

## 平成29年度 学内研究助成金 研究報告書

研究種目	<input checked="" type="checkbox"/> 奨励研究助成金	<input type="checkbox"/> 研究成果刊行助成金
	<input type="checkbox"/> 21世紀研究開発奨励金 (共同研究助成金)	<input type="checkbox"/> 21世紀教育開発奨励金 (教育推進研究助成金)
研究課題名	ウイルス性心筋炎モデルのトランスクリプトームビッグデータを用いた心筋炎診断バイオマーカー同定パイプラインの構築	
研究者所属・氏名	研究代表者： 微生物学講座・尾村誠一 共同研究者：	

### 1. 研究目的・内容

突然死の主要原因の一つである心筋炎はほとんどがウイルス感染により引き起こされる。ウイルス性心筋炎は3つの病期に分かれることが提唱されており、それぞれの病期で治療法が異なるが、高感度かつ高精度に病期を判別できるマーカーは確立されていない。本研究ではウイルス性心筋炎マウスモデルのRNAシーケンシングデータを解析し、病期判別マーカーを同定することを目的として、そのデータ解析パイプラインの構築を行った。

### 2. 研究経過及び成果

RNAシーケンシングデータ解析は一部の高額なソフトウェアを除くと、Linux上でのみ動作するプログラムがほとんどであり、また大容量メモリを必要とする。そこでOSのインストールされていない32GBメモリ搭載PCを購入し、Linuxをインストールした。

RNAシーケンシングにより得られた生データ(リード断片データ)から解析可能な遺伝子発現データを抽出するには、以下の4つのステップにより構成されるパイプラインを構築する必要があった。A) 公開されているリファレンスゲノム配列上に対応する配列をもつリード断片データを配置するアラインメント(マッピング)。B) 配置されたリード断片データの数を数えるリードカウント。C) 遺伝子を同定するアノテーション。D) サンプル間および遺伝子間の比較を行うための基準を設定してデータを調整する正規化。これらのステップにはそれぞれ複数の手法、プログラム、データベースが存在したので、この中から目的に適した、容易に運用可能なプログラムを選定した。リファレンスゲノムにマッピングするプログラムは大きく分けて、スプライスジャンクションを含むリード断片をマッピングしないbasic alignerと、スプライスジャンクションを含むリード断片をマッピングできるsplice-aware alignerがある。マッピングの効率と動作速度、要求されるPCのスペックなどを考慮して、QuasR、Bowtie2など複数のプログラムを比較し、splice-aware alignerであるSpliced Transcripts Alignment to a References (STAR)が本研究に適すると判断した。遺伝子のアノテーションに用いるデータベースには、頻用されているものに米国University of California Santa Cruz (UCSC)により提供されているものと、ヨーロッパのEuropean Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI)により提供されているものがあるが、両者を比較した結果、EMBL-EBIのデータベースを用いた場合には多くの抗体可変部やマイクロRNAが含まれているのに対して、UCSCのデータベースではほとんど含まれていなかったため、EMBL-EBIデータベースを用いてアノテーションを行うこととした(Omura et al., Neuroinfection, 2018)。データ正規化の方法としては、サンプル中の総リード数を100万リードに統一し、全ての遺伝子の長さを1000 baseに調整するReads per kilobase of exon per million mapped reads (RPKM)が頻用されているが、本研究においては全てのサンプルでリード数が一定であるとは考えにくいので、発現が変化しない遺伝子のデータからサンプル間の正規化を行うiDEGES/edgeRを用いた。得られるデータを用いて主成分分析、因子分析などのバイオインフォマティクス解析を行う環境も整えた(Omura et al., Arch Virol, 2018; Sato et al., Sci Rep, 2017)。

予備的解析として、マウス脊髄のRNAシーケンシングデータを本研究で構築したパイプラインを用いて解析したが、全リードの約80%がリファレンスゲノムにマップされ、抗体可変部やT細胞受容体可変部等を含むカウントデータが得られた。

### 3. 本研究と関連した今後の研究計画

今後は実際にウイルス性心筋炎マウスモデルを作製し、各病期で解剖して心筋および血液を採取する。RNA を抽出して RNA シーケンシングを行い、本研究において構築した解析パイプラインを用いて遺伝子発現データを抽出する。得られたデータを用いてバイオインフォマティクス解析を行い、病期判別バイオマーカーおよび血中代替マーカーの同定を行う。

さらに、バイオマーカー同定までをパイプラインに含められるよう検討し、他の疾患のバイオマーカー同定への応用を試みる。

### 4. 成果の発表等

発表機関名	種類 (著書・雑誌・口頭)	発表年月日(予定を含む)
Neuroinfection (神経感染症学会)	学術論文	2018年4月30日
Archives of Virology	学術論文	2018年1月23日
Scientific Reports	学術論文	2017年9月5日
Clinical and Experimental Neuroimmunology	総説	2017年7月23日
第65回日本ウイルス学会年会	ポスター発表	2017年10月24 - 26日
第22回日本神経感染症学会学術大会	口頭発表	2017年10月13 - 14日
第6回生命医薬情報学連合大会	ポスター発表	2017年9月27 - 29日
XXIII World Congress of Neurology	ポスター発表	2017年9月16 - 21日
7 <sup>th</sup> Sendai Conference 2017	招待講演	2017年7月8日