



## FBS Seminar

May 16 (Wed), 2018

5月16日(水)

16:00 - 17:00

生命システム棟 2F セミナー室

池村淑道 博士

長浜バイオ大学 客員教授: 国立遺伝学研究所/総研大 名誉教授

### AI(人工知能)に導かれたビッグデータ時代におけるゲノム研究

広範な生物種について多様なゲノム関係の情報が集積しビッグデータ化している。このビッグデータから新規性の高い知識発見をする上で、AIは力強い味方と言える。AIのなかでも「教師なし機械学習」は、予備知識やモデルや仮説を持たずに、知識発見をAIに任せる手法であり、想像もしなかった現象を見つけてくれる。我々のグループは十数年前より、教師なし機械学習のBLSOM（一括学習型自己組織化マップ）によりオリゴヌクレオチド頻度を解析し(1)、ゲノム配列に潜む多様な特徴を見いだして来た。オリゴヌクレオチドは、転写因子等の塩基配列特異的に結合するタンパク質類の結合モチーフであることから、機能的に興味深いゲノムの特徴を、BLSOMが視覚的に理解しやすい形式で表示してくれる。

今回はBLSOMが可能にする多様な解析例を紹介し、特に、セントロメア近傍のヘテロクロマチン領域に見られる、転写因子の結合配列やCpGを含むオリゴヌクレオチド類の集中の実態と、そのゲノム核内配置における役割を説明する(2と最近の研究)。応用的な研究例としては、エボラやインフルエンザのウイルスに見られる、方向性と再現性のあるゲノム配列の時系列変化を紹介する(3)。有効性が失われ難い核酸医薬や診断試薬のデザインに必須の知見を提供できる(4)。

(1) Informatics for unveiling hidden genome signatures. *Genome Res.* 13:693-702, 2003.

(2) Notable clustering of transcription-factor-binding motifs in human pericentric regions and its biological significance. *Chromosome Res.* 21, 461-474, 2013.

(3) Directional and reoccurring sequence change in zoonotic RNA virus genomes visualized by time-series word count. *Scientific Reports* 2016; 6: 36197.

(4) Time-series oligonucleotide count to assign antiviral siRNAs with long utility fit in the big data era. *Gene Therapy*, 24, 668-673, 2017.

世話人: 深川竜郎 ([tfukagawa@fbs.osaka-u.ac.jp](mailto:tfukagawa@fbs.osaka-u.ac.jp), 06-6879-4428)

セミナー前後に池村博士と個別に面談したい方は、深川までご連絡ください。