若手研究者インターナショナル・トレーニング・プログラム(ITP)

バイオインフォマティクスとシステムズバイオロジーの国際連携教育研究プログラム 応募書類

| |
|---|
| Name：David duVerle |
| Title：Ph.D. candidate, 1st year |
| Institute: Bioinformatics Center (Mamitsuka Laboratory), Kyoto University Institute for Chemical Research; Kyoto, Uji, Gokasho 611-0011, Japan |
| Partner institute: Max Planck Institute for Molecular Genetics (Vingron Department), Ihnestraße 63-73, 14195 Berlin, Germany |
| Duration: December 2009 - March 2009 (3 months) |

Plan:

*Background*

Under Prof. Mamitsuka's supervision, my research has focussed on improving methods for prediction of cleavage sites in substrate sequences by Calpain (a calcium-activated cysteine protease involved in muscular dystrophies, diabetes, and tumorigenesis in human, embryonic lethality in mouse, neurogenesis deficiency in fly, incomplete sex determination in nematodes, defect of aleurone cell development in maize, and alkaline/osmo-shock susceptibility in yeast). Among other sequence-based approaches, a method based on Position-Specific Scoring Matrix (PSSM) made use of Prof. Vingron's past work on improving such methods through the use of pseudo-counts.

Recently, we explored the possibility of using a Hidden Markov Model-based approach to represent candidate sequences and accurately predict possible Calpain substrates. However, the difficulty of formulating a Markovian model powerful enough to represent biological specificities, while still computationally practical, has so far prevented us from reaching satisfying results. We believe it might be possible to streamline the model through the use of higher-order information (such as secondary structure or amino-acid properties) or by reducing the complexity of the probabilistic framework.

*Goal*

Department Vingron's large emphasis on gene regulation matters, such as prediction of transcription factor binding sites (TFBS), makes it an ideal environment to study novel methods for sequence-based prediction. While these methods generally apply to different biological mechanisms and different support (DNA vs. protein sequence), it might be possible to draw some general techniques that can be reused in the case of cleavage site prediction by Calpain.