

jh230035

大規模比較ゲノム解析による病原細菌の進化と病態発症機構の解明

山口雅也（大阪大学）

概要

肺炎球菌と化膿レンサ球菌は、致命的な感染症を引き起こす一方で、健常人からも分離される。本研究では、侵襲性疾患由来株ならびに無症候性定着株の全ゲノム情報を比較し、ゲノムワイド関連解析（GWAS）と分子進化解析を行うことで、病態形成に必要な分子または無症候的に定着するための機構の解明を試みた。化膿レンサ球菌は遺伝子型 *emm89* 型に系統を絞り 666 ゲノムの解析を、肺炎球菌は表現型の情報が入手可能な 12,599 ゲノムの解析を行った。さらに、化膿レンサ球菌においては、得られた情報をもとに感染モデルを用いた実験で因果関係を検証した。

公募型共同研究として実施した意義として、大規模計算システム上で解析パイプラインを構築し、ゲノム情報と臨床情報を収集することで多様な細菌種に対して同様の解析を迅速にできるようになる事が挙げられる。解析対象となる細菌ゲノムは多数存在するため、今後も継続的に大規模計算機システムを活用した研究を推進しうると考える。

1. 共同研究に関する情報

- (1) 共同利用・共同研究を実施している拠点名
大阪大学 サイバーメディアセンター

大野誠之 大阪大学 大学院歯学研究科 微生物学講座・助教

ゲノム解析・細菌学的解析の実施

- (2) 課題分野
大規模計算科学課題分野

東孝太郎 大阪大学 大学院歯学研究科 微生物学講座・大学院生

ゲノム解析・細菌学的解析の実施

- (3) 共同研究分野 (HPCI 資源利用課題のみ)
超大規模数値計算系応用分野

小林桃子 大阪大学 大学院歯学研究科 微生物学講座・大学院生

ゲノム解析・細菌学的解析の実施

- (4) 参加研究者の役割分担
山口雅也 大阪大学 大学院歯学研究科 バイオインフォマティクス研究ユニット・准教授
研究総括・ゲノム解析の実施

2. 研究の目的と意義

川端重忠 大阪大学 大学院歯学研究科 微生物学講座・教授
情報解析結果に基づく細菌学的解析（計算機利用なし）

肺炎球菌と化膿レンサ球菌は、致命的な感染症を引き起こす一方で、健常人からも分離される。しかし、そのような無症候的に定着している菌株と重篤な疾患を引き起こす菌株の間に、どのような違いがあるのかという

点は明らかになっていない。本研究では、侵襲性疾患由来株ならびに無症候性定着株の全ゲノム情報を比較し、ゲノムワイド関連解析 (GWAS) と分子進化解析を行うことで、病態形成に必要な分子または無症候的に定着するための機構の解明を試みた。得られた情報をもとに、感染モデルを用いた細菌および宿主のシングルセル解析などの実験で因果関係を検証するとともに、侵襲性病態の制御手段を探索することとした。

日本において、病態と関連する因子を探索する大規模な細菌の GWAS の論文報告はなされていない。GWAS と分子進化学的解析を行うことで、ゲノム情報から病態と関連する因子を探索するとともに、自然淘汰に基づいた種における重要性を評価する。すなわち、感染モデルを用いた解析のみでは難しい、ヒトの臨床病態を反映した高解像度の解析戦略が可能となる。さらに、細菌および宿主のシングルセル解析を実施することで、感染過程での両者の詳細な変化を明らかにする。

2018 年の報告では世界の 1 年間における下気道感染症による死亡の約半数は肺炎球菌が原因であるとされている (GBD 2016 L.R.I. Collaborators. *Lancet Infect. Dis.* 18: 1191-1210. 2018)。また、化膿レンサ球菌による咽頭炎や劇症型感染症は日本で増加傾向にある。情報解析を活用して、これらの細菌による病態形成機構を探索することは極めて重要であると考えられる。

3. 当拠点の公募型共同研究として実施した意義

大規模計算システム上で解析パイプラインを構築し、ゲノム情報と臨床情報を収集することで多様な細菌種に対して同様の解析を迅速にできるようになる。我々は日本各地の公衆衛生研究所と連携し、日本の化膿レンサ球菌臨床分離株の収集とドラフトゲノム解読を行い、細菌ゲノム情報の拡充を試みている。多数の細菌ゲノム情報が収集されるこ

とで、全体の中の特定の集団 (例えば化膿レンサ球菌における特定の遺伝子型集団) の n 数が増加し、GWAS を行うことが可能となる。

ヒトでは主に抽出された約 100 万塩基対の SNP を用いて、GWAS が行われている。一方で、細菌ゲノムは比較的小さいため、レンサ球菌の全ゲノムである約 200 万塩基対について、遺伝子単位、SNP 単位、 k -mer 単位による病態との相関を計算できる。結果として、同規模のヒトでの一般的な解析よりも、多くの計算資源が必要となる場合が多い。数千株以上のレベルでの解析を行うには、十分なストレージ容量と解析資源が必要である。すなわち、個人用 PC では非現実的な解析時間が必要となっている。また、シングルセル解析においても同様のことが言える。なお、前年の萌芽型研究課題での採択計画の GWAS において、当初の想定よりも大幅に少ない解析時間で一定の目的を達成できた。解析資源量を最小化した上で、大規模計算機による解析を必要としている。

本研究で実施する、細菌種内における進化の選択圧に着目して病原因子を探すアプローチは、研究代表者が確立して報告したものである (Yamaguchi M. et al. *Commun. Biol.* 2: 96. 2019; Yamaguchi M. et al. *Front. Microbiol.*, 11: 582437. 2020)。さらに、一菌体ごとのゲノム変異についても進化解析を適用できる可能性を探索し、感染過程における進化の一端を明らかにする。同一菌株の一菌体ごとの遺伝的多様性については世界でも報告がなく、病原体の変異機構を解明する基盤となりうる。

これらのことから、本計画は大規模計算機を用いて解析することで今後大きな成果が得られることが期待される研究であると考えられる。また、解析対象となる細菌ゲノムは多数存在するため、今後も継続的に大規模計算機システムを活用した研究を推進しうると考える。

4. 前年度までに得られた研究成果の概要

該当なし

5. 今年度の研究成果の詳細

(A) 侵襲性化膿レンサ球菌感染症についてのゲノムワイド関連解析

化膿レンサ球菌は健康なヒトの咽頭や皮膚に存在する常在菌である。咽頭炎や膿痂疹などの非侵襲性感染症の原因菌である一方で、壊死性筋膜炎やレンサ球菌性毒素性ショック症候群などの致死的な侵襲性感染症を引き起こし、“人喰いバクテリア”とも呼ばれている。化膿レンサ球菌による侵襲性感染症の致死率は約 40%と高いことが知られている。しかし発症機序には不明な点が多く、効果的な予防手段が確立されていないのが現状である。

化膿レンサ球菌の分類には、病原因子である M タンパクをコードする遺伝子の配列に基づく *emm* 分類が採用されており、現在 240 種類以上が特定されている。なかでも近年、欧米や日本において、侵襲性感染症患者から *emm89* 型化膿レンサ球菌が分離される頻度が高まっている。*emm89* 型は遺伝学的にさらに 3 つの *clade* に分類されており、近年では先進国の侵襲性疾患患者より分離される *emm89* 型化膿レンサ球菌の大部分が *clade 3* となっている。*Clade 3* は莢膜ヒアルロン酸合成遺伝子群 *hasABC* の欠失と、*nga-ifs-slo* オペロンのプロモーター領域に生じた変異による病原因子 NADase およびストレプトリジン O の発現増加を特徴とする。しかし、日本において侵襲性感染症由来株に加え、非侵襲性感染症由来株でも *clade 3* に分類される菌株の割合が支配的であり、侵襲性感染症特異的な変異ではないことが示唆された。

本研究では、侵襲性感染症の発症機構の解明を目的として、*emm 89* 型化膿レンサ球菌の全ゲノム情報から病態に相関する細菌因

子の探索を行なった。

化膿レンサ球菌について、日本で近年侵襲性もしくは非侵襲性感染症から分離された *emm89* 型臨床分離株 311 株を収集し、全ゲノム情報の解読を行ったほか、すでに公開されている世界の臨床分離株 355 株の全ゲノム情報を追加して 666 株に対する全ゲノム解析を行った。

プログラム *Roary* を用いたパンゲノム解析により 99%以上の株が保有するコア遺伝子を特定したのち、プログラム *snp-sites* により一塩基多型 (SNP) を抽出した。パンゲノム解析の結果、*emm89* 型株 666 株において 4743 種類の遺伝子が検出され、そのうち 1327 遺伝子がコア遺伝子であった。コア遺伝子の SNP をもとに分子系統樹を算出したところ、666 株は 4 つのクラスターに分類された。さらに、*multilocus sequence typing (MLST)* による遺伝子型を調べたところ、666 株のうち 522 株が ST101 であった。日本由来株では、ST101 について ST646 が検出されたが、ST646 は日本由来株に特異的な遺伝子型であった。

次に、全ゲノムから 31 塩基対の長さの配列である *k-mer* を抽出した。得られた SNP および *k-mer* から、病態と関連する因子を *Pyseer* ならびに *DBNGWAS* を用いたゲノムワイド関連解析により探索した。47060 の SNP が検出され、そのうち 1075 の SNP が侵襲性と有意に相関することが示された。また、169 の遺伝子について、侵襲性との相関が示唆された。*k-mer* による解析では 5 つの *complex* と侵襲性との相関が示され、そのうちの一つは SNP による解析でも検出された *fhuB* 遺伝子の変異であった。検出された変異について、*AlphaFold2* を用いて変異がタンパク質の機能や構造に影響を与えうるかの検討を行った。タンパク質の機能に影響を与える可能性が高い変異のうち、鉄イオントランスポーターをコードする *fhuB* 遺伝子に存在する SNP は非侵襲性感染症とより関連し、日

本で分離された株にのみ存在した。

そこで、日本における侵襲性由来株に、非侵襲性由来株に存在する点変異を、温度感受性ベクターを用いて導入した。野生株と変異導入株について、THY 培地中とヒト血液中での増殖能を比較したところ、THY 培地中での増殖率に有意な差は見られなかったが、変異によりヒト血中における菌の増殖率は低下した。

本研究では、世界的に拡大している侵襲性感染症を引き起こす *emm89* 型化膿レンサ球菌に焦点を当てた。細菌 GWAS のワークフローを構築し、重症感染症や侵襲性感染症に関連する細菌遺伝因子を探索した。日本および世界の cohorts について GWAS を実施し、侵襲性化膿レンサ球菌感染症の関連を調べた。化膿レンサ球菌については、これまでにいくつかの GWAS が実施されている。Davies らは、幅広い *emm* 遺伝子型を持つ化膿レンサ球菌の 1,944 株のゲノムから k-mer を抽出し、侵襲性との関連を評価し、Kachroo らは、侵襲性または非侵襲性感染症から分離された 442 株の *emm28* 株を用いて k-mer に関する GWAS を行った。しかし、*emm89* 株を用いて、SNP、遺伝子、k-mer を複合的に解析した研究はない。我々が行った解析では、コア遺伝子の SNP だけでなく、アクセサリ遺伝子や遺伝子間領域にまたがる挿入を含む、膨大な数の原因変異が発見された。

さらに、*fhuB* 変異がヒト血液中の細菌の生存に寄与することを実験的に検証した。この研究は、侵襲性感染症の病態を解明するためのゲノム統計的アプローチの可能性を示すものである。本研究で同定された侵襲性関連因子のさらなる解析は、侵襲性感染症に対する新たな治療法や予防戦略を確立するための基盤となりうる。

(B) 侵襲性肺炎球菌感染症についてのゲノムワイド関連解析

肺炎球菌は口腔レンサ球菌の *mitis* 群に分類されるグラム陽性細菌で、世界 195 カ国の下気道感染症による死亡のうち、半数以上が肺炎球菌によると推計されている。さらに、肺炎球菌は肺炎や中耳炎などに加え、致死的な侵襲性肺炎球菌感染症を引き起こす。侵襲性感染症は複数の因子により複数の機序で発症することが示唆されている。本研究では遺伝統計学的解析により、病態と関連する肺炎球菌の因子の探索を行った。

肺炎球菌について、我々が以前解読した配列と併せて、1978 年から 2018 年までに世界 20 カ国で分離された、臨床情報を備えた 12,599 株のゲノム配列を収集した。12,599 株のうち、侵襲性由来株は 6,616 株、非侵襲性由来株は 5,983 株であった。プログラム Roary を用いたパンゲノム解析により集団中に存在する遺伝子の分布を解析した結果、全体で 42,473 遺伝子が検出された。なかでも 99% 以上の株が保有するコア遺伝子は 687 遺伝子存在することが明らかとなった。続いて、プログラム snp-sites によりコア遺伝子から 274,786 箇所の SNP を抽出した。病態と関連する SNP と遺伝子は、それぞれプログラム Pyseer により探索した。SNP に関する解析の結果、14,788 箇所の変異が病態と有意に関連した。SNP のタンパク質機能への影響を解明するために、AlphaFold2 を利用した。AlphaFold2 の結果から、フェニルアラニン tRNA リガーゼ β サブユニットをコードする *pheT* 遺伝子座の 4 つの SNP は、柔軟なリンカー領域の 3 つのアミノ酸を変化させ、おそらく PheTS 複合体の構造安定性、ひいてはその機能に影響を及ぼすと予測された。検出された SNP の一部は *tarI* や *murA1_2* など細胞壁合成に関与する複数の遺伝子座に存在した。

また遺伝子の分布と病態の関連について検討を行ったところ、42,473 遺伝子のうち、3,102 遺伝子が有意に病態と関連することが

示された。侵襲性関連遺伝子の 60% は hypothetical protein をコードする遺伝子としてアノテーションされたが、その他の遺伝子はガラクトシダーゼ、薬剤排出ポンプ、付着因子などをコードすることが示唆された。

SignalP を用いて細胞外に局在する分子をコードする遺伝子をスクリーニングし、遺伝子欠失株を作製してマウス敗血症モデルにおける病原性を比較した。その結果、肺炎球菌 TIGR4 野生株と比較して、菌体表層分子をコードする *group_3907* 遺伝子の欠失株を感染させたマウスの生存率が有意に高いことが示された。

侵襲性肺炎球菌感染症の発症に関連する遺伝子の分布パターンが複数存在することから、特定の遺伝子の集積した集団ごとに異なる発症機序が存在する可能性が示された。また、遺伝統計学的解析より病態との相関関係が示された *group_3907* が、マウス敗血症モデルにて病原性に寄与することが示唆された。

6. 進捗状況の自己評価と今後の展望

これまでの研究で、肺炎球菌と化膿レンサ球菌について、細菌ゲノム情報とメタデータとしての臨床症状を用いたゲノムワイド関連解析 (GWAS) を実施し、病態と相関する細菌側因子を検出した。さらに一部の分子について、分子生物学的解析にて病原性に寄与することを示した。これらの結果を踏まえ、当初の計画に対して概ね順調に計画を進展できたと考える。

来年度の継続課題では、これまでに得られたデータを検証するとともに、次の二点の解明を試みる。一点目は、肺炎重症化を引き起こす宿主要因の解明である。マウス感染モデルを用いて、加齢による免疫応答の変化ならびに肺微生物叢の影響、ならびに一菌体ごとのゲノム変異を検討する。

二点目は、口腔常在細菌叢の菌叢構造と網羅的な遺伝情報の解明を行う。肺炎球菌は外部遺伝子を取り込みやすい性質を持ち、特に類縁の口腔レンサ球菌を外部遺伝子リザーバーとして用いて、選択圧に応じて薬剤耐性遺伝子や莢膜抗原を変化させることが報告されている。すなわち、口腔微生物叢に薬剤耐性遺伝子や病原因子が含まれている場合、肺炎球菌が感染後にそれらの遺伝資源を利用する可能性がある。実際のヒトの口腔細菌叢の構造と耐性遺伝子の分布などを明らかにすることで、これまで認識されていなかった、健康なヒトの口腔が持っている、肺炎球菌の耐性化リスクを検討する。

継続課題は、バイオインフォマティクス研究者と、細菌学研究者、臨床家が共同して実施する多分野融合の研究計画である。マウス感染モデルを用いた実験は細菌学研究者が実施する。ヒト由来検体の収集については臨床医・臨床歯科医師と連携し、収集を行う。また、細菌ゲノムおよびメタゲノムの解読、宿主細胞のシングルセル RNA-seq などの次世代シーケンス解析は、大阪大学微生物学研究所ゲノム解析室の支援を受けて実施し、情報解析はバイオインフォマティクス研究者が行う。

7. 研究業績

(1) 学術論文 (査読あり)

Nagata N., Kurosaka H., Higashi K., Yamaguchi M., Yamamoto S., Inubushi T., Nagata M., Ishihara Y., Yonei A., Miyashita Y., Asano Y., Sakai N., Sakata Y., Kawabata S., Yamashiro T. Characteristic craniofacial defects with novel truncated variant of USP9X. *Hum Genome Var.* *in press.*

Yamaguchi M.[†], Uchihashi T., Kawabata S.[†] Hybrid sequence-based analysis reveals the

distribution of bacterial species and genes in the oral microbiome at a high resolution. **Biochem. Biophys. Rep.** Vol. 38. 101717. 2024. †*corresponding author.*

Ikeda E., Yamaguchi M., Kawabata S., Gut microbiota-mediated alleviation of dextran sulfate sodium-induced colitis in mice. **Gastro Hep Advances.** Vol. 3. 461-470. 2024.

Takebe K., Suzuki M., Sangawa T., Kreikemeyer B. (+), Yamaguchi M., Uzawa N., Sumitomo T., Kawabata S., Nakata M. Analysis of FctB3 crystal structure and insight into its structural stabilization and pilin linkage mechanisms. **Arch. Microbiol.**, Vol. 206. 4. 2023.

Kubota S., Nakata M., Hirose Y., Yamaguchi M., Kreikemeyer B. (+), Uzawa N., Sumitomo T., Kawabata S. Involvement of ribonuclease Y in pilus production by M49 *Streptococcus pyogenes* strain via modulation of mRNA level of transcriptional regulator. **Microbiol. Immunol.**, Vol. 67. 319-333. 2023.

Kawanishi K.†, Naito-Matsui Y.†, Soares Zaramela L.† (+), van Sorge N.M.† (+), Yamaguchi M.†: Editorial: A sweet deal – Glycobiology in host-pathogen interactions. **Front. Microbiol.**, Vol. 14. 1341820. 2023. †*corresponding author. Editorial*

(2) 国際会議プロシーディングス (査読あり)
該当なし

(3) 国際会議発表 (査読なし)
シンポジウム・特別講演
Yamaguchi M. Elucidation of mechanism by which aging exacerbates pneumococcal

pneumonia. The 2nd International Conference for Future Dentistry (ICFD). 2024 年 4 月 5 日. Bangkok, Thailand.

Ono M. Large-scale Genome-wide Association Study for Exploring Pneumococcal Factors Relating to Invasive Infections. The 4th Asian Pneumococcal Symposium. 2023 年 12 月 7 日. Seoul, Korea.

ワークショップ

Yamaguchi M., Kobayashi M., Kawanishi K., Ono M., Motooka D., Okuzaki D., Kawabata S. Insight into neutrophil phenotypic alterations in relation to the severity of pneumococcal infection with age. The 21st Awaji International Forum on Infection and Immunity. 2023 年 9 月 3 日. Karuizawa, Nagano, Japan.

ポスター発表

Higashi K., Yamaguchi M., Takebe K., Nakata M., Sumitomo T., Suzuki M., Nizet V. (+), Kawabata S. Hyaluronidase in *Streptococcus pyogenes* – analysis based on structural biology and molecular phylogenetics. Society for Glycobiology 2023 Annual meeting. 2023 年 11 月 6 日. Hawaii, USA.

Yamaguchi M., Kobayashi M., Kawanishi K., Ono M., Motooka D., Okuzaki D., Kawabata S. Elucidation of the impact of age-related changes in the host response on the severity of pneumococcal infections. FEMS2023, 2023 年 7 月 10 日. Hamburg, Germany.

(4) 国内会議発表 (査読なし)
シンポジウム・特別講演

山口雅也. 大規模情報解析を活用した細菌感染症の病態解明. Cyber HPC Symposium 2024. 2024 年 3 月 1 日. 吹田市, 大阪府.

山口雅也, バイオインフォマティクス研究ユニットの使い方 2023. 歯学研究科バイオインフォマティクス ショートセミナー 2023. 2023 年 11 月 27 日. 大阪府.

山口雅也. 異分野融合による肺炎球菌感染症の重症化機構の解明. 第 35 回微生物シンポジウム. 2023 年 9 月 1 日. 岡山市.

口頭発表

池田恵莉, 山口雅也, 川端重忠. 腸内細菌叢が潰瘍性大腸炎モデルマウスの疾患重症度に与える影響の検討. 第 76 回日本細菌学会関西支部総会. 2023 年 11 月 18 日. 大阪府.

大野誠之, 山口雅也, 元岡大祐, 広瀬雄二郎, 東孝太郎, 秋山徹, 住友倫子, 池辺忠義, 山口貴弘, 河原隆二, 奥野ルミ, 大塚仁, 松本裕子, 賀澤優, 川端重忠. オミクス解析を用いた *emm89* 型化膿レンサ球菌侵襲性感染症の発症に関与するメカニズムの解明. 2023 年口腔微生物研究会. 2023 年 9 月 19 日. 東京都.

大野誠之, 山口雅也, 川端重忠. 大規模ゲノム解析による侵襲性肺炎球菌感染症の発症機構の解明. 第 65 回歯科基礎医学会学術大会. 2023 年 9 月 17 日. 東京都.

小林桃子, 山口雅也, 川西邦夫, 元岡大祐, 奥崎大介, 川端重忠. 肺炎球菌感染症の重症化に寄与する宿主の肺環境の解析. 第 17 回細菌学若手コロッセウム. 2023 年 8 月 18 日. 福岡.

大野誠之, 山口雅也, 元岡大祐, 広瀬雄二郎, 東孝太郎, 秋山徹, 住友倫子, 池辺忠義, 山口

貴弘, 河原隆二, 奥野ルミ, 大塚仁, 松本裕子, 賀澤優, 川端重忠. 日本および世界の *emm89* 型化膿レンサ球菌侵襲性感染症に関与する遺伝因子の比較. 第 53 回レンサ球菌研究会. 2023 年 7 月 23 日. 前橋市.

ポスター発表

山口雅也, 小林桃子, 川西邦夫, 大野誠之, 元岡大祐, 奥崎大介, 川端重忠. シングルセル RNA-seq 解析による細菌性肺炎重症化機構の解明. NGS EXPO 2023. 2023 年 11 月 15 日. 大阪府.

武部克希, 鈴木守, 東孝太郎, 山口雅也, 住友倫子, 川端重忠, 中田 匡宣. *Streptococcus sanguinis* が産生する線毛タンパク質の X 線結晶構造解析. 第 65 回歯科基礎医学会学術大会. 2023 年 9 月 16 日. 東京.

山口雅也, 川端重忠. 細菌シングルセルゲノム解析による唾液細菌叢の解析. 第 65 回歯科基礎医学会学術大会. 2023 年 9 月 17 日. 東京

中田匡宣, 窪田星子, 広瀬雄二郎, 山口雅也, 住友倫子, 川端重忠. 化膿レンサ球菌の RNase Y による線毛産生量の調節. 第 65 回歯科基礎医学会学術大会. 2023 年 9 月 17 日. 東京

山口雅也, 大野誠之, 東孝太郎, 小林桃子, 川端重忠. 大規模比較ゲノム解析による病原細菌の進化と病態発症機構の解明. 第 15 回 JHPCN シンポジウム. 2023 年 7 月 6 日. 東京 (オンライン)

大野誠之, 山口雅也, 川端重忠. 大規模ゲノムワイド関連解析による侵襲性肺炎球菌感染症の発症因子の探索. チーム阪大リトリート 2023. 2023 年 6 月 19 日. 吹田市, 大阪府.

(5) 公開したライブラリなど

該当なし

(6) その他（特許，プレスリリース，著書等）

山口雅也，川端重忠．基礎研究から見た化膿レンサ球菌による病態形成機構と宿主の防御機構．*医学のあゆみ*，医歯薬出版．Vol.288．pp869-874．2024．