

## 【総説】

## 次世代シーケンス (NGS) による劇症型レンサ球菌の病原性解析

竹本 訓彦<sup>1)</sup>・小倉 康平<sup>1)</sup>・渡邊 真弥<sup>2)</sup>・秋山 徹<sup>1)</sup><sup>1)</sup> 国立研究開発法人 国立国際医療研究センター (研究所) 病原微生物学研究室\*<sup>2)</sup> 自治医科大学感染・免疫学講座細菌学部門

(平成 29 年 1 月 16 日受付・平成 29 年 4 月 19 日受理)

次世代シーケンスの登場により病原微生物の全ゲノム解析はきわめて身近なものとなった。本稿では次世代シーケンスを活用したレンサ球菌解析の事例を紹介させていただく。① *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) は近年、劇症型症例 (STSS) からの分離が増加している。われわれはこれらの菌種の全ゲノム配列を世界に先駆けて明らかにし、SDSE が他のレンサ球菌と比較して、A 群レンサ球菌 (GAS) と最も近縁であるが、いくつかの GAS の重要病原因子を欠落していることを明らかにした。② GAS はヒトやマウスにおいて *in vivo* で *covRS* などの制御系因子の遺伝子を破壊することで高病原性化することが知られているが、それ以外の遺伝子にも変異が起こるのかについてはほとんど知見がない。そこで、マウス皮下投与モデルで変異を誘発させた菌株を、NGS により多数解析し、変異の入った位置の特定を試みた。③ 種々の制御因子を破壊した GAS 株をマウスに投与後、RNA-seq により *in vivo* でのトランスクリプトームの変化を解析している。発現データを元に、各遺伝子をクラスタリングし、ネットワーク解析を実施している。④ GAS では、病原性因子の制御にかかわる転写因子が複数同定されている。しかし、各転写因子が直接的に制御する regulon の網羅的解析はほとんど行われておらず、転写制御因子間の制御ネットワークや各転写因子の STSS 発症過程における役割についての理解はほとんど進んでいない。そこでわれわれは既知の病原性因子の発現を制御する転写因子 CovR の ChIP-seq 解析と *covR* 遺伝子破壊株における RNA-seq 解析を行い、CovR が制御する遺伝子群を探索している。⑤ 当センターはエボラ、デング、MERS、ジカ熱などの輸入感染症に対応する指定医療機関である。今回、アラブ首長国連邦とモロッコに観光のために渡航し、ラクダに乗った後、帰国後、MERS-CoV 感染が疑われる症状を発症し、当センターで治療された症例が、実は GAS による感染であることが判明した症例を経験した。同菌の全ゲノム配列を解析した事例を紹介する。

**Key words:** *Streptococcus pyogenes* (Group A streptococci), *Streptococcus dysgalactiae* subsp. *equisimilis*

次世代シーケンス (NGS) の出現は病原体や細菌の研究のパラダイムシフトの原動力となっている。NGS は価格もかなり低下し、簡便性、得られる情報量からしても、サンガー法による PCR-direct sequencing に比べてわかりつつある。本稿では NGS を活用したレンサ球菌研究の事例を紹介させていただく。

*Streptococcus pyogenes* [または A 群レンサ球菌 (GAS)] は咽頭炎、猩紅熱、膿痂疹、産褥熱、リユーマチ熱、糸球体腎炎、そして致死率が 50% にも及ぶ劇症型レンサ球菌感染症 (STSS) まで、さまざまな感染症や続発症の原因となる。レンサ球菌の型別は Lancefield 抗原により行われ、GAS 内の型別は M 蛋白質をコードする *emm* 遺伝子の配列解析により実施される。M 蛋白質の型と GAS が起こす疾患の間にはある程度の相関があることが知られている。一方で近年、Lancefield

抗原で G または C に分類される、*Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) による侵襲性感染症の報告が増加しており、その病原体としての重要性が認識されている。STSS は数年前までは国内での報告は年間 200 例程度だったが、近年増加の一途をたどっており、昨年は約 500 例が報告された。地球規模での予測では、先進国と発展途上国の年間発生数はそれぞれ 29,000 例と 2,243,000 例、死亡数は 4,400 例と 561,000 例とされている<sup>1)</sup>。その初期症状は四肢の疼痛、腫脹、発熱、血圧低下であり、発病から病状の進行が非常に急激かつ劇的で、発病後数十時間以内には軟部組織壊死、急性腎不全、成人型呼吸窮迫症候群、播種性血管内凝固症候群、多臓器不全を引き起こす。死亡率は 30~50% と予後は非常に不良である。診断は通常無菌部からのレンサ球菌の分離により行われる。治療はペニシリン・クリンダマイシンの投与、生菌の病巣

\*東京都新宿区戸山 1-21-1

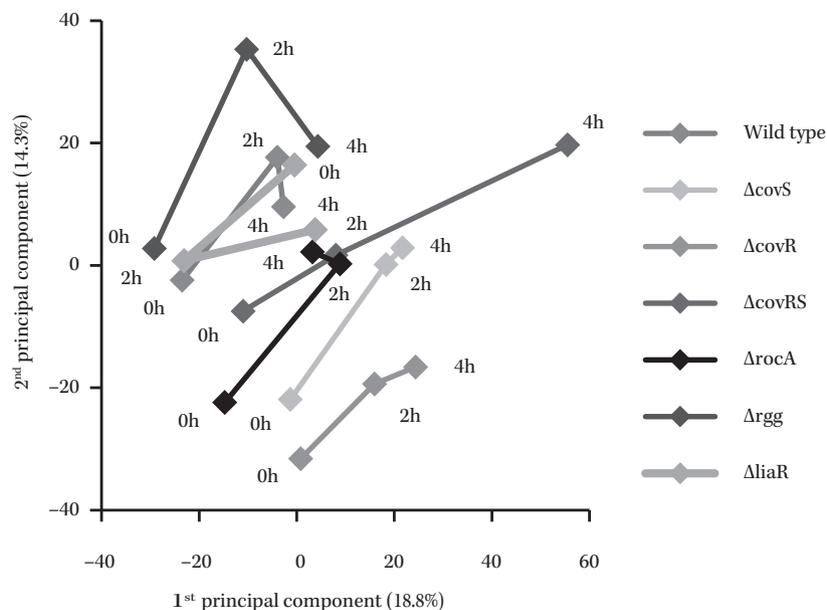


Fig. 1. Principle component analysis of GAS mutants. Based on the RNA-seq analysis of deletion mutants of transcriptional regulators indicated, principle component analysis was performed.

である壊死部の切除により実施される。GASはM蛋白質、フィブロネクチン結合蛋白質、リボタイコ酸、ヒアルロン酸夾膜、ストレプトキナーゼ、DNase、ストレプトリジン、スーパー抗原(発熱性毒素)、プロテアーゼ類などの多様な病原因子をもっており、これがGASの感染症制御や研究を困難にしている。STSS症例から分離されるGASは*emm1*型が優勢であり、過去には*emm3*型も優勢であったことが国立感染症研究所レンサ球菌レファレンスセンターにより明らかにされている。われわれは、マウスの感染実験で*emm1*や*emm3*型のGASが他の型のGASよりも高病原性であることを報告した<sup>2)</sup>。ヒトで劇症型感染症を起こしやすい*emm1*と*emm3*型がマウスにおいて高病原性であることから、このマウス感染モデルはGASのヒト感染のある局面におけるモデルとして使用可能であると考えられ、われわれは、マウスモデルを積極的に利用している。以下では、マウスモデルを使用した研究事例についてもいくつか紹介する。

#### I. *In vivo*での*covRS*の変異誘発にかかわる研究

*CovRS*は2成分制御系として知られるGASなどの発現制御因子である。*CovRS*は主に負の制御因子として機能していると考えられており、その破壊により、GASにおいては全遺伝子の約15%の発現が変化する。それらのなかには多数の病原因子が含まれており、それらの発現は総じて上昇する<sup>3)</sup>。*covS*の破壊では、マウスにおける病原性は親株と比較して、約10,000倍上昇する<sup>4)</sup>。Walker等は、マウスに投与されたGASにおいて、*covRS*の破壊が発生し、この破壊にはDNaseであるGASの*sda1*という遺伝子が関与していることを報告した<sup>5)</sup>。大変興味深いことに、STSS症例から分離されたGASの約50%でも

*covRS*が破壊されていることが明らかにされている<sup>6)</sup>。しかし、どのようにして遺伝子破壊が起こるのか、そして他のゲノム領域における破壊のスペクトラムは明らかになっていない。そこでわれわれはNGSを用いて、マウス皮下で最大6日間の維持後に分離した多数のGASの全ゲノム解析を実施した。すると大変興味深いことに、野生型株では、分離された株すべてにおいて*covS*に変異を認めたのに対し、*sdaD2*(使用株における*sda1*に相当)破壊株では、*covS*変異株の頻度が低下し、他の転写制御因子にも変異が発生することが明らかになった。その一方、転写制御因子以外の部分には、*covS*に100%変異が入る条件でも、90%の株ではまったく変異が発生せず、*covS*の遺伝子破壊がきわめて特異的であることが明らかとなった。このような特異的な遺伝子破壊に関与する因子については現在探索中である。

#### II. 腹腔内投与後のGASのトランスクリプトーム解析

良く知られているようにNGSでは全ゲノム解析の他にも、さまざまな応用が可能である。RNA-seqもその一つである。われわれは*covS*などのさまざまな転写制御因子を破壊したGAS菌株をマウス腹腔に投与し、経時的に回収して、RNAを調製し、GASの*in vivo*におけるトランスクリプトームを解析した。その結果、転写制御因子の破壊により発現が増加する因子として、ストレプトリジンOなどの9個の病原因子の遺伝子、ファージ蛋白質/CRISPRオペロン、金属獲得系、プリン/ピリミジン合成系の遺伝子、そしていくつかのhypothetical proteinが同定された。また、これらの遺伝子のなかには、構成的に発現が上昇する遺伝子11個が存在し、その内訳は、

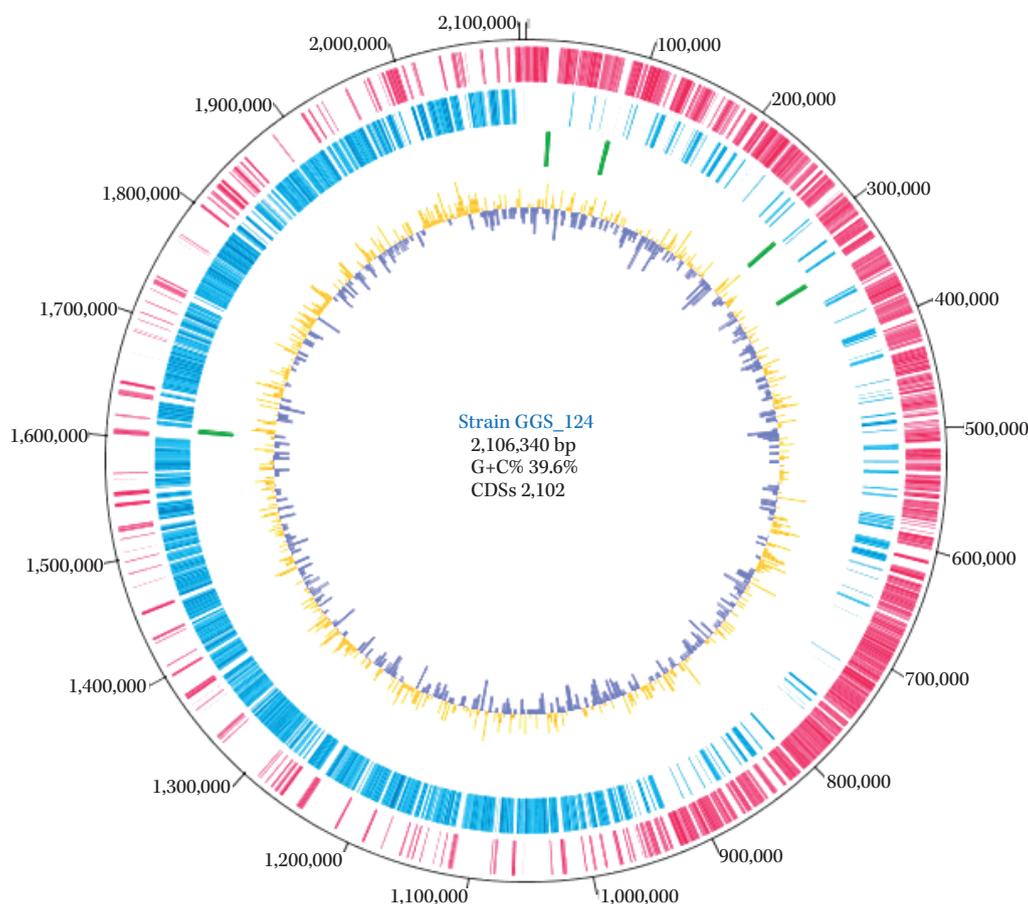


Fig. 2. Genome map of SDSE\_GGS124.

Blue bars and red bars indicate the coding segment in positive strand and negative strand, respectively. There are five *rrs* operon, which are indicated in green.

病原性遺伝子8個と3個のhypothetical proteinだった。11個のうち、8個が既知病原因子であることから、この3個のhypothetical proteinも新たな病原因子である可能性があると考えている。発現データを利用した主成分分析では、第1成分は全体の変動の18.8%を、第2成分は14.3%を説明し、前者は主にマウス腹腔内での経時的変化によるもの、後者は導入した変異によるものと考えられた (Fig. 1)。GASはモデル生物ではないため、パスウェイ解析のツールはほとんど準備されていない。そこで、発現制御ネットワークを解析するため、階層的クラスタリング後に、グラフィカルガウシアンモデルでネットワーク推定を実施した。その結果、遺伝子群は13個のクラスターに分かれ、そのなかで、病原因子が集中する合計200個あまりの遺伝子を含むクラスターが同定され、このクラスターに含まれる遺伝子は *in vivo* での *covRS* 破壊後の高病原性化に寄与していると推定される。

### III. A群レンサ球菌以外のレンサ球菌 (SDSE) による STSS

先述したように、近年SDSEによる侵襲性レンサ球菌

感染症の報告が増加している。厚生労働省研究班が実施した調査でもSDSEは高頻度に侵襲性感染症から分離されている (<http://strep.umin.jp>)。しかしながら、SDSEの全ゲノム情報は当時、解明されていなかった。そこでわれわれは次世代シーケンサを利用して、侵襲性感染症から分離された Lancefield G群に属するSDSE GGS\_124株の全ゲノム配列を決定した (Fig. 2)<sup>7)</sup>。GGS\_124のゲノムサイズは約2.1 Mbpであり、G+C含量は39.6%だった。SDSEは16S rRNA分類では *S. agalactiae* に近い。しかし、全ゲノムレベルで比較を行ったところ、SDSEはレンサ球菌のなかでGASと最も近縁な菌種であることが明らかとなった (Table 1)。SDSEはGASにおいて重要とされている病原因子、例えば溶血毒素であるストレプトリジンO (SLO)なども共有している (Table 2)。しかしながら、セリンプロテアーゼをコードする *speB*、スーパー抗原をコードする *smeZ*、ヒアルロン酸合成酵素をコードする *hasA* と *hasB* を欠落している。この知見はGASの遺伝子群を搭載したマイクロアレイによる多数の菌株の解析でも確認されている<sup>8)</sup>。SDSEはスーパー抗原遺伝子として唯一 *speG* を保有しているが、

Table 1. Summary of genomes of *Streptococcus*

Bacterium	Lancefield	Genome size (Mbp)	G + C %	Number of CDS	Number of the <i>rrs</i> operons	Genome cover rate by SDSE (%)
SDSE GGS_124	G	2.11	39.6	2,102	5	—
GAS (13 strains)	A	1.84–194	38.3–38.7	1,697–1,987	5–6	63–66
<i>S. equi</i> subsp. <i>zooepidemicus</i>		2.02	41.3	1,961	5	30
MGCS10565	C					
<i>S. agalactiae</i> (3 strains)	B	2.13–2.21	36.5–36.7	1,995–2,114	7	16–17
<i>S. thermophilus</i> (3 strains)		1.8–1.86	38.1	1,710–1,915	6	10

Cover rate of each genome by SDSE is based on results of the BLAST analysis.

Table 2. Existence or non-existence of major GAS virulence factors in SDSE

Item	Virulence factor in GAS	Existence/non-existence in SDSE
Adhesion factor	Fibronectin binding protein	+
	Collagen binding protein	+
	Laminin binding protein	+
Toxic factor	Streptolysin O	+
	Streptolysin S	+
	Streptococcal pyrogenic exotoxin type G (SpeG)	+
	Streptococcal mitogenic exotoxin Z (SmeZ)	–
	Hemolysin	+
Immunity evasion	Streptococcal protein interfering with complement function (Sic)	+
	Myosin-cross reactive streptococcal antigen	+
	Erythrogenic toxin B/Cysteine proteinase (SpeB)	–
	Hyaluronate synthetase (HasA, HasB)	–
	Streptococcal C5a peptidase	+
Regulatory factor	C3-degrading proteinase	+
	M protein transacting positive regulator (Mga)	+
	covRS	+

SDSE の *speG* 産物はヒト細胞に対するスーパー抗原活性を欠落しており<sup>9)</sup>, SDSE にはヒト病原性にかかわるスーパー抗原は存在しない。これらの病原因子を欠落しているにもかかわらず、マウス感染モデルでは、SDSE は GAS と同等の病原性を示す。このため、SDSE による STSS は、病態としては GAS によるものと類似しているが、その詳細な発症機構は異なる可能性があると考えられる。

#### IV. CovR を標的とした Chip-seq による CovRS システムにより直接的に制御される遺伝子群の同定

NGS では ChIP-seq により、転写制御因子の標的となる遺伝子の同定も可能である。われわれは、ChIP-seq により、前述の CovRS システムが制御する遺伝子群の同定を試みている。ゲノム DNA に結合している制御因子などの蛋白質を架橋により固定化・断片化した後、CovRS に対する抗体により免疫沈降を行う。得られた沈降物から DNA を回収した後、NGS により配列解析を行うことで CovRS の標的となる遺伝子候補を明らかにすることができる。解析の結果、ゲノム全体で 119 カ所の

結合部位候補が認められ、特に、病原因子と代謝酵素と転写因子が高頻度に存在した。これらのなかには、これまで CovR の標的としては知られていなかった新たな因子が多数含まれていた。これらの因子の遺伝子の上流への CovR の結合をゲルシフトアッセイで確認するとともに、CovR ノックアウト株と野生株の比較で、標的遺伝子の発現が変化していることを確認している。CovR による遺伝子発現制御では病原因子の発現制御が特に注目されているが、今回の結果は、CovR は糖代謝系の制御にも重要な役割を果たしていることを示唆している。

#### V. NGS の臨床研究への応用

NCGM はエボラ、デング、MERS、ジカ熱などの輸入感染症に対応する指定医療機関であり、それらの疑い症例の治療が実施されている。今回、アラブ首長国連邦とモロッコに観光のために渡航し、ラクダに乗った後、帰国後、MERS-CoV 感染が疑われる症状を発症し、当センターで治療された症例が、実は GAS による感染であることが判明した症例を経験した。GAS が MERS 様症状を発生させるきわめて興味深い症例であったため、同菌

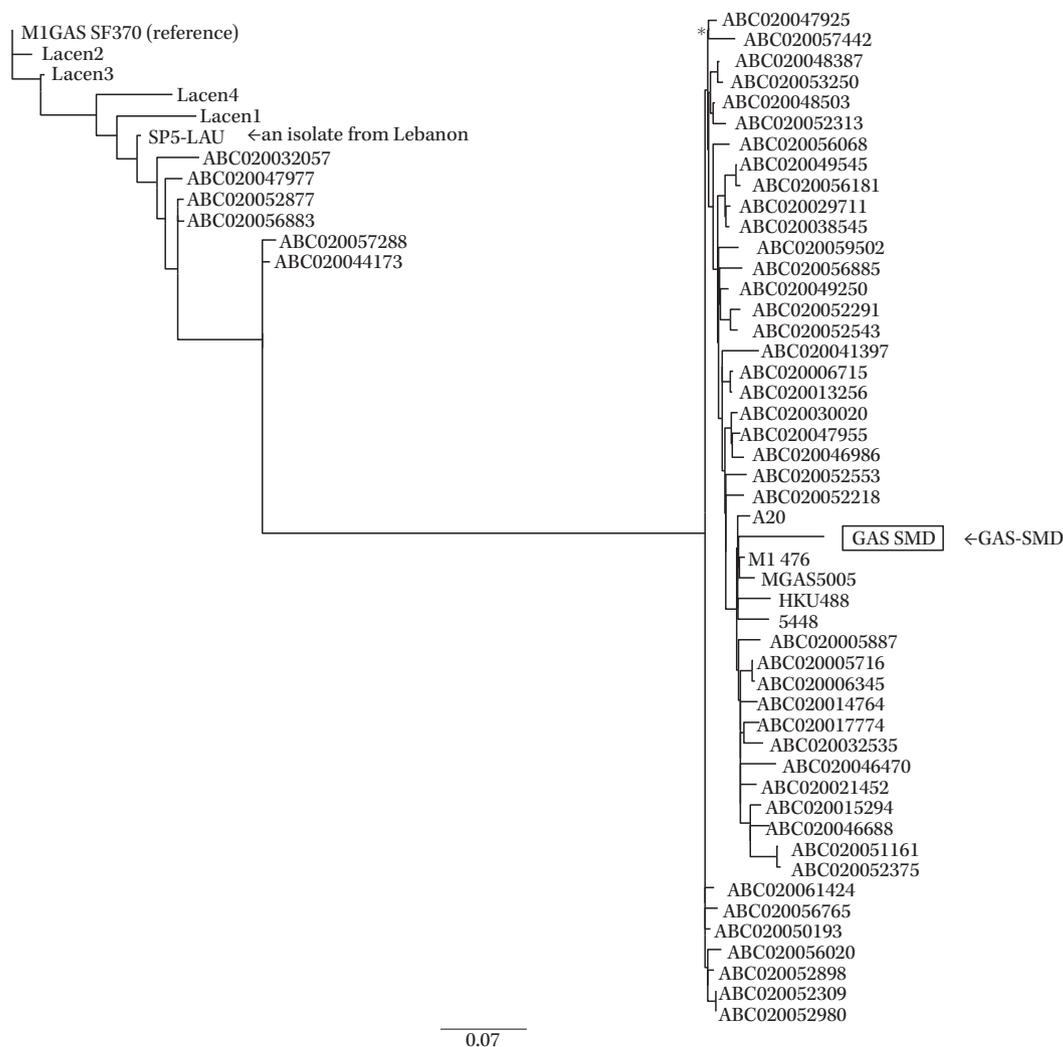


Fig. 3. Phylogenetic analysis of *emm1*-type GAS isolated from MERS suspected case.

The isolate named GAS-SMD is close to isolates from Japan, Taiwan and some of them from USA, and is not to close to a isolate from Lebanon.

の全ゲノム配列を解析した事例を紹介する。

GAS-SMD と名付けられた分離菌の全ゲノム解析の結果では、本菌は *scpA*, *cpa*, *fbp54*, *emm*, *slo*, *ska*, *speB*, *hasB*, *hasA*, *fctA*, *ideS/mac*, *hasC*, *lmb*, *mfs/spd*, *mfs3*, *spea*, *srtC1*, *speg*, *spej*, *smeZ*, *fctB*, *grab*, *lepA* といった典型的な GAS の病原因子を保有しており、*emm1* 型だった。全ゲノム配列を元に、SNP コンカテマーを作製し、データベースに登録されている *emm1* 型とゲノムレベルでの比較を行ったところ GAS-SMD は日本型・台湾型・一部の米国型に近く、特に当センターで全ゲノム解析をした日本における STSS 症例の由来株である GAS476 と最も近縁だった。一方、レバノン型のものとは、近縁ではなかった (Fig. 3)。このことから本症例では、患者は帰国してから罹患した可能性が高いと推定された。

## VI. おわりに

感染症研究において、NGS は、遺伝子情報がすべて得

られるきわめて強力なツールであり、PCR でわかることはすべてわかるため、研究手法のパラダイムシフトを起こしている。発現解析 (RNA-seq) など、これまではカスタムマイクロアレイでしかできなかった解析が安価・簡便に実施可能であり、研究の幅を大きく広げることは間違いない。解析用のソフトウェアも急速に整備されており、バイオインフォマティクスの専門家でなくとも、解析が可能な環境が整備されつつあり、さらに応用されていくものと予想される。

## 謝辞

第 65 回日本感染症学会東日本地方会学術集会・第 63 回日本化学療法学会東日本支部総会合同学会にて講演の機会をいただきました。菊池賢先生・大塚喜人先生、ならびに本稿執筆の機会をいただきました館田一博先生に感謝いたします。

これらの研究の一部は国立研究開発法人日本医療研究開発機構の新興・再興感染症に対する革新的医薬品等開

発推進研究事業の支援を受けて実施された。

利益相反自己申告：申告すべきものなし。

文 献

- 1) Carapetis J R, Steer A C, Mulholland E K, Weber M: The global burden of group A streptococcal diseases. *Lancet Infect Dis* 2005; 5: 685-94
- 2) Miyoshi-Akiyama T, Zhao J, Kikuchi K, Kato H, Suzuki R, Endoh M, et al: Quantitative and qualitative comparison of virulence traits, including murine lethality, among different M types of group A streptococci. *J Infect Dis* 2003; 187: 1876-87
- 3) Graham M R, Smoot L M, Migliaccio C A, Virtaneva K, Sturdevant D E, Porcella S F, et al: Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and *in vivo* infection modeling. *Proc Natl Acad Sci U S A* 2002; 99: 13855-60
- 4) Miyoshi-Akiyama T, Ikebe T, Watanabe H, Uchiyama T, Kirikae T, Kawamura Y: Use of DNA arrays to identify a mutation in the negative regulator, *csrR*, responsible for the high virulence of a naturally occurring type M3 group A streptococcus clinical isolate. *J Infect Dis* 2006; 193: 1677-84
- 5) Walker M J, Hollands A, Sanderson-Smith M L, Cole J N, Kirk J K, Henningham A, et al: DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. *Nat Med* 2007; 13: 981-5
- 6) Ikebe T, Ato M, Matsumura T, Hasegawa H, Sata T, Kobayashi K, et al: Highly frequent mutations in negative regulators of multiple virulence genes in group A streptococcal toxic shock syndrome isolates. *PLoS Pathog* 2010; 6: e1000832
- 7) Shimomura Y, Okumura K, Murayama S Y, Yagi J, Ubukata K, Kirikae T, et al: Complete genome sequencing and analysis of a Lancefield group G *Streptococcus dysgalactiae* subsp. *equisimilis* strain causing streptococcal toxic shock syndrome (STSS). *BMC Genomics* 2011; 12: 17
- 8) Davies M R, McMillan D J, Beiko R G, Barroso V, Geffers R, Sriprakash K S, et al: Virulence profiling of *Streptococcus dysgalactiae* subspecies *equisimilis* isolated from infected humans reveals 2 distinct genetic lineages that do not segregate with their phenotypes or propensity to cause diseases. *Clin Infect Dis* 2007; 44: 1442-54
- 9) Zhao J, Hayashi T, Saarinen S, Papageorgiou A C, Kato H, Imanishi K, et al: Cloning, expression, and characterization of the superantigen streptococcal pyrogenic exotoxin G from *Streptococcus dysgalactiae*. *Infect Immun* 2007; 75: 1721-9

## Investigations of invasive streptococcal infection based on next generation sequencing

Norihiko Takemoto<sup>1)</sup>, Kohei Ogura<sup>1)</sup>, Shinya Watanabe<sup>2)</sup> and Tohru Miyoshi-Akiyama<sup>1)</sup>

<sup>1)</sup> Pathogenic Microbe Laboratory, Research Institute, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku-ku, Tokyo, Japan

<sup>2)</sup> Division of Bacteriology, Department of Infection and Immunity, School of Medicine, Jichi Medical University

With the advent of the next generation sequencer (NGS), whole genome analysis of pathogenic microorganisms has become very familiar. This review introduces our investigations of an invasive streptococcal infection using an NGS. 1) The isolation frequency of *Streptococcus dysgalactiae* subsp. *equisimilis* (SDSE) from invasive streptococcal infection cases has been increasing recently. We reported for the first time on the complete genome sequence of SDSE strains, and found that SDSE is the closest relative of group A streptococcus (GAS), while SDSE lacks several important virulence factors. 2) Although it is known that GAS achieves hypervirulence during the infectious stage in both humans and mice through disrupting *covRS*, a two-component regulatory system, it is unclear if the other genes are also disrupted in the *in vivo* environment. We used an *in vivo* mutagenesis scheme to obtain mutant GAS strains, and identified the location of the generated mutations with an NGS. 3) We generated several GAS mutants, of which various transcriptional regulators were deleted. After mouse infection of the mutants, their transcriptome data were obtained by RNA-seq, genes were clustered, and their network was analyzed. 4) Although there are a number of transcriptional regulators in GAS, there are only a few reports describing the regulons, which are directly regulated by certain transcriptional regulators. The relationships among the regulatory network and the role(s) of transcriptional factors during the pathogenesis of STSS are therefore not well-understood. We utilized ChIP-seq and RNA-seq to explore the genes regulated by *CovRS*. 5) Our institute is one of the designated medical institutions for specified infectious diseases such as Ebola, Dengue, MERS and Zika fever. We encountered a patient who showed symptoms suspected as being caused by MERS-CoV infection since he had traveled in the UAE and Morocco for sightseeing and had ridden a camel. MERS-CoV, in turn, is actually one of the infectious diseases caused by GAS. We were able to analyze the whole genome sequence of the isolate.