

ORIGINAL ARTICLE

Evolution of the contact phase of vertebrate blood coagulation

M. B. PONCZEK,*¹ D. GAILANI† and R. F. DOOLITTLE*

*Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA; and †Departments of Pathology and Medicine, Vanderbilt University, Nashville TN, USA

To cite this article: Ponczek MB, Gailani D, Doolittle RF. Evolution of the contact phase of vertebrate blood coagulation. *J Thromb Haemost* 2008; 6: 1876–83.

Summary. *Background:* Previous reports have noted that factor (F) XI and FXII and prekallikrein (the contact phase proteases) are absent in fish. *Objectives:* A broad survey of recently completed genomes was undertaken to find where during the course of vertebrate evolution these coagulation factors appeared. *Methods:* BLAST searches were conducted for the various factors on genomes of lamprey, puffer fish, zebra fish, frog, chicken, platypus, and opossum. *Results:* It was confirmed that FXII is absent from fish; it is present in frog, platypus, and opossum, but is absent in chicken, an apparent example of gene loss. A single gene corresponding to the evolutionary predecessor of FXI and prekallikrein occurs in frog, chicken, and platypus. The opossum (a marsupial) has both prekallikrein and FXI, completing the full complement of these genes that occurs in eutherian mammals. *Conclusions:* The step-by-step accrual of genes for these factors by a series of timely gene duplications has been confirmed by phylogenetic analysis and other considerations.

Keywords: chicken, factor XI, factor XII, opossum, platypus, prekallikrein.

Introduction

The availability of numerous whole genome sequences has ushered in a new era of comparative biology whereby it is now possible to reconstruct the step-by-step evolution of complex pathways such as vertebrate blood coagulation [1–4]. For example, lampreys may have only predecessor genes for factor (F) V and FVIII on the one hand, and FIX and FX on the other [3]. Moreover, all fish appear to lack the genes for proteases that constitute the ‘contact’ phase of clotting [4].

Correspondence: Russell F. Doolittle, Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA. Tel.: +1 858 534 4417. fax: +1 858 534 4985; e-mail: rdoolittle@ucsd.edu

¹Present address: Department of General Biochemistry, University of Lodz, Banacha 12/16, 90-237 Lodz, Poland.

Received 27 June 2008, accepted 13 August 2008

In mammals, the contact phase proteases are FXI, FXII, and plasma prekallikrein (PK). Functionally, FXII and PK are reciprocally activated *in vitro* in the presence of charged surfaces such as kaolin; activated FXII (FXIIa) converts FXI to the protease FXIa, which in turn converts FIX to FIXa (Fig. 1). While the bleeding disorder associated with deficiency of FXI is relatively mild, and no bleeding diathesis is associated with FXII or PK deficiency [5], there is renewed interest in the role of this system in pathologic coagulation based on observations that FXII-deficient mice are resistant to vascular thrombosis [6,7]. As such, determining when during evolution the ‘contact factors’ appeared may bear on the longstanding debate about how important this system is to coagulation.

The notion that the ‘contact phase’ of clotting is absent from infra-mammalian vertebrates is longstanding. In the case of birds, reports dating back over the course of a century have made the point that chicken blood is slow to clot upon exposure to foreign surfaces [8,9], the observed behavior being attributed to an absence of FXI and FXII [10]. Other birds are also reported to lack FXII/contact activation, including vultures [11] and ostriches [12].

It has long been apparent that FXI and PK are so similar – both in sequence and domain arrangement – that the duplication that gave rise to these paralogs.² must have been quite recent, and it was suggested that one or the other might be absent in non-mammalian vertebrates [13]. Similarly, the cloning of hepatocyte growth factor activator (HGFA) revealed a close relationship to FXII and suggested an analogous situation for these proteins [14], although in this case the sequence divergence is considerably greater, even though the domain arrangements in HGFA and FXII are the same (Fig. 2).³

Here, we show where in the course of vertebrate evolution FXI, PK, and FXII make their first appearances. Three types of evidence are provided to support the conclusions:

²The terminology surrounding homologous genes can be confusing. Homologs, which are by definition descended from a common ancestor, are divided into two kinds: orthologs and paralogs. Orthologs are straight-line genetic descendants that are thought to have the same function. Paralogs are the result of gene duplications and may have different, albeit often similar, functions.

³In way of direct comparison, the human sequences for FXI and PK are 68 percent identical in their serine protease domains, but the same domains of human FXII and human HGFA are only 46 percent identical.

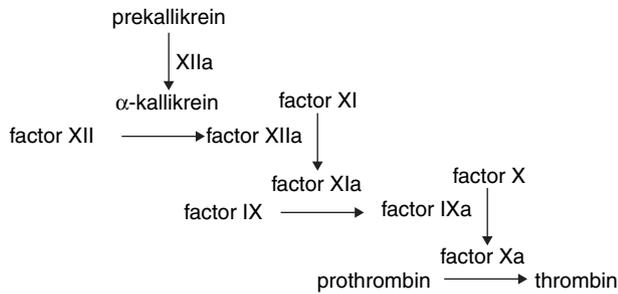


Fig. 1. Simplified outline of ‘contact factors’ involvement in mammalian blood coagulation (activation of factors IX and X by other routes not shown).

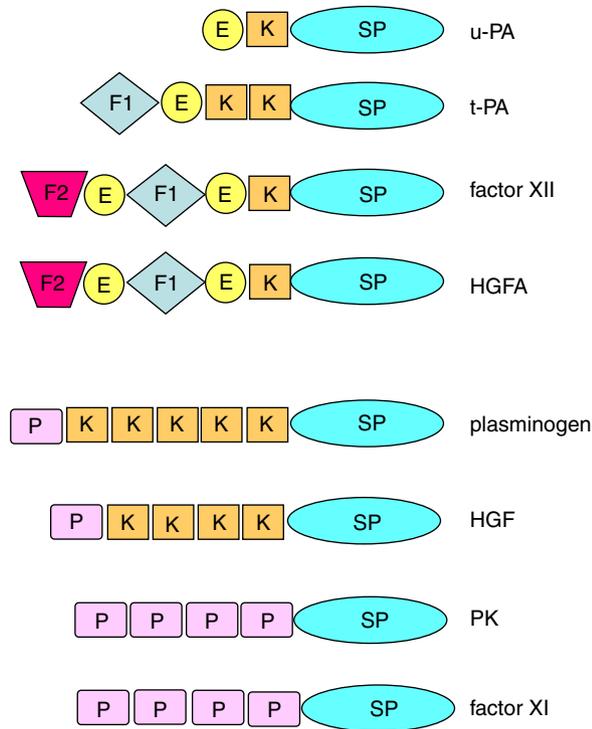


Fig. 2. Cartoon depiction of domains that occur in mammalian contact phase coagulation factors and some paralogs. SP, serine protease domain; K, kringle domain, E, EGF domain; P, PAN domain; F1, fibronectin type 1 domain; F2, fibronectin type 2 domain.

(i) sequence similarity and phylogenetic trees based on aligned sequences; (ii) the occurrence and arrangement of peripheral domains, including PAN (apple) domains, kringles, EGF, and fibronectin domains (Fig. 2); and (iii), when possible, synteny, as demonstrated by the arrangement of neighboring genes. Synteny is the co-localization of sets of genes on chromosomes.

Although the current paper is restricted to proteases, a recent phylogenetic study [15] has examined the evolution of kininogens, including high-molecular-weight kininogen (HK), the non-protease component of the contact system, providing a complementary backdrop to our analysis.

Methods

Several data sources were used during the course of this bioinformatics analysis, including the National Center for Biotechnology Information (NCBI), European Bioinformatics Institute (EMBL–EBI) and Washington University of St Louis Genome Center (<http://genome.wustl.edu>). Although these databases mostly mirror each other, each offers a different regimen of searching and map-viewing tools with different advantages depending on the degree of completeness of a particular whole genome sequence. The protocols available at the EMBL–EBI and Wellcome Trust–Sanger Centre (WTSI), including ENSEMBLE [16], were particularly helpful.

Surveys began with BLAST [17] searches of each of the human sequences (FXI, FXII, and PK) against the whole genome sequence databases; the highest scoring hits were then blasted reciprocally against the NCBI non-redundant database to see if there were better matches for the uncovered sequence. Preliminary reconstructions between exons were made with GeneScan [18], but final versions were made manually in the light of multiple amino acid sequence alignments. Alignments and phylogenetic trees were made both with CLUSTAL [19] and by an older progressive method [20,21]. Phylogenetic trees were drawn on the phylodendron website maintained at Indiana University (<http://iubio.bio.Indiana.edu/>).

The genomes analyzed include those of the opossum (*Monodelphis domestica*), platypus (*Ornithorhynchus anatinus*), chicken (*Gallus gallus*), frog (various species of the genus *Xenopus*), zebra fish (*Danio rerio*), puffer fish (*Fugu rubripes*), and lamprey (*Petromyzon marinus*). A simple phylogeny of these creatures is presented in Fig. 3.

The status of the various genomes varied from the draft assembly stage (e.g. platypus), where genes are mostly localized to contigs and supercontigs only [22], to virtually complete mapping of chromosomes (e.g. chicken) [23]. The possibility of syntenic gene arrangements for these various organisms was explored with ENSEMBLE [16].

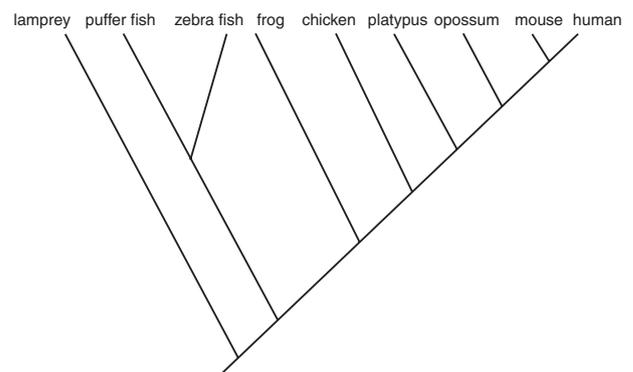


Fig. 3. Diagrammatic depiction of the phylogenetic relationships of organisms discussed in this article. A whole genome sequence is not yet available for any reptile.

Results

Genes for FXII in vertebrate genomes

The absence of a gene for FXII in fish was confirmed by a re-examination of the latest version of the puffer fish genome and a search of the more recently described zebra fish and lamprey genomes (Table 1). Genes for HGFA, a paralog of FXII, were found in lamprey, puffer fish (but not zebra fish), frog, chicken, platypus, and opossum. Genes for authentic FXII were identified in frog, platypus, and opossum genomes; the gene is not present in the chicken genome (Table 1). Phylogenetic evidence and chromosomal considerations presented below indicate that the gene has been lost on the lineage leading to birds.

Genes for FXI and PK in vertebrate genomes

A single paralog of FXI and PK occurs in frog, chicken, and platypus, suggesting its first appearance among early tetrapods (Table 1). In contrast, the opossum genome has genes for both PK and FXI, indicating that the gene duplication leading to separate factors occurred early in mammalian evolution but after the divergence of monotremes (platypus).

Chicken PK–FXI predecessor gene

The case for there being one PK–FXI gene in the chicken genome [23] is greatly strengthened not only by there being a single gene with appropriate sequence similarity and domain arrangement, but also by its detailed chromosomal location relative to neighboring genes. In humans, the genes for PK and FXI are adjacent to each other and between genes for a cytochrome P-450 and a melatonin receptor (MTNR1). In the chicken, only a single gene occurs between the same two genes (Fig. 4).

Platypus PK–FXI predecessor gene

Of all the identifications reported here, those involving sequences related to PK or FXI in the platypus genome are

the most problematic. There are two reasons. First, the platypus genome sequence is still at the 'draft' stage [22], and numerous short regions remain unsequenced (NNNNN regions). Secondly, the contigs on which the relevant exons occur are not fully assembled. As such, caution needs to be observed in certain circumstances.

One such circumstance involves an anomalous termination codon (TAG) in the middle of the protease portion of the PK–FXI predecessor sequence; in reality, it most likely encodes a tryptophan (TGG). There are two reasons to think this anomaly was introduced at the cloning stage. First, the rest of the region is free of termination codons, and secondly the sequence following the anomalous codon fits exactly into the expected pattern at the same high level of similarity (> 60% identity). Moreover, when the putative sequence is examined phylogenetically with the serine protease domains of other clotting factors, it assumes an appropriate place at the base of the cluster composed of FXI and PK (Fig. 5). This argument is presented more fully in the Discussion section.

The terminal three PAN domains of the platypus PK–FXI predecessor are clustered on a single supercontig (co29087); the fourth PAN domain is on a contig that also contains the amino-terminal half of the serine protease domain (co41273). No terminator codons occur in either of these contigs. The region containing the second half of the serine protease domain is on a third contig (co73301) containing the anomalous terminator codon.

Separate genes in the opossum

There is no ambiguity in the assignment of the two distinct genes in the opossum to FXI and PK. An examination of the opossum FXI gene identified several features that are highly similar to FXI in placental mammals. First, for the sequence available, human and opossum FXI are 70% identical at the amino acid level. By comparison, human and mouse FXI are 78% identical. Secondly, a conserved sequence (amino acids 183–191) in the third PAN domain in placental mammals, which likely represents a binding exosite for the substrate FIX, is conserved in the opossum protein [24]. Finally, FXI is a disulfide bond-linked homodimeric protein with the fourth

Table 1 Occurrence of genes for contact phase factors and some paralogs in assorted vertebrate genomes

	Factor XI	Prekallikrein	Factor XII	HGFA	HGF	Plasminogen	t-PA
Human	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Opossum	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Platypus	Pred*	Pred*	Yes	Yes	Yes	Yes	Yes
Chicken	Pred*	Pred*	Yes	Yes	Yes	Yes	Yes
Frog	Pred*	Pred*	Yes	Yes	Yes	Yes	Yes
Zebra fish	No	No	No	No [†]	Yes	Yes	Yes
Puffer fish	No	No	No	Yes	Yes	Yes	Yes
Lamprey	No	No	No	Yes	Yes	Yes	Yes

HGF, hepatocyte growth factor; HGFA, hepatocyte growth factor activator; t-PA, tissue-type plasminogen activator. *Pred denotes single gene for evolutionary predecessor of factor XI and prekallikrein. [†]We did not find a gene corresponding to HGFA in the zebra fish, but it does have one for HGF, which would need activation.

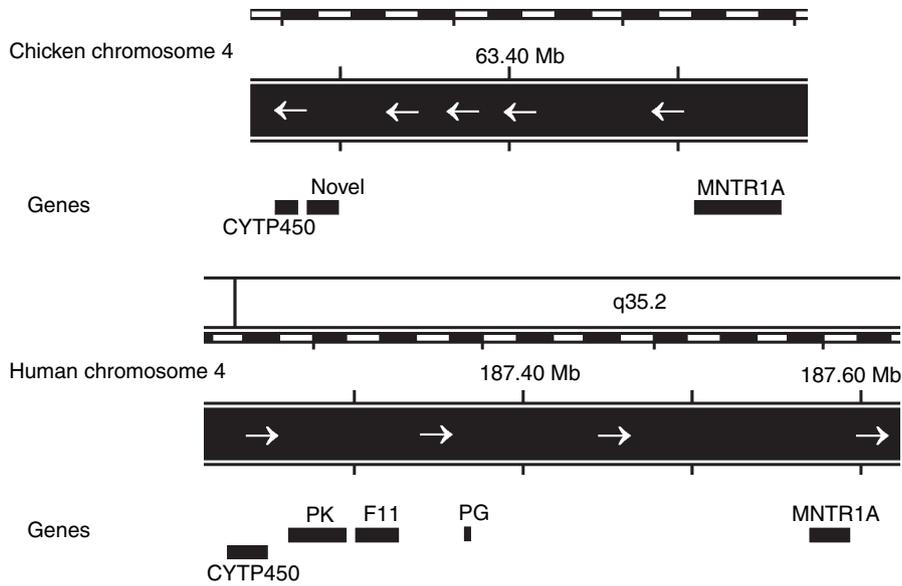


Fig. 4. Locations of genes neighboring factor XI-prekallikrein on human and chicken chromosomes (both coincidentally numbered 4). Genes: CYTP450, cytochrome P450; NOVEL, PK-fXI predecessor; PK, prekallikrein; F11, factor XI; PG, pseudogene; MNTR1A, melatonin receptor type 1A (slightly simplified from the EMBL-EBI website as viewed with ENSEMBLE).

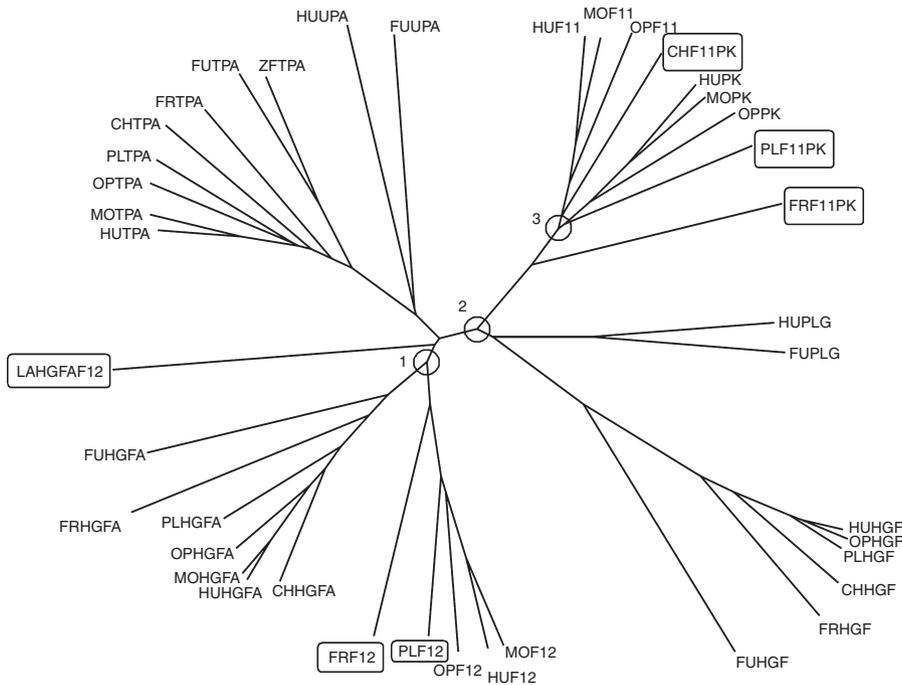


Fig. 5. Phylogenetic tree generated from the alignment of serine protease portions of various coagulation factors. In deference to the artifactually truncated platypus factor XI-prekallikrein predecessor, which lacks 13 residues at its carboxy-terminus, all sequences were shortened to the same point. The circled regions labeled 1, 2 predecessor, which lacks 13 residues at its carboxy-terminus, all sequences were shortened to the same point. The circled regions labeled 1, 2 and 3 correspond to key gene duplication sites. Trees with virtually the same topology were generated independently by other methods [20, 21]. Organism abbreviations: HU, human; MO, mouse; OP, opossum; PL, platypus; CH, chicken; FR, frog, ZF, zebra fish; FU, puffer fish; LA, lamprey. Factor abbreviations: PK, prekallikrein; F11, factor XI; F11PK, PK-fXI predecessor; F12, fXII; PLG, plasminogen; HGF, hepatocyte growth factor; HGFA, hepatocyte growth factor activator; TPA, tissue plasminogen activator; UPA, urokinase.

PAN domain forming the interchain interface [25,26]. Amino acids in human FXI, which are critical for forming the homodimer, are conserved in the opossum, including residues

Leu284, Ile290, and Tyr329 that form the interface, and an unpaired Cys321 that forms the interchain disulfide bond. Like human PK, opossum PK appears to be a monomer, with a

cysteine residue at position 326 that forms an intrachain disulfide bond with Cys321 [27–29].

Phylogenetic trees

Phylogenetic trees were calculated from a multiple alignment of the serine protease domains of numerous proteases, including some not involved in the contact phase of clotting (Fig. 5). In a perfect phylogenetic tree, predecessor genes are expected to appear prior to duplications leading to new proteins. If the duplication occurs within a short interval of a species divergence, however, the tree can be slightly muddled. In the case of the putative PK–FXI predecessor, the frog sequence appears well in advance of the duplication, as expected, but the chicken and platypus sequences are near the bottoms of the FXI and PK clusters, respectively. The internal branch lengths are very short, however, and either entry could be shifted on the tree merely by a few amino acid replacements.

At this point, it is not clear if the predecessor proteins are more like PK or FXI from placental mammals and the opossum. Indeed, the amino acid sequence of the protease domain of the platypus predecessor is about 68% identical with the corresponding regions of either human FXI or PK.

The platypus, frog, and chicken proteins all lack the putative FIX-binding site found in the FXI third PAN domain. An examination of sequences for the fourth PAN domain shows cysteine residues at positions 321 and 326 (human FXI numbering system; Fig. 6). This feature, found in PK but not FXI, indicates that the predecessor protein, like PK, is a monomer. These features notwithstanding, there is insufficient understanding of functional epitopes for mammalian PK proteins to determine if the predecessor proteins of frog, chicken or platypus are likely to have activity similar to PK. The results do strongly suggest, however, that the dimeric

structure of FXI is a feature acquired after duplication of the predecessor gene.

Discussion

The contact activation system of blood plasma consists of three serine protease zymogens (FXII, FXI, and PK), and the multidomain glycoprotein HK. HK is cleaved by activated PK (α -kallikrein) during contact activation, releasing the vasodilator and pro-inflammatory peptide BK.

In humans, the four components are required for surface-initiated formation of a blood clot in assays such as the partial thromboplastin time. Even so, deficiency of FXII, PK or HK is not associated with a predisposition to excessive bleeding, and abnormal hemostasis in FXI deficiency is variable and relatively mild [5]. As such, the importance of contact activation to normal hemostasis is questionable. However, thrombosis models using primates and genetically engineered mice, and epidemiologic studies in human populations suggest that the contact system may be involved in pathologic coagulation *in vivo* under some circumstances [6,7].

Our analysis shows that the expansion of the vertebrate clotting system to include FXI, PK, and FXII has occurred by way of a series of widely spaced gene duplications during the course of several hundred million years of evolution, beginning with the appearance of FXII and a PK–FXI predecessor in amphibians. Fortunately, a recent independent phylogenetic analysis of kininogens [15] has revealed that, although fish have a kininogen comparable to HK, the protein lacks the histidine-rich segment that is essential for co-factor activity in the contact system [30]. Indeed, that portion of the protein makes its first appearance at the level of amphibians [15], fully consistent and complementary with our findings.

Evolutionary appearance of FXII

Factor XII made its first appearance with the evolution of amphibians. Both lamprey and puffer fish have genes that are clearly orthologs of HGFA, but neither has FXII. The lamprey HGFA gene is especially noteworthy in that it appears on the phylogenetic tree well in advance of the gene duplication that divides the groups (Fig. 5). It is also worth commenting that the lamprey and other fish have hepatocyte growth factor itself, as well as its activator, although we should note in passing that we did not find an HGFA sequence in the zebra fish database, even though HGF is present (Table 1).

The fact that the putative frog FXII is a genuine ortholog of mammalian FXII is corroborated by the syntenic arrangement of neighboring genes in frog and mammals (Fig. 7). Thus, in humans, paralogous genes are situated on either side of human FXII and HGFA, even though they occur on different chromosomes, and the exact same situation occurs in frogs. In chickens, the gene arrangement is identical except that FXII is missing. A duplication of a chromosomal segment followed by a translocation is clearly indicated in the evolutionary introduction of FXII.

Opossum FXI	MC	IRDIFPNTVF	AD
Human FXI	AC	IRDIFPNTVF	AD
Chimpanzee FXI	AC	IRDIFPNTVF	AD
Rhesus monkey FXI	AC	IRDVFPNTVF	AD
Mouse FXI	AC	IRDIFPNTVL	AD
Rat FXI	AC	IRDIFPNTVL	AD
Rabbit FXI	AC	IRDIFPSTVF	AD
Cattle FXI	AC	IRDIFPRTAF	VD
Horse FXI	AC	IRDIFPSTVF	AD
Dog FXI	AC	IRDIFPSTVF	AD
Frog Predecessor	GC	QNDLFQNMEL	PG
Chicken Predecessor	DC	QMDIFEHEEF	SG
Platypus Predecessor	FC	HPSLYSDTDF	LG
Human PK	GC	HMNIFQHLAF	SD

Fig. 6. Alignment of factor IX-binding regions of various mammalian factors XI [based on ref. 24] and corresponding regions of human PK and putative predecessor proteins from frog, chicken and platypus.

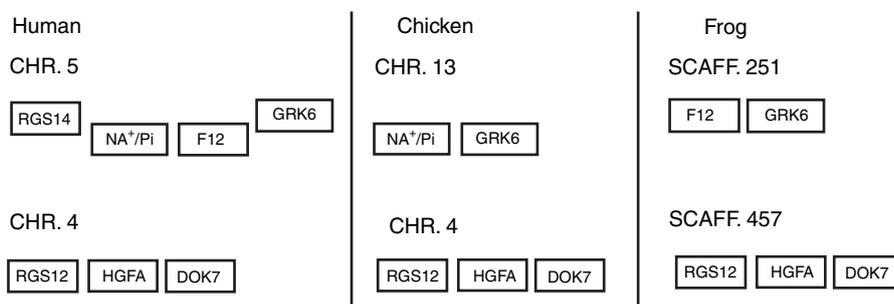


Fig. 7. Arrangement of neighboring genes for factor FXII and/or HGFA in human, chicken (HGFA only) and frog. Genes: RGS14, regulator of G-protein signaling 14; Na⁺/Pi, sodium-dependent phosphate transport protein 2A; F12, fXII; GRK6, G protein-coupled receptor kinase 6; RGS12, regulator of G-protein signaling 12; HGFA, hepatocyte growth factor activator; DOK7, protein Dok-7 (downstream of tyrosine kinase 7).

In humans, the FXII gene is sandwiched between genes denoted Na⁺/Pi and GRK6. In chickens, these two genes are adjacent to each other with no gene in between, but in frogs, the FXII gene is adjacent to GRK6. As it happens, the FXII gene in frogs is at the end of a scaffold and its upstream neighbor cannot yet be identified. Interestingly, in all three genomes (frog, chicken, and human) the arrangement around the HGFA gene is identical (Fig. 7). Loss of the FXII gene on the lineage leading to birds is also consistent with the topology of the phylogenetic tree (Fig. 5).

Evolutionary appearance of PK–FXI predecessor

As with FXII, the predecessor of mammalian FXI and PK first appears in amphibians. The most notable structural feature of FXI, PK, and their predecessor is a series of four disulfide-constrained 90–91 amino acid repeats initially called apple domains. Apple domains are now properly considered members of the PAN domain family. Patthy *et al.* [31] have written extensively on the evolutionary origins of the apple/PAN domains of FXI and PK, and have provided convincing evidence that they are derived from the N-terminal domain of plasminogen and hepatocyte growth factor (HGF; Fig. 2). While the domain is widely scattered among various animal proteins, including sundry tandem repeats in nematode adhesive proteins, the most likely ancestors for FXI and PK are serine proteases of the plasminogen/HGF sort (Fig. 5). In addition to the frog and chicken, the PK–FXI predecessor is present in the platypus, one of five extant species of monotremes (egg-laying mammals) that diverged from the lineage leading to marsupial and placental mammals relatively early in mammalian evolution.

Separate genes in the opossum

The observation that distinct genes for PK and FXI are present in the opossum indicates they are the result of a gene duplication involving the predecessor gene that occurred after the divergence of monotremes from other mammals. PK and FXI in placental mammals and the opossum have structural features that must account for specific functions, and it is tempting to speculate about the functional properties of the

predecessor based on a comparison of amino acid sequences. Factor XI is a dimeric protein, an unusual conformation for a trypsin-like serine protease. Key residues in the fourth PAN domain involved in FXI dimer formation are not conserved in PK or the PK–FXI predecessor. Furthermore, Cys³²¹, which forms an interchain disulfide bond in FXI, forms an intrachain bond with Cys³²⁶ in PK and the predecessor proteins of frog, chicken, and platypus. Finally, sequences in the FXI third PAN domain that are required for binding to FIX and to platelets are not conserved in PK or the predecessors. Taken as a whole, the data suggest that the PK–FXI predecessor would lack the ability to function in coagulation in a similar manner to FXI (i.e. to activate FIX). This is consistent with the observation that FXI activity is not found in the plasmas of birds [8–12]. It must be noted, however, that studies examining coagulant activities in bird plasma were based on assays using human factor-deficient plasmas [10], and failure to identify a specific activity may be related to species incompatibility rather than true absence. Despite the similarities with PK, it is not known at this point if the predecessor has kininogenase activity. Experiments with wholly homologous systems (all components from the same organism) should provide the answer.

About losing genes

It is well known that cetaceans (whales, dolphins, and porpoises) are deficient in FXII, and yet these creatures seem perfectly well adapted in an evolutionary sense [32]. In this case, the history of events leading to the loss of activity is made apparent by the presence of a corresponding pseudogene [33]. In the case of the chicken, the gene is completely missing.

The determination of whole genome sequences during the past decade has shown that gene loss is common in the biological world. For example, almost 400 genes have been lost along each of the lineages leading from the common ancestor of fission and budding yeasts [34]. Seen in this light, the loss of the FXII gene along the lineage leading to birds, or its conversion to an inactive pseudogene in cetaceans, is not necessarily a cataclysmic event.

Interestingly, these two cases of FXII loss bear on the questionable occurrence of the anomalous terminator in the platypus PK–FXI predecessor. It might be argued that

the terminator codon merely indicates that the predecessor gene has been lost. Natural selection is necessarily and immediately relaxed on a non-functioning gene, however, and the coding sequence quickly deteriorates. In the case of the whale FXII pseudogene, two terminator codons are immediately followed by a single-nucleotide insertion that changes the remainder of the reading frame to a nonsensical one with multiple terminators. All of this decay had to occur within the 60 million year interval since cetaceans diverged from other mammals [35]. In the case of FXII loss from chickens, the gene has vanished completely during the 300 million year period since the divergence of birds, as shown by its absence between syntenic genes in other vertebrate classes (Fig. 7). What these observations tell us is, if the reported terminator codon in the platypus PK–FXI gene is genuine, the mutation giving rise to it must have occurred very recently, and the gene must have been active for most of the time since monotremes diverged from other mammals 166 million years ago [22]. In other words, even allowing for the small population size of the Australian platypus, there would doubtless be a ‘mutant’ specimen where the terminator codon has been replaced with a tryptophan codon, thereby resulting in an active gene.

Adding complexity

Gene duplications lie at the heart of adding complexity to any genetic system. Often such duplications merely enhance the fine-tuning and balance of an already highly regulated system, and it is difficult to determine what natural advantage was gained. In this regard, the independent appearances of the PK–FXI predecessor, FXII, and the histidine-rich co-factor section of HK at the dawn of tetrapods is not likely mere coincidence (Fig. 8). Indeed, the fact that HGFA and the PK–FXI predecessor gene are related to about the same degree as HGFA and FXII (Fig. 5) strongly suggests these events were coupled in

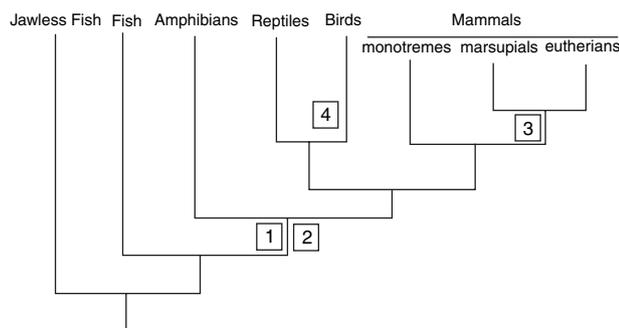


Fig. 8. Diagrammatic depiction of vertebrate evolutionary tree showing proposed history of gene duplications and gene loss giving rise to observed distribution of contact phase coagulation factors. **1.** Duplication of HGF- or plasminogen- related protease and tandem duplication of PAN domains gives rise to evolutionary predecessor of PK and FXI. **2.** Duplication of gene for HGFA gives rise to fXII. **3.** Duplication of gene for PK–FXI predecessor gives rise to separate genes for factors XI and PK. **4.** Loss of fXII gene on lineage leading to birds (the loss could have occurred before the split leading to reptiles and birds, a matter that should be resolved when a whole genome sequence for a reptile becomes available).

time. Both HGFA and HGF, the putative ancestors of the contact proteases, appear to be virtually ubiquitous among vertebrates, as had been predicted [36]. Something about the evolutionary migration of vertebrates from a strictly aqueous environment to land is likely associated with these new genetic acquisitions, but it would be premature to speculate on what that may have been.

Meanwhile, it will be of considerable interest to find if, *in vitro*, in a wholly homologous system composed of components from frog plasma, the PK–FXI predecessor is activated by frog FXIIa and, if so, whether the activated predecessor protein can activate frog HK, FIX, or both. Similarly, it will be useful some day to study the contact phase of clotting in the platypus. It must be emphasized that such studies need be conducted on strictly homologous systems, and biological material from key organisms such as the platypus is not always readily available.

Acknowledgement

M. Ponczek was the recipient of an EMBO short-term fellowship.

Disclosure of Conflict of Interests

The authors state that they have no conflict of interest.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Sequence alignment of serine protease portions of contact phase coagulation factors and various paralogs. Proteins: FXI, factor XI; FXII, factor XII; PREK, prekallikrein; HGFA, hepatocyte growth factor activator; PG, plasminogen; TPA, tissue plasminogen activator. Organisms: HU, human; OP, opossum; PL, platypus; CH, chicken; FR, frog; PF, puffer fish.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

- Davidson CJ, Hirt RP, Lal K, Elgar G, Tuddenham EGD, McVey JH. Molecular evolution of the vertebrate blood coagulation network. *Thromb Haemost* 2003; **89**: 420–8.
- Davidson CJ, Tuddenham EG, McVey JH. 450 million years of hemostasis. *J Thromb Haemost* 2003; **1**: 1487–94.
- Doolittle RF, Jiang Y, Nand J. Genomic evidence for a simpler clotting scheme in jawless vertebrates. *J Mol Evol* 2008; **66**: 185–96.
- Jiang Y, Doolittle RF. The evolution of vertebrate blood coagulation as viewed from a comparison of puffer fish and sea squirt genomes. *Proc Natl Acad Sci U S A* 2003; **100**: 7527–32.
- Gailani D, Broze G. FXI and the contact system. In: Scriver C, Beaudet A, Sly W, Valle D, Childs B, Kinzler K, Vogelstein B, eds.

- Metabolic and Molecular Basis of Inherited Disease*. New York: McGraw-Hill, 2001: 4433–53.
- 6 Renne T, Pozgajova M, Gruner S, Schuh K, Pauer HU, Burfeind P, Gailani D, Nieswandt B. Defective thrombus formation in mice lacking coagulation FXII. *J Exp Med* 2005; **202**: 271–81.
 - 7 Renne T, Gailani D. Role of fXII in hemostasis and thrombosis: clinical implications. *Expert Rev Cardiovasc Ther* 2007; **5**: 733–41.
 - 8 Delezenne C. Recherches sur la coagulation du sang chez les oiseaux. *Arch Physiol Norm Path*, 1897; **29**: 333–77.
 - 9 Didisheim P, Hattori K, Lewis JH. Hematologic and coagulation studies in various animal species. *J Lab Clin Med* 1959; **53**: 866–75.
 - 10 Soulier J-P, Wartelle O, Menache D. Hageman Trait and PTA deficiency; the role of contact of blood with glass. *Br J Haematol* 1959; **5**: 121–38.
 - 11 Weir-M J, Acurero Z, Salas-A R, Arteaga-Vizcaino M. Blood coagulation factors in the black headed vulture (*Coragyps atratus*), a potential animal model for the study of haemostasis. *Thromb Res* 2004; **113**: 269–73.
 - 12 Frost CL, Naudé RJ, Oelofsen W, Jacobson B. Comparative blood coagulation studies in the ostrich. *Immunopharmacology* 1999; **45**: 75–81.
 - 13 Doolittle RF, Feng DF. Reconstructing the evolution of vertebrate blood coagulation from a consideration of the amino acid sequences of clotting proteins. *Cold Spring Harb Symp Quant Biol* 1987; **52**: 869–74.
 - 14 Miyazawa M, Shimomura T, Kitamura A, Kondo J, Morimoto Y, Kitamura N. Molecular cloning and sequence analysis of the cDNA for a human serine protease responsible for activation of hepatocyte growth factor. Structural similarity of the protease precursor to blood coagulation fXII. *J Biol Chem* 1993; **268**: 10024–8.
 - 15 Zhou L, Li-Ling J, Huang H, Fei M, Li Q. Phylogenetic analysis of vertebrate kininogen genes. *Genomics* 2008; **91**: 129–41.
 - 16 Birney E, Andrews TD, Bevan P, Caccamo M, Chen Y, Clarke L, Coates G, Cuff J, Curwen V, Cutts T, Down T, Eyra E, Fernandez-Suarez XM, Gane P, Gibbins B, Gilbert J, Hammond M, Hotz H, Iyer V, Jekosch K *et al.* An overview of ENSEMBLE. *Genome Res* 2004; **14**: 925–8.
 - 17 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389–402.
 - 18 Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. *J Mol Biol* 1997; **268**: 78–94.
 - 19 Thompson JD, Higgins DG, Gibson TJ. Clustal W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**: 4673–80.
 - 20 Feng D-F, Doolittle RF. Progressive alignment and phylogenetic tree construction of protein sequences. *Meth Enzymol* 1996; **266**: 368–72.
 - 21 Doolittle RF, Feng D-F. Nearest neighbor procedure for relating progressively aligned amino acid sequences. *Meth Enzymol* 1990; **183**: 659–69.
 - 22 Warren WC, Hillier LW, Marshall Graves JA, Birney E, Ponting CP, Grützner F, Belov K, Miller W, Clarke L, Chinwalla AT, Yang S, Heger A, Locke DP, Miethke P, Waters PD, Veyrunes F, Fulton L, Fulton B, Graves T, Wallis J *et al.* Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 2008; **453**: 175–83.
 - 23 International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 2004; **432**: 695–716.
 - 24 Sun MF, Zhao M, Gailani D. Identification of amino acids in the fXI apple 3 domain required for activation of factor IX. *J Biol Chem* 1999; **274**: 36373–8.
 - 25 Bouma BN, Griffin JH. Human blood coagulation factor XI, purification, properties, and mechanism of activation by activated fXII. *J Biol Chem* 1977; **252**: 6432–7.
 - 26 Papagrigoriou E, McEwan PA, Walsh PN, Emsley J. Crystal structure of the fXI zymogen reveals a pathway for transactivation. *Nat Struct Mol Biol* 2006; **13**: 557–8.
 - 27 Wu W, Sinha D, Shikov S, Yip CK, Walz T, Billings PC, Lear JD, Walsh PN. FXI homodimer structure is essential for normal proteolytic activation by fXIIa, thrombin and factor XIa. *J Biol Chem* 2008; **283**: 18655–64.
 - 28 Meijers JC, Mulvihill ER, Davie EW, Chung DW. Apple 4 in human blood coagulation fXI mediates dimer formation. *Biochemistry* 1992; **31**: 4680–4.
 - 29 McMullen BA, Fujikawa K, Davie EW. Location of the disulfide bonds in human plasma prekallikrein: presence of four novel apple domains in the amino-terminal portion of the molecule. *Biochemistry* 1991; **30**: 2050–6.
 - 30 Waldmann R, Scicli AG, McGregor RK, Carretero OA, Abraham JP, Kato H, Han YN, Iwanaga S. Effect of bovine high molecular weight kininogen and its fragments on Fitzgerald trait plasma. *Thromb Res* 1976; **8**: 785–95.
 - 31 Tordai H, Banyai L, Patthy L. The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins. *FEBS Lett* 1999; **461**: 63–7.
 - 32 Robinson AJ, Kropatkin M, Aggeler PM. Hageman factor (fXII) deficiency in marine mammals. *Science* 1969; **166**: 1420–2.
 - 33 Semba U, Shibuya Y, Okabe H, Yamamoto T. Whale Hageman factor (fXII) prevented production due to pseudogene conversion. *Thromb Res* 1998; **90**: 31–7.
 - 34 Aravind L, Watanabe H, Lipman DJ, Koonin EV. Lineage specific loss and divergence of functionally linked genes in eukaryotes. *Proc Natl Acad Sci USA* 2000; **97**: 11319–24.
 - 35 Bajpai S, Gingerich PD. A new Eocene archaeocete (Mammalia, Cetacea) from India and the time of origin of whales. *Proc Natl Acad Sci USA* 1998; **95**: 15464–8.
 - 36 Gherardi E, Manzano RG, Cottage A, Hawker K, Aparicio S. Evolution of plasminogen-related growth factors (HGF/SF and HGF1/MSP). In: Gherardi E, ed. *Plasminogen-related Growth Factors*. New York: John Wiley & Sons, 1997: 24–47.