氏　　名	伊東 潤平
学位(専攻分野)	博士(理学)
学 位 記 番 号	総研大甲第 2009 号
学位授与の日付	平成 30 年 3 月 23 日
学位授与の要件	生命科学研究科 遺伝学専攻 学位規則第6条第1項該当
学 位 論 文 題 目	Systematic identification of functional elements derived from human endogenous retroviruses
論 文 審 査 委 員	主　査　教授 斎藤 成也 教授 角谷 徹仁 教授 黒川 顕 准教授 池尾 一穂 教授 鈴木 善幸 名古屋市立大学 大学院システム自然科学研究科

(別紙様式 2)
(Separate Form 2)

論文の要旨
Summary (Abstract) of doctoral thesis contents

Human endogenous retrovirus (HERV) belongs to a class of transposable elements (TEs) and occupies approximately 8% of the entire human genome. Although HERVs were initially thought to be non-functional and merely parasitic sequences in the genome, evidences have been accumulated that the human genome carries enormous HERV-derived functional elements, which affect host physiologies and diseases. These elements play a role in the DNA level (i.e., transcriptional regulatory elements), RNA level (i.e., non-coding RNA), or protein level. Recent efforts using high-throughput sequencing have generated a massive amount of genomic, epigenomic, and transcriptomic data. In this doctoral thesis, I aimed to identify HERV-derived functional elements (particularly regulatory elements and transcripts) by reanalyzing epigenomic and transcriptomic sequencing data accumulated in the public databases.

In Chapter 1 (general introduction), I describe basic knowledge about HERVs and recent progresses that have showed associations of HERV-derived functional elements with host physiologies and diseases.

In Chapter 2, I investigated HERV-derived regulatory elements using 519 ChIP-Seq data for 97 transcription factors (TFs) provided by ENCODE and Roadmap Epigenomics. I identified 794,972 TF-binding events on HERVs and 2,201 specific HERV-TF associations. Using unsupervised clustering analysis, I demonstrated that HERVs could be grouped according to TF binding patterns: HERV groups bound by pluripotent TFs (e.g., SOX2, POU5F1, and NANOG), embryonic endoderm/mesendoderm TFs (e.g., GATA4/6, SOX17, and FOXA1/2), hematopoietic TFs (e.g., SPI1 (PU1), GATA1/2, and TAL1), and CTCF were identified. By analyzing the three-dimensional chromosomal interactions, I demonstrated that HERV-derived regulatory elements tend to interact with host genes relating to the innate immune response. This suggests that the HERV-derived regulatory elements play a role in the modulation of this biological pathway. We further demonstrated heterogeneities of regulatory elements within LTR7 group: SOX2, POU5F1, and KLF4-binding sites were highly enriched in the youngest subgroup of LTR7, which had the highest transcriptional activity in pluripotent cells. This suggests that the subgroup acquired those regulatory activities for efficient replication in the host germ cells. Furthermore, my colleagues and I constructed dbHERV-REs (http://herv-tfbs.com/), a database of HERV-derived regulatory elements.

In Chapter 3, I investigated HERV-derived transcripts in tumors by reanalyzing RNA-Seq data of 5,550 patients across 12 solid tumors provided by TCGA. I identified 10,060 transcribed HERV loci in tumors and the corresponding normal tissues. In nine out of 12 tumor types, the overall transcription levels of HERVs significantly increased. Particularly, transcription levels of HERVH group were highly up-regulated in a broad range of tumors. In unsupervised clustering analysis based on the HERV transcriptome, RNA-Seq samples clustered according to tumor and tissue types, and even molecular subtypes within a type of tumors. This indicates that HERVs had unique transcription profiles among tumor/tissue types and the subtypes. The

transcriptionally up-regulated HERVs in tumors were associated with TFs that were overexpressed in the tumors. In case of breast cancer, the up-regulated HERVs tended to be bound by ESR1 (estrogen receptor 1), PGR (progesterone receptor), GATA3, and FOXA1, which were overexpressed in the cancer type. Furthermore, transcription levels of HERVs in tumors were positively correlated with those of genes targeted by ZNF274, TRIM28, and/or SETDB1. These are known to form a protein complex and suppress the transcription of TEs in mouse embryos. This result suggests that these genes work on the transcriptional silencing of HERVs also in human tumors. Some HERVs were transcribed as parts of mRNA of genes and contributed to produce non-canonical transcripts of those genes. For example, the fused transcript of ERVL-B4-int and TMPRSS4, a major causal gene of prostate cancer, was highly up-regulated in prostate adenocarcinoma. The ERVL-B4-int locus worked as an alternative transcription start site of TMPRSS4 and contributed to the overexpression of this gene in the tumors.

This doctoral thesis depicts the landscape of HERV-derived functional elements providing insights into effects of these elements on host physiologies and diseases.

(別紙様式 3)
(Separate Form 3)

博士論文審査結果の要旨

Summary of the results of the doctoral thesis screening

可動因子は、逆転写酵素を用いてRNAを介しコピーを増やすレトロトランスポゾンと自身が移動するDNAトランスポゾンに大別される。レトロトランスポゾンには、LTRタイプ、LINE、SINEの３種類が存在する。

出願者は、ヒトゲノム中に内在しているLTRタイプのレトロトランスポゾンであるレトロウイルス（HERV）の塩基配列について、網羅的な解析をおこなった。HERV配列はヒトゲノムのおよそ８％を占めている繰り返し配列であり、かつてはまったく機能がないと思われていた。しかし近年いくつかのHERVは機能を持つことがわかってきた。

出願者は、ENCODE計画とRoadmap Epigenomics計画で97転写因子(TF)にたいして決定された519種類のChip-Seqデータから、HERV上に約80万個のTF結合イベントと2200個のHERVとTFについての特異的結合様式を発見した。これらのTF結合パターンから、HERVは次の４種類に分類された：多能性TF結合タイプ、内胚葉・中胚葉TF結合タイプ、造血幹細胞TF結合タイプ、CTCF結合タイプ。

さらに、染色体の三次元相互作用情報解析から、HERV由来の調節配列は、内在性免疫応答に関連した遺伝子の発現にかかわる傾向があることを明らかにした。また、出願者は共同研究者とともに、これらの解析結果を閲覧できるデータベースdbHERV-Resを開発した。

出願者はTCGAプロジェクトが提供している12種類の腫瘍の計5550人の患者サンプル由来のRNA-Seqデータも分析した。９種類の腫瘍において、HREVの転写レベルは有意に上昇しており、またHREVのRNA-Seqデータは腫瘍と組織のタイプに対応していた。

出願者によるこれら一連の解析結果から、ヒトゲノムにおける HERV 配列に関して新しい知見が得られた。したがって、審査委員会は全員一致して、出願者が博士（理学）の学位を取得することを承認した。