Modular Evolution of the Blood clotting Cascade
– Project supervisor: Erich Bornberg-Bauer, N.N.

Background
Blood clotting is an essential and life saving feature of all higher vertebrates since, through forming fibrin clots, it can close wound upon contact with air only. The cascade is extant in a simpler form in jawless fish and in a more complex form in jawed vertebrates. It involves several serine proteases called coagulation factors, which activate other factors by peroteolytic cleavage. The factors can be grouped into mainly three different groups of domain arrangements: Gla-Kringle-Kringle-Trypsin, Gla-EGF-EGF-Trypsin and Copper_oxidase-Copper_oxidase-Copper_oxidase-dicoidin-discoidin. The Gla-EGF-EGF-Trypsin group is believed to have developed from the Gla-Kringle-Kringle-Trypsin group, which in turn originated in the classical complement system. The system is a classical example of how increasingly complex mechanisms evolved from simpler ones by duplication and fusion of proteins and recycling of their parts (domains).

Goal:
The goal of the project is to trace the evolution of the domain arrangement of each group and mainly bases on information gained from the Pfam database. While several studies exist, all of them are more than 15 years old and in the meanwhile many more genomes for comparative analysis have become available.

Methods:
The early stages of the project will comprise reading the literature and assembling an up-to-date collection of domain models and corresponding protein IDs.
In a second step, appropriate genomes will need to be collected which may represent possible intermediate evolutionary steps. In a third step, individual HMMs will need to be built and genomes scanned. Building on these data, a preliminary model of evolutionary duplications and re-arrangements shall be built. Finally, all alignments of domains will be investigated w.r.t. possible signals of adaptation with the goal to understand if, e.g. more recent acquisitions are still under selection pressure to integrate into the overall complex.

Required skills:
Basics in Bioinformatics, Basics in Molecular Evolution, Basics in Molecular Biology

Skills that can be learned:
Improvement of all the above, in particular advanced sequence analysis, phylogeny and a deeper understanding of molecular evolution of essential physiological systems in humans; basics of PYTHON scripting can and should be learned along the lines under supervision; usage of domain-based analyses such as RADS and DoMosaics.

Literature: