0.8526	0.02/0.57
Kindlin-3					
Mouse	0.03/0.43				
K3-Rattus	0.03/0.45	0.008/0.212			
K3-Xenopus	0.25/0.76	0.26/0.80	0.26/0.76		
K3-Diano	0.241/0.71	0.24/0.71	0.24/0.72	0.28/0.79	
K3-Fugu	0.25/0.65	0.25/0.70	0.25/0.69	0.3038/0.7795	0.16/0.70

Kindlins. Similarly the analysis within and between mammalian and non-mammalian groups clearly showed the same pattern observed in individual pairwise comparisons between different species ie, the absence of any positive selection whatsoever in all the groups. However, ω values were slightly higher in non-mammals than in mammals. This difference may result from the high substitution rate often seen in fish genomes, leading to higher dn/ds ratios. Kindlin-2 in both mammals and non-mammals showed values of ds similar to Kindlin-1 and Kindlin-3, but with much lower values of dn, thereby producing greatly lowered values of ω (Table 4).

To gain further insight into the lineage specific nature of the selective pressures acting on each branch of the phylogenetic tree, we performed a genetic algorithm, namely (Ga)-branch analysis. The GA-branch method is an alternative to the branch site method. This method, unlike the branch site method, does not require the manual selection of branches of interest to identify evidence for positive or negative selection. Because the GA branch method does not require the user to select branches of interest, or that testing be performed one branch at a time, it experiences reduced statistical instability while also offering improved interpretability for poorly supported models. It achieves this by mining the data for good-fitting models. In addition, inferences based on

multiple models (as opposed to a null-alternative pair) are less vulnerable to model misspecification. In our study, Ga-branch analysis selected a model with five classes of ω . In total, 72% of branches are assigned to a ω of 0.037, named as Class D, with the remaining 28% of branches assigned to four additional Classes, designated A,B,C and E, with ω values of 0.102, 0.050, 0.037 and 0.026, respectively. None of the branches studied show any trend for positive selection, with the probability of positive selection being 0% for each branch of the tree. Notably, here again the very low value of ω (0.013) for Kindlin-2 indicates that it is evolving under the influence of much stronger negative selection than Kindlin-1 and Kindlin-3 (Fig. 4 and Table 5).

In short, whether calculated by pairwise comparison or by lineage specific analysis, dn/ds ratios consistently indicate that vertebrate Kindlins have been evolving under the influence of purifying selection, with Kindlin-2 under much stronger negative selection than Kindlin-1 and Kindlin-3. Functional studies on Kindlins also very much support this trend. For instance, if adaptive selection was the main force for Kindlin divergence rather than purifying selection, we would expect that the functional roles of Kindlins may be diverse. However, such diversity is not evident from data available on Kindlin function. In fact, the hallmark function of Kindlins is integrin

Table 4. Average Ka and Ks values between and within mammalian - non-mammalian lineages for Kindlin orthologs.

	ds	dn	dn/ds	T-value	P value		Conclusion
					for positive selection	for negative selection	
Kindlin-1							
Within mammals	0.384 (0.13)	0.041 (0.02)	0.1 (0.02)	-58.9188	0.999856	0.000144	Negative selection
Within non-mammals	0.79 (0.06)	0.18 (0.03)	0.223 (0.02)	-64.6226	0.99988	0.00012	Negative selection
Between mammals and non-mammals	0.812 (0.02)	0.176 (0.01)	0.217 (0.04)	-55.397	1	0.00001	Negative selection
Kindlin-2							
Within mammals	0.405 (0.15)	0.007 (0.001)	0.02 (0.003)	-553.15	0.999998	0.000002	Negative selection
Within non-mammals	0.75 (0.15)	0.046 (0.02)	0.056 (0.02)	-70.75	0.9999	0.0001	Negative selection
Between mammals and non-mammals	0.78 (0.06)	0.04 (0.01)	0.06 (0.016)	-169.4494	1	0.000001	Negative selection
Kindlin-3							
Within mammals	0.36 (0.1)	0.023 (0.1)	0.06 (0.01)	-94	0.999943	0.000057	Negative selection
Within non-mammals	0.76 (0.05)	0.25 (0.08)	0.326 (0.08)	-13.71	0.99736	0.0026	Negative selection

Note: P-values for both positive and negative selection are included here. Standard deviations are enclosed in brackets.

activation and all the other higher order functions associated with Kindlins, including cell migration, cell spreading, cell adhesion, cellular signaling and cancer promotion are associated with the ability of Kindlins to activate integrin. On the other hand, these higher order processes impacted by Kindlins – through integrin activation – are so essential for organism survival and viability that very strong functional constraints exist over Kindlin evolution. The more stringent functional constraint of Kindlin-2 compared to either of its two counterpart paralogs is likely due to the fact that it is expressed ubiquitously in the body while Kindlins 1 and 3 are tissue specific.

Divergence rate between paralogs

For studying the evolutionary distances between different Kindlin paralogs, Tajima's relative rate test was employed. This test involves pairwise comparison of protein sequences from Kindlin paralogs of each species while using the orthologous Kindlin sequences from amphioxus (*Branchiostoma floridae*) as an out-group (Table 6). These tests produced intriguing results which, in correlation with phylogenetic data, show that Kindlin-2 has undergone relatively limited divergence compared to both Kindlin-1 and Kindlin-3, of which Kindlin-3 is indicated to be most divergent. Interestingly, it seems that Kindlin-2 is the representative of the original ancient Amphioxus Kindlin which underwent two duplications in fish giving rise to Kindlin-1 and Kindlin-3. There are two important reasons to believe that Kindlin-2 is the representative of the unduplicated ancestral Kindlin gene, whose duplication in vertebrates gave rise to two other paralogs. Firstly the Tajima test clearly shows that Kindlin-2 is closest to Amphioxus Kindlin, with very low levels of divergence evident compared to Kindlin-1 and Kindlin-3. Secondly, Kindlin-2 is not only a ubiquitously expressed protein but also the only Kindlin protein expressed in embryonic stem cells. Although no study has been conducted to explore the expression pattern of Kindlins in Amphioxus, *C. elegans* studies suggest that this unduplicated Kindlin is expressed ubiquitously in all tissues including the embryonic stem cell.⁵ Therefore if the Amphioxus Kindlin is an ortholog of *C. elegans* Kindlin, it should also exhibit the same expression pattern, making vertebrate Kindlin-2 a strong candidate as the representative of the ancestral unduplicated Kindlin protein. No species specific

**Table 5.** Lineage specific dn/ds values Kindlins.

Branch name	Mean dn/ds	Std. Dev.	2.5%	Median	97.5%	Prob {dN > dS} [†]
K1_HUMAN	0.102	0.002	0.093	0.103	0.103	0.000
K1_MOUSE	0.102	0.003	0.093	0.103	0.103	0.000
K1_RATTUS	0.042	0.007	0.037	0.040	0.054	0.000
Node5	0.102	0.002	0.093	0.103	0.103	0.000
Node3	0.038	0.002	0.037	0.037	0.045	0.000
K1_XENOPUS	0.038	0.001	0.037	0.037	0.042	0.000
Node2	0.037	0.001	0.037	0.037	0.040	0.000
K1_DIANO	0.038	0.001	0.037	0.037	0.041	0.000
K1_FUGU	0.038	0.001	0.037	0.037	0.040	0.000
Node9	0.037	0.001	0.037	0.037	0.039	0.000
Node1	0.038	0.001	0.037	0.037	0.040	0.000
K2_HUMAN	0.013	0.002	0.011	0.013	0.017	0.000
K2_MOUSE	0.028	0.006	0.014	0.026	0.042	0.000
K2_RATTUS	0.014	0.002	0.011	0.013	0.020	0.000
Node16	0.014	0.002	0.011	0.013	0.021	0.000
Node14	0.026	0.003	0.017	0.026	0.031	0.000
K2_XENOPUS	0.026	0.003	0.023	0.026	0.037	0.000
Node13	0.026	0.003	0.018	0.026	0.031	0.000
K2_DIANO	0.013	0.002	0.011	0.013	0.017	0.000
K2_FUGU	0.014	0.002	0.013	0.013	0.024	0.000
Node20	0.027	0.003	0.024	0.026	0.038	0.000
Node12	0.037	0.001	0.037	0.037	0.039	0.000
K3_HUMAN	0.043	0.007	0.037	0.040	0.055	0.000
K3_MOUSE	0.036	0.006	0.026	0.037	0.045	0.000
K3_RATTUS	0.037	0.004	0.027	0.037	0.044	0.000
Node27	0.043	0.008	0.037	0.040	0.056	0.000
Node25	0.037	0.001	0.031	0.037	0.038	0.000
K3_XENOPUS	0.037	0.000	0.037	0.037	0.038	0.000
Node24	0.037	0.003	0.027	0.037	0.040	0.000
K3_DIANO	0.038	0.001	0.037	0.037	0.041	0.000
K3_FUGU	0.038	0.001	0.037	0.037	0.040	0.000
Node31	0.038	0.001	0.037	0.037	0.040	0.000
Node23	0.037	0.001	0.037	0.037	0.039	0.000

Note: [†]Not a *P*-value! It only represents probability in percentage of any positive selection ie, dn/ds > 1 in a given branch.

asymmetry was found in the divergence pattern of any of the three Kindlin paralogs which is in agreement with the generally accepted principle that paralogs evolve at similar rates in different species.³⁰

Relatively high divergence of Kindlin-1 and Kindlin-3 relative to Kindlin-2 can be explained by analyzing the expression pattern of the proteins.¹⁰ During evolution, from Amphioxus to higher vertebrates, the expression pattern of the two paralogs have diverged. Our analysis of Kindlin expression patterns in human tissue samples clearly shows that unlike Kindlin-2, which is expressed ubiquitously, both Kindlin-1 and Kindlin-3 present tissue specific expression patterns (Fig. 1). Similarly, Ussar et al have shown in their study that both Kindlin-1 and Kindlin-3 are expressed predominantly in epithelial tissues and the hematopoietic system respectively.¹⁰

The distinct expression patterns of all three Kindlins in vertebrates thus supports the subfunctionalization model of gene duplication.³¹ It seems likely that the function of the ancestral unduplicated Kindlin was subfunctionalized in vertebrates in part due to the divergence of Kindlin expression location. Thus, while Kindlin-1 and Kindlin-3 are expressed exclusively in epithelial and hematopoietic tissues, respectively, Kindlin-2 – being the representative of ancestral Kindlin gene – is expressed in a variety of tissues, in a pattern less ubiquitous than the original unduplicated Kindlin gene. Ultimately, it seems that this subfunctionalization of Kindlin expression patterns may have provided a degree of selective advantage associated with the diversification of higher order functions performed by integrins in various tissues.

Table 6. Tajima's relative rate test for the comparison of evolutionary distance between Kindlin paralogs in different species using Amphioxus BRAFLDRAFT_285279 as an outgroup.

Tajima's evolutionary test							
Paralogs	Identical sites in all three sequences	Divergent sites in all three sequences	Unique differences in Sequence A	Unique differences in Sequence B	Unique differences in Sequence C	X ²	df P
Human							
Kindlin-2/Kindlin-1	293	129	42	75	130	9.31	1 0.00228*
Kindlin-2/Kindlin-3	257	133	50	107	97	20.69	1 0.00001*
Kindlin-1/Kindlin-3	251	144	56	81	111	4.56	1 0.03269
Mouse							
Kindlin-2/Kindlin-1	294	124	43	74	134	8.21	1 0.00416*
Kindlin-2/Kindlin-3	258	138	47	106	97	22.75	1 0.00001*
Kindlin-1/Kindlin-3	255	154	50	79	107	6.52	1 0.01067
Rat							
Kindlin-2/Kindlin-1	294	126	43	75	131	8.68	1 0.0032*
Kindlin-2/Kindlin-3	258	135	48	107	98	22.46	1 0.00001*
Kindlin-1/Kindlin-3	253	155	53	81	103	5.85	1 0.01557
Frog							
Kindlin-2/Kindlin-1	288	119	54	78	128	4.36	1 0.03671
Kindlin-2/Kindlin-3	250	140	47	112	96	26.57	1 0.0001*
Kindlin-1/Kindlin-3	243	161	54	96	88	11.76	1 0.00061*
Fugu							
Kindlin-2/Kindlin-1	257	135	52	97	117	13.59	1 0.00023*
Kindlin-2/Kindlin-3	241	125	56	110	98	17.57	1 0.00003*
Kindlin-1/Kindlin-3	235	153	62	69	107	0.37	1 0.54081
Zebra fish							
Kindlin-2/Kindlin-1	268	123	46	81	133	9.65	1 0.00190*
Kindlin-2/Kindlin-3	250	120	51	100	105	15.9	1 0.00007*
Kindlin-1/Kindlin-3	243	144	57	66	105	0.66	1 0.41708

Note: Sequence A means protein on the left hand side of slash (/) while the sequence B means the protein on right hand side. Sequence C represents Amphioxus (acting as out group here). P-value less than 0.05 is used to reject the null hypothesis of equal rates between lineages.



Conclusion

Kindlins represent a recently identified family of integrin interacting proteins. They play an important role in cell migration, cell spreading and cancer progression through a core molecular mechanism of integrin activation. In this study we have shown that the ancestral Kindlin gene was an unduplicated single gene found in organisms as primitive as hydra. This gene then underwent a lineage specific duplication in insects, giving rise to two Kindlin paralogs, while the three paralogs of Kindlin found in Vertebrates are the result of duplication events that occurred in fish. Of the three Kindlins, Kindlin-2 has undergone the least evolutionary divergence, probably due to stringent functional constraints it associated with its virtually ubiquitous expression pattern in body tissues and especially in embryonic stem cells. On the other hand both Kindlin-1 and Kindlin-3 showed significantly greater divergence as a result of significantly weaker functional constraints, possibly resulting from their subfunctionalization into very specific portions of the body. The comparison of synonymous to non-synonymous substitutions both by a pairwise method as well as a lineage specific method also indicate that all three Kindlins have been evolving under strong negative selection.

Author Contributions

AAK conceived the idea. AAK, AJ, TS performed the analysis. AAK, AJ, TS and HZ analyzed the data. AAK, AJ and HZ wrote the paper.

Acknowledgments

This work was supported by Grants to H.Z. from The Swedish Cancer Foundation, The Swedish Society of Medicine, The Swedish Research Council, The Karolinska Institute, The Nature Science Foundation of China grant 30830048, the Basic Research Project of the Ministry of Science and Technology of China (973) 2010CB912203 and 2010CB529402, and the Peking University 985 Project, to SS from The Center for Biosciences at Karolinska Institutet and to AAK from Higher education commission of Pakistan. We are thankful to Beston Nore for his helpful suggestions and to John Lock for reading and editing the manuscript.

Disclosure

This manuscript has been read and approved by all authors. This paper is unique and is not under consideration by any other publication and has not been published elsewhere. The authors and peer reviewers of this paper report no conflicts of interest. The authors confirm that they have permission to reproduce any copyrighted material.

References

1. Siegel DH, Ashton GH, Penagos HG, et al. Loss of kindlin-1, a human homolog of the *Caenorhabditis elegans* actin-extracellular-matrix linker protein UNC-112, causes Kindler syndrome. *Am J Hum Genet.* 2003;73:174–87.
2. Lai-Cheong JE, Parsons M, McGrath JA. The role of kindlins in cell biology and relevance to human disease. *The International Journal of Biochemistry and Cell Biology.* 2010;42:595–603.
3. Kloeker S, Major MB, Calderwood DA, Ginsberg MH, Jones DA, Beckerle MC. The Kindler syndrome protein is regulated by transforming growth factor-beta and involved in integrin-mediated adhesion. *J Biol Chem.* 2004;279:6824–33.
4. Shi X, Ma YQ, Tu Y, et al. The MIG-2/integrin interaction strengthens cell-matrix adhesion and modulates cell motility. *J Biol Chem.* 2007;282: 20455–66.
5. Rogalski TM, Mullen GP, Gilbert MM, Williams BD, Moerman DG. The UNC-112 gene in *Caenorhabditis elegans* encodes a novel component of cell-matrix adhesion structures required for integrin localization in the muscle cell membrane. 2000;150(1):253–64.
6. Larjava H, Plow EF, Wu C. Kindlins: essential regulators of integrin signaling and cell-matrix adhesion. *EMBO Rep.* 2008;9(12):1203–8.
7. Svensson L, Howarth K, McDowall A, Patzak I, Evans R, Ussar S, Moser M, Metin A, Fried M, Tomlinson I, Hogg N. Leukocyte adhesion deficiency-III is caused by mutations in KINDLIN3 affecting integrin activation. *Nat Med.* 2009;15(3):306–12.
8. Kindler T. Congenital poikiloderma with traumatic bulla formation and progressive cutaneous atrophy. *The British Journal of Dermatology.* 1954;66, 104–11.
9. Sharma RC, Mahajan V, Sharma NL, Sharma AK. Kindler syndrome. *International Journal of Dermatology.* 2003;42:727–32.
10. Ussar S, Wang HV, Linder S, Fassler R, Moser M. The Kindlins: subcellular localization and expression during murine development. *Experimental Cell Research.* 2006;312:3142–51.
11. Dowling JJ, Gibbs E, Russell M, et al. Kindlin-2 is an essential component of intercalated discs and is required for vertebrate cardiac structure and function. *Circ Res.* 2008a;102:423–31.
12. Dowling JJ, Vreede AP, Kim S, Golden J, Feldman EL. Kindlin-2 is required for myocyte elongation and is essential for myogenesis. *BMC Cell Biol.* 2008b;9:36.
13. Kruger M, Moser M, Ussar S, et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell.* 2008;134:353–64.
14. Moser M, Nieswandt B, Ussar S, Pozgajova M, Fassler R. Kindlin-3 is essential for integrin activation and platelet aggregation. *Nature Medicine.* 2008;14:325–30.
15. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research.* 1994;22:4673–80.
16. Jobb G, von Haeseler A, Strimmer K. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology.* 2004;4:18.
17. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution.* 2001;18:691–9.



18. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 1987;4:406–25.
19. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*. 2007;24:1596–9.
20. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*. 1986;3:418–26.
21. Pond SL, Frost SD. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. *Molecular Biology and Evolution*. 2005;22:478–85.
22. Tajima F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics*. 1993;135:599–607.
23. Hallstrom BM, Janke A. Gnathostome phylogenomics utilizing lungfish EST sequences. *Molecular Biology and Evolution*. 2009;26:463–71.
24. Aburomia R, Khaner O, Sidow A. Functional evolution in the ancestral lineage of vertebrates or when genomic complexity was wagging its morphological tail. *Journal of Structural and Functional Genomics*. 2003; 3:45–52.
25. Holland PW, Garcia-Fernandez J, Williams NA, Sidow A. Gene duplications and the origins of vertebrate development. *Development (Cambridge, England)*. 1994:125–33.
26. Ruddle FH, Bentley KL, Murtha MT, Risch N. Gene loss and gain in the evolution of the vertebrates. *Development (Cambridge, England)*. 1994; 155–61.
27. Wagner GP, Amemiya C, Ruddle F. Hox cluster duplications and the opportunity for evolutionary novelties. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100:14603–6.
28. Zhu J, He F, Hu S, Yu J. On the nature of human housekeeping genes. *Trends Genet*. 2008;24:481–4.
29. Montanez E, Ussar S, Schifferer M, et al. Kindlin-2 controls bidirectional signaling of integrins. *Genes and Development*. 2008;22:1325–30.
30. Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV. Selection in the evolution of gene duplications. *Genome Biology*. 2002;3:RESEARCH0008.
31. Lynch M, Force A. The probability of duplicate gene preservation by subfunctionalization. *Genetics*. 2000;154:459–73.

Publish with Libertas Academica and every scientist working in your field can read your article

“I would like to say that this is the most author-friendly editing process I have experienced in over 150 publications. Thank you most sincerely.”

“The communication between your staff and me has been terrific. Whenever progress is made with the manuscript, I receive notice. Quite honestly, I’ve never had such complete communication with a journal.”

“LA is different, and hopefully represents a kind of scientific publication machinery that removes the hurdles from free flow of scientific thought.”

Your paper will be:

- Available to your entire community free of charge
- Fairly and quickly peer reviewed
- Yours! You retain copyright

<http://www.la-press.com>