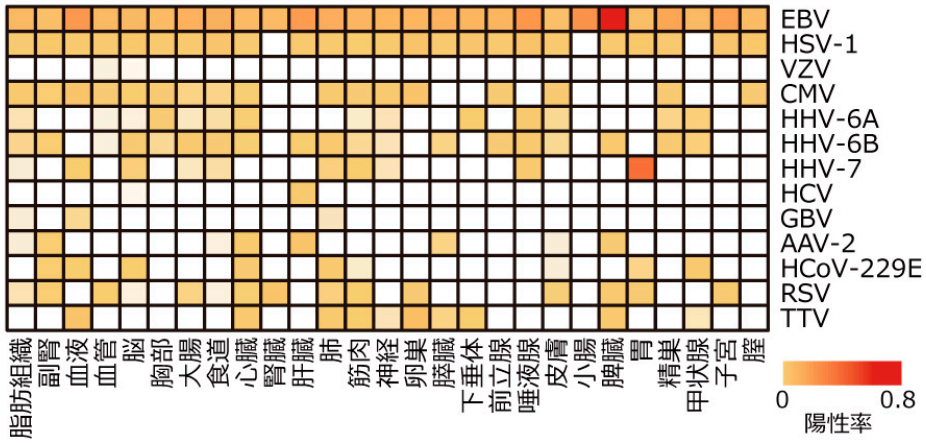
199　全ウイルスの網羅的検出法の開発と疾患との関連解析	佐藤　佳

【目的】多因子疾患や生活習慣病は、環境要因と遺伝要因が相加・相乗的に作用することにより発症する。多因子疾患の多くが、免疫・炎症の異常に関連があり、またこれらの異常は、ウイルス感染によって惹起されうる。そのため、ウイルス感染が疾患のリスクファクターの一つとなる可能性は充分に想定される。たとえば、RS ウイルス感染が、多因子疾患の一つである小児喘息と関連があることが指摘されている。他方、「衛生仮説（hygiene hypothesis）」で提唱されるように、ウイルスの不顕性感染が、疾患の発症を抑制する因子となる可能性も考えられる。実際に、ガンマヘルペスウイルス感染マウスモデルにおいて、このウイルスの感染により、気管支喘息の発症が抑制されることが示されている。このように、常在するウイルスの感染がヒトの健康に影響を与えている可能性は十分に考えられるが、常在ウイルスの解析のために健常人に侵襲的操作を加えることは倫理的にも極めて困難であり、解析が進んでいなかった。そこで本研究では、米国ブロード研究所が主導する Genotype-Tissue Expression（GTEx）プロジェクトに着目し、これまでは見出されなかった疾患の要因としてのウイルス感染の可能性を網羅的に検出することを目的とした。

【方法】GTEx には、検死検体などから取得された、547 人の 51 種類の組織からなる合計 8,991 サンプルの RNA シークエンス（RNA-seq）データが含まれている。もし、その検体のいずれかにウイルスが感染しているのであれば、その RNA-seq データの中にもウイルス由来の配列が含まれていることが想定される。そこで、RNA-seq データから 5,139 種類のウイルス由来の配列を識別し、定量化するためのパイプラインを構築し、GTEx に含まれる RNA-seq データ全てに対して網羅的にメタトランスクリプトーム解析を実施した。

【結果】GTEx が提供する計 8,991 サンプルの RNA-seq データの大規模メタ解析により、39 種類のウイルスが、ヒトのさまざまな組織に常在していることを明らかにした。また、ヒト遺伝子発現の関連解析により、ウイルスの感染が、ヒト遺伝子発現パターンに影響を与えていることを示唆する結果を得、研究成果を論文として発表した（Kumata et al., *BMC Biol*, 2020）。

描出した健常人のヴァイローム



199. 全ウイルスの網羅的検出法の開発と疾患との関連解析

佐藤　佳

東京大学 医科学研究所 感染症国際研究センター システムウイルス学分野

Key words：ヴァイローム，ウイルス叢，次世代シークエンス

緒　言

疾患の環境要因のひとつであるウイルスは、その元来の病原性から、感染症学的には、重症化し死に至る可能性がある高病原性ウイルス（エボラウイルスなど）、一過的な病態を示した後に自然寛解する低病原性のウイルス（RS ウイルスなど）、健常人にはほとんど病原性を示さない不顕性ウイルス（EB ウイルスなど）に大別される。しかし、感染病態の発現には、ウイルス元来の病原性に加え、感染者の免疫力や体質、すなわち、ヒトの形質を形成する遺伝要因も密接に関与することに留意すべきである。

多因子疾患や生活習慣病は、環境要因と遺伝要因が相加・相乗的に作用することにより発症する。多因子疾患の多くが、免疫・炎症の異常に関連があり、またこれらの異常は、ウイルス感染によって惹起されうる。そのため、ウイルス感染が疾患のリスクファクターのひとつとなる可能性は充分に想定される。たとえば、RS ウイルス感染が、多因子疾患のひとつである小児喘息と関連があることが指摘されている［2］。他方、「衛生仮説（hygiene hypothesis）」［3］で提唱されるように、ウイルスの不顕性感染が、疾患の発症を抑制する因子となる可能性も考えられる。実際に、ガンマヘルペスウイルス感染マウスモデルにおいて、このウイルスの感染により、気管支喘息の発症が抑制されることが示されている［4］。

我々は、ヒト体内に常在するウイルスを網羅的に同定する（ヒト体内のヴァイローム［virome：全ウイルス情報］を明らかにする）ことを目的として、GTEx（Genotype-Tissue Expression project：米国ブロード研究所をはじめとする国際コンソーシアムが提供する、ヒトの体組織ごとの遺伝子発現と遺伝的多型を網羅的に調べたプロジェクト）が提供する、ヒト体組織のトランスクリプトーム情報を用いた解析を行った。具体的には、RefSeq に登録された 5,139 種類のウイルスをクエリーとして、549 人の 53 か所の体組織のトランスクリプトーム情報から、ウイルス配列を抽出した。すなわち、次世代シークエンスデータを用い、ヒトのさまざまな組織におけるウイルスの存在を描出できる「ヒトヴァイローム解析パイプライン（Human Virome Analysis Pipeline：HVAP）」の構築を進めている。本研究では、HVAP に加え、ヴァイローム情報を既存の多階層オミクス情報に重層することで、これまでは見出されなかった疾患の要因としてのウイルス感染の可能性を網羅的に検出することを目的とした。

方　法

1. HVAP パイプラインの構築とバイオインフォマティクス解析

健常人のさまざまな体組織における「ヴァイローム（virome）」を網羅的に明らかにするために、我々は、健常人のさまざまな体組織における遺伝子発現が登録されている、米国ブロード研究所が主導する Genotype-Tissue Expression（GTEx）プロジェクト［5］に着目した。GTEx には、検死検体などから取得された、547 人の 51 種類の組織からなる合計 8,991 サンプルの RNA シークエンス（RNA-seq）データが含まれている。もし、その検体のいずれかにウイルスが感染しているのであれば、その RNA-seq データの中にもウイルス由来の配列が含まれていることが想定される。そこで、RNA-seq データから 5,139 種類のウイルス由来の配列を識別し、定量化するためのパイプラインを構築し、GTEx に含まれる RNA-seq データ全てに対して網羅的にメタトランスクリプトーム解析を実施した。

結　果

1. HVAP の構築とヒトヴァイロームの描出

GTEx に含まれる RNA-seq データ全てに対して網羅的にメタトランスクリプトーム解析を実施した（図 1A）。我々が開発したパイプラインを用いることにより、同一の RNA-seq データに含まれるヒト遺伝子発現とウイルス量の両方を定量化し、それらの関連解析を行うことが可能になった。また、本研究では、効率的に多量のデータを処理するために、東京大学医科学研究所のスーパーコンピューターSHIROKANE を活用した。解析の結果、少なくとも 39 種類のウイルスがヒトに常態的に感染していることを明らかにした（図 1B）（ヒトに感染する代表的なウイルスのみ記載）[1]。検出されたウイルスの中には、すでにヒトに感染することが知られているウイルスの他に、ヒト以外の脊椎動物や、昆虫、植物等に感染するウイルスも含まれていた。また、EBV やヒトパピローマウイルス（HPV）など、全身のさまざまな組織からウイルス由来の配列が検出されたウイルスがある一方で、C 型肝炎の原因ウイルスである C 型肝炎ウイルス（HCV）やヒトヘルペスウイルス 7 型（HHV-7）は、組織特異的にウイルス由来の配列が検出された（後述）。

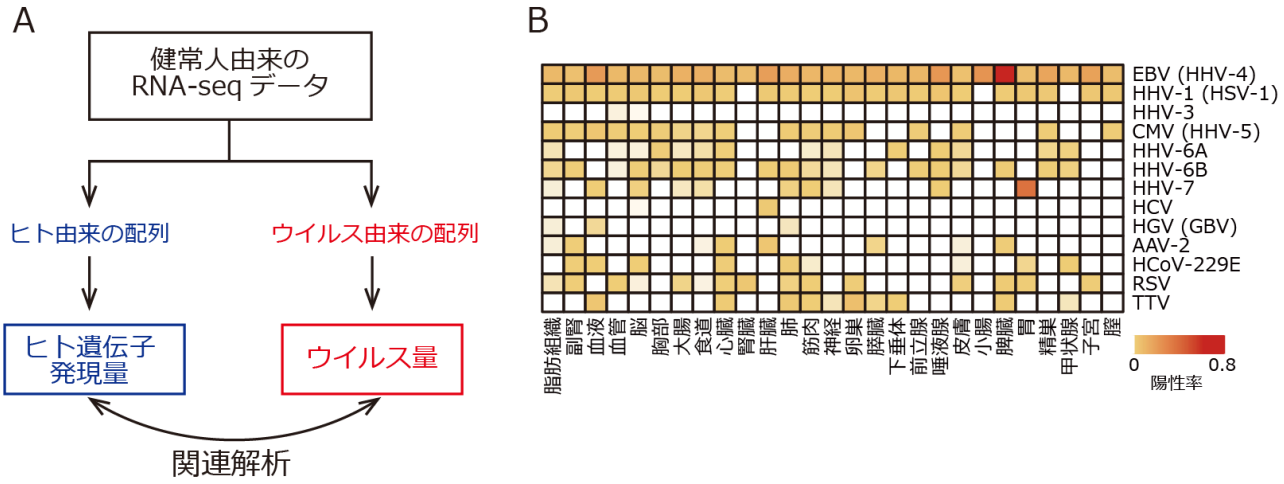


図 1. ヒトヴァイローム（virome）解析パイプラインとその結果

A) ヒトヴァイローム(virome)解析パイプライン。GTEx データに登録されている RNA-seq データからヒト由来の遺伝子配列（青）とウイルス由来の遺伝子配列（赤）をそれぞれ抽出し、それぞれの体組織におけるヒト遺伝子の発現プロファイルの描出と、ウイルス量の定量を行った。そして、両者の関連解析を行うことで、ウイルス感染によって変化するヒト遺伝子の発現変化を解析した。

B) ヒト体組織のヴァイローム（virome）解析結果 2)。各ウイルスの組織ごとの陽性率をヒートマップで示した。本研究 2）で同定した 39 種類のウイルスのうち、代表的なヒトに感染するウイルスを抜粋して示した。

2. C 型肝炎ウイルス感染とインターフェロン応答

まず、健常人の体内に存在するウイルスが、ヒトの免疫に関連する遺伝子の発現にどのような影響を与えているのかを調べるために、ヒト遺伝子発現とウイルス由来の配列の有無の関連解析を行った。HCV は、肝臓特異的にウイルス由来の配列が見られたことから、肝臓で特異的に感染、複製していることが示唆された。なお、肝臓 136 検体のうち、3 つの検体が HCV 陽性であった（図 2A）。GTEx が提供する組織切片の画像データから、HCV 陽性の肝臓の画像データを確認したところ、HCV 陽性の 3 検体すべてで、慢性肝炎様の炎症が確認された（図 2B）。このことは、この提供者は、病徴の自覚はないものの、不顕性（気づかない）状態で HCV が感染していたことを示唆する。次に、HCV 陽

性の検体と陰性の検体の間での、ヒト遺伝子の発現を比較した。その結果、ウイルス感染により誘導されるインターフェロン誘導性遺伝子（interferon-stimulated gene：ISG）の発現上昇が見られた（図 2C）。これらの結果から、今回の解析で得られた HCV 陽性の肝臓において、実際に HCV 感染が生じており、ウイルス感染に応答して誘導されるインターフェロン系の免疫応答が顕著に誘導されていることが示唆された。

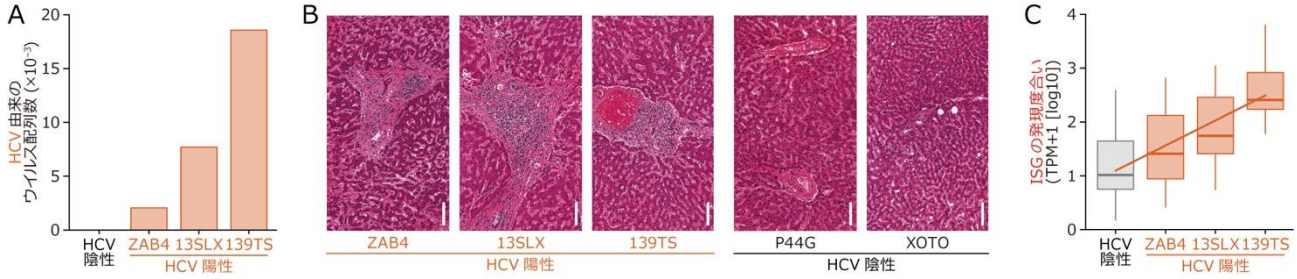


図 2. 肝臓における HCV 感染がインターフェロン誘導性遺伝子（ISG）の発現に与える影響

A） HCV 由来の配列の同定と定量。肝臓の 3 検体（GTEx ID：ZAB4、13SLX、139TS）において、HCV 由来のウイルス配列が検出された。

B） HCV 陽性の肝臓の 3 検体（GTEx ID：ZAB4、13SLX、139TS）と、HCV 陰性の 2 検体（GTEx ID：P44G、XOTO）の肝臓の組織病理像。HCV 陽性の検体において、顕著な慢性肝炎の所見が観察される。スケールバー：100μm。

C） HCV 由来の配列の量（x 軸）と、インターフェロン誘導性遺伝子（ISG；y 軸）発現量の関連。両者に正の相関が確認される。

3. EB ウイルスの溶解感染と B 細胞の分化

EBV は、全身のさまざまな組織で、ウイルス由来の配列が確認された（図 1B）[1]。しかし、EBV の主要な感染標的細胞は、血液細胞の一種である B 細胞である。GTEx の検体は、その検体調整の際、生理食塩水を用いた還流操作が行われているため、組織への血液の残存はあまりないと考えられるが、この広範な組織からの EBV 由来の配列の検出は、微量の血液が残存してしまったことの影響であることも否定できない。いずれにせよ、EBV 由来の配列の主要な感染標的が B 細胞であるという事実をふまえ、今回は、血液細胞を多く含む脾臓と血液において、EBV 由来の配列の有無と、ヒト遺伝子の発現変化の関連を調べた。その結果、EBV 陽性検体において、補体や食作用、抗体産生系などのさまざまな免疫系に関連する遺伝子群が発現上昇していることが確認された（図 3A）。さらに、ヒト遺伝子発現情報から、サンプルに含まれる各血液細胞の割合を推定する cell deconvolution 解析を行ったところ、脾臓と血液の両方において、EBV 陽性検体における形質細胞の割合の有意な上昇が見られた（図 3B）。この結果は、EBV 感染が B 細胞の形質細胞への分化に関与するという近年の報告と合致しており [6]、健常者の体内においても、この関連が見られることを示している。なお、健常人においては、EBV は潜伏感染していることが知られている [7]。EBV は潜伏感染の際、ウイルス複製を行わないため、ウイルス由来の mRNA は産生されないはずである。にもかかわらず、今回の GTEx のデータセットを用いたメタトランスクリプトーム解析により、EBV に由来するウイルス mRNA が複数の検体で検出された。この事実は、ごく一部の EBV が健常人の体内においても溶解感染（新規ウイルスを産生するウイルス複製サイクル）を起こしている可能性、あるいは、遺伝子の欠損などによって複製能を失った EBV が abortive infection（ウイルスとしては複製せず、ウイルスに由来する mRNA のみが転写されている状態）している可能性が示唆される。

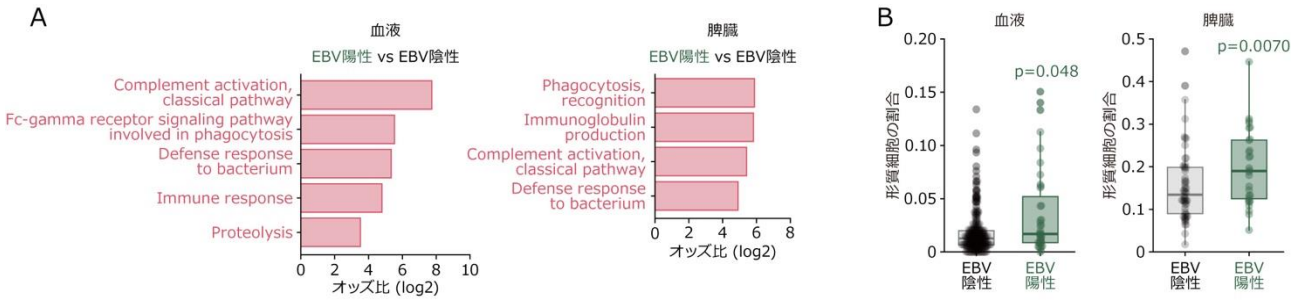


図 3. 血液と脾臓における EBV 感染が免疫応答と B 細胞の分化に与える影響

A) EBV 陽性の検体と陰性の検体の間でのヒト遺伝子発現の遺伝子オントロジー（gene ontology：GO）解析。血液（左）と脾臓（右）において、EBV 陰性の検体に比して EBV 陽性の検体で発現上昇が確認された GO 情報と、それぞれのオッズ比を示した。

B) 血液（左）と脾臓（右）の検体に含まれる各血液細胞の割合を推定する cell deconvolution 解析［8］の結果。統計検定は Welch の t 検定を用いた。

4. 胃における HHV-7 感染

今回の解析で新たに明らかになった点として、ヒトヘルペスウイルス 7 型（HHV-7）に由来するウイルスの配列が、胃に局在して検出された点が挙げられる（図 1B）。HHV-7 は、小児における突発性発疹の原因ウイルスであることが知られている［9］が、成体における病原性や感染の分布、および、その感染の生物学的意義はほとんど不明であった。そこで、HHV-7 が胃に与える影響を調べるために、ヒト遺伝子発現との関連解析を行った。ヒト遺伝子発現を基に、胃の検体をクラスタリング（分類）したところ、大きくふたつの群に分かれた（図 4A 上段）。そして興味深いことに、その片方の群に、HHV-7 陽性の検体が集中していた（図 4A 中段）。これら 2 群の間でのヒト遺伝子発現を詳細に比較したところ、HHV-7 陽性が多い群において、特に消化酵素をコードする遺伝子の顕著な発現上昇が見られた（図 4B）。この結果は、HHV-7 感染が胃の状態と関連する可能性、あるいは、消化酵素群をコードする遺伝子の発現が高い胃の部位において HHV-7 が感染、複製している可能性を示している。

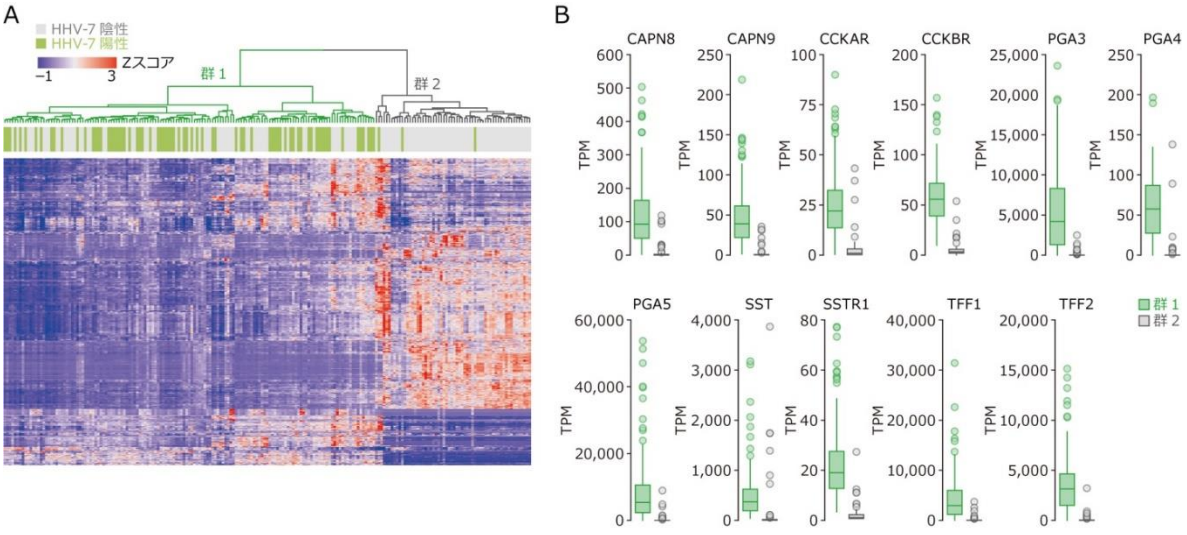


図 4. 胃における HHV-7 感染とヒト遺伝子発現の関連解析

A) 胃におけるヒト遺伝子の発現パターンと HHV-7 の分布。203 個の胃の検体を、発現上位 1,000 遺伝子（下段）の発現パターンでクラスタリング解析を行った。その結果、胃の検体は群 1 と群 2 に分けられ（上段）、HHV-7 陽性の検体は特に群 1 に顕著に分類された（中段）。

B) 群 1、2 それぞれにおける消化酵素関連遺伝子の発現レベル。TPM:：transcripts per million（100 万配列あたりの転写産物量）。

考　察

本研究により、健常者のさまざまな体組織におけるウイルス叢（virome）の実態が初めて描出され、ヒトとウイルスの共生関係の一端が明らかになった［1］。腸内細菌叢をはじめとしたマイクロバイオーム解析が急速に進展している中で、ウイルス叢（virome）については充分に研究が進んでいるとは言いがたい状況にある。その一因として、ウイルスには、生物におけるリボソーム RNA 配列のように、共通して存在するコンセンサス配列が存在しないことが挙げられる。腸内細菌叢の NGS 解析の場合、菌類に共通するリボソーム RNA の配列を解読し、それを基に菌種を分類することが多いが、ウイルスにはそのような方法を適用できないことを意味する。本研究で行った、RNA-seq データを用いたウイルス叢（virome）の解析は、通常のメタゲノム解析で行う、PCR による対象の配列の特異的増幅を伴わないため、低感度になるという欠点も生じる。しかしこれは、アンバイアスにウイルス由来の配列を描出することが可能である、という長所にも通じ、また実際に、本研究において、健常人の体内からさまざまなウイルスを捕捉することに成功している。さらに、ヒトの体組織に由来する RNA-seq データであることから、ウイルス配列の解析に用いるのと同じデータセットから、宿主（ヒト）の遺伝子発現のデータを同時に取得、解析することが可能である。HCV や EBV の例を上述したように、ウイルス由来の配列の有無とヒト遺伝子発現情報の関連解析を行うことで、ウイルスが宿主（ヒト）に与える影響を調べることが可能となる。

共同研究者

本研究の共同研究者は、国立感染症研究所感染病理部の鈴木忠樹博士、高橋健太博士である。

文　献

1) Kumata R et al. A tissue level atlas of the healthy human virome. BMC Biol. 2020. 18:55. PMID: 32493363 DOI: 10.1186/s12915-020-00785-5.
2) Piedimonte G. Respiratory syncytial virus and asthma: speed-dating or long-term relationship? Curr Opin Pediatr. 2013. 25:344-349. PMID: 23657245 DOI: 10.1097/MOP.0b013e328360bd2e.
3) Strachan DP. Hay fever, hygiene, and household size. BMJ. 1989. 299:1259-1260. PMID: 2513902 DOI: 10.1136/bmj.299.6710.1259.
4) Machiels B et al. A gammaherpesvirus provides protection against allergic asthma by inducing the replacement of resident alveolar macrophages with regulatory monocytes. Nat Immunol. 2017. 18:1310-1320. PMID: 29035391 DOI: 10.1038/ni.3857.
5) Consortium G. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013. 45:580-585. PMID: 23715323 DOI: 10.1038/ng.2653.
6) Mrozek-Gorska P et al. Epstein-Barr virus reprograms human B lymphocytes immediately in the prelatent phase of infection. Proc Natl Acad Sci U S A. 2019. 116:16046-16055. PMID: 31341086 DOI: 10.1073/pnas.1901314116.
7) Longnecker RM et al. Epstein-Barr virus. In: Knipe DM, Howley PM, editors. Fields Virology. 2. 6th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2013. p. 1898-1959.
8) Newman AM et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015. 12:453-457. PMID: 25822800 DOI: 10.1038/nmeth.3337.
9) Yamanishi K et al. Human herpesviruses 6 and 7. In: Knipe DM, Howley PM, editors. Fields Virology. 2. 6th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2013. p. 2058-2079.