

論文 / 著書情報
Article / Book Information

題目(和文)	ヒト細菌叢の個人特異性と常在因子の解明
Title(English)	A personality and commensal factors of human microbiome
著者(和文)	渡邊日佳流
Author(English)	Hikaru Watanabe
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第11072号, 授与年月日:2019年3月26日, 学位の種別:課程博士, 審査員:山田 拓司,北尾 彰朗,伊藤 武彦,本郷 裕一,平沢 敬
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第11072号, Conferred date:2019/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

平成30年度

ヒト細菌叢の個人特異性と 常在因子の解明

東京工業大学 生命理工学院

生命理工学系 生命理工学コース

平成30年度入学

渡邊 日佳流

指導教員 山田 拓司 准教授

目次

第1章 背景	5
第2章 ヒト皮膚細菌叢を用いた個人特定	9
2.1 緒言	9
2.2 材料と方法	10
2.2.1 細菌叢サンプル採取	10
2.2.2 メタデータ採取	10
水分量の計測	10
油分量の計測	11
酸性度の計測	11
2.2.3 16S rRNA 遺伝子の PCR 増幅	11
2.2.4 PCR 産物の精製	12
2.2.5 PCR 産物の定量	12
2.2.6 16S rRNA 遺伝子アンプリコンシーケンス	13
2.2.7 16S rRNA 遺伝子系統解析	13
2.2.8 97% OTU の系統組成テーブル作成	13
Our dataset について	13
Public dataset について	14
2.2.9 多様性指数の計算	15
2.2.10 距離法を用いたサンプル間比較	15
2.2.11 個人特定手法について	16
2.2.12 拡張版個人特定手法について	16
2.2.13 個人特定に関連したパラメータについて	17
個人特定における reference 数の影響	17
個人特定における時系列の影響	17
個人特定における少数種の影響	18
2.2.14 個人特定に重要な系統 personal OTU の探索	18

2.2.15	ヒト皮膚メタデータと OTU0001 の相対存在量の相関解析	18
2.3	結果	19
2.3.1	ヒト皮膚細菌叢の系統組成について	19
2.3.2	ヒト皮膚細菌叢を用いた個人特定精度	19
2.3.3	OTU0001 とヒト皮膚環境メタデータの相関解析	20
2.4	考察	21
2.5	図表	25
第 3 章	個人特異的な細菌と、細菌のヒト常在因子の解明	30
3.1	緒言	30
3.2	材料と方法	31
3.2.1	16S rRNA 遺伝子配列データ取得と系統解析	31
3.2.2	細菌の属レベルでのヒト常在性の定量化	32
3.2.3	被験者メタデータの整形	33
3.2.4	各体部位における Specific の数と、被験者メタデータの相関解析 . .	34
3.2.5	細菌の生息環境データ取得	35
3.2.6	General と Drifter が持つ機能遺伝子データの比較解析	35
3.3	結果	37
3.3.1	各体部位にて観測される General, Specific, Drifter の属の数	37
3.3.2	Specific と被験者メタデータの関連性と環境への分布	37
3.3.3	ヒト常在細菌と非ヒト常在細菌の機能遺伝子の比較解析	38
3.4	考察	40
3.5	図表	45
第 4 章	総括と展望	53
	謝辞	55
	付録	56
	参考文献	140

第1章 背景

生物の皮膚、腸内、土壌、海洋など、地球上のほぼ全ての環境において、様々な種類の細菌が群集を形成している。それらは細菌叢と呼ばれ、細菌叢中にどのような細菌がどの程度存在するのかを示す系統組成は、16S ribosomal RNA (16S rRNA) 遺伝子系統解析によって明らかにすることができる。本手法は、異なる系統間で塩基配列の保存性が高いとされる細菌の 16S rRNA 遺伝子の特異的に増幅した後、次世代シーケンサーで解読し、得られた配列群を用いて系統組成の推定を行うものである。特に、ヒトを宿主とする細菌叢は、ヒトの健康に大きく関与していると考えられており、腸内細菌と病気を対象とした研究が多くなされている [1,2]。

ヒト皮膚についても、ヒト皮膚細菌叢がヒト皮膚の健康に深く関わっていることが報告されている [3-5]。ヒト皮膚細菌叢は、時系列によって変動することが確認されているが、その変動 (個人内での二年半の間に渡る変動) は個人間の差異に比べると小さいことも明らかとなった [6,7]。その性質から、ヒトを宿主とした細菌叢は個人を規定する情報を持つことが示唆されている。実際に、ヒト皮膚細菌叢を用いた個人特定手法の提案もなされており、3人の被験者の手と、その被験者の所持するキーボードやマウスから数十サンプルの細菌叢を採取し、細菌叢の系統組成を用いたオーナーの特定が可能であることが報告された [8]。ヒト皮膚細菌叢は法医学の分野で注目されている [9]。さらに近年の研究では、ヒト皮膚上の細菌をターゲットとした single nucleotide polymorphism (SNP) または short tandem repeat (STR) を用いて個人間のヒト皮膚細菌叢の差異を機械学習によって明らかにし、10人以上の被験者を高精度に分類できることが報告された [10,11]。その個人識別能力は、サンプリング期間が二年半以上という長期に渡って保持されていることが明らかとなった。これらの手法を用いた個人特定手法は非常に高精度なもので、ヒト皮膚において非常に優占な細菌 (*Propionibacterium* 属や *Staphylococcus* 属等) のみを対象としている。近年の研究では、より大規模な人数で、一回目に採取した細菌叢の系統組成を用いて二回目に採取した細菌叢 (一回目と二回目のサンプリングの期間は 30-300 日) の系統組成を用いた個人特定が試みられた [12]。結果、50人の被験者から取得した糞便を用いた場合、80%近い精度で個人を特定することができたが、61-85人から取得したヒト皮膚細菌叢 (e.g., antecubital fossa, retroauricular crease, and anterior nares) を用いた場

合、その個人特定精度は30%に満たない結果となった。以上の報告より、大規模な人数に対してヒト皮膚細菌叢の系統組成を用いた個人特定が可能であるか否かについては未だ明らかとなっていない。すなわち、どのような系統の存在が個人特定に寄与するのかを徹底的に議論する必要がある。

第二章の研究では、a) 高精度な新規個人特定手法の開発、b) 個人特定に寄与する細菌の解明、の二点を目的とした。本研究では我々が独自に採取したデータ (our dataset: 被験者 11 人) と、公開済みのデータ (public dataset: 被験者 89 人) を用いた。ヒト皮膚細菌叢として額の細菌叢のモデルとして選択した。額は、衣服や他の個人と接触する可能性が低く、常に外気へと暴露されている定常的な部位であることに加え、額と近い眉間の細菌叢の系統組成は比較的安定していることからモデルとして適していると考えられる [13]。本研究で開発した個人特定手法によって、ヒト皮膚における低い相対存在量を示す細菌の存在が、個人特定には重要であることが明らかとなった。また、個人特定を行う際に、個人を識別するために用いる学習データとして被験者一人あたりに複数サンプル用意することで高精度な個人特定を実現できることが明らかとなった。

第三章の研究では、ヒト細菌叢における個人特異的な細菌の存在についての解明と、細菌をヒト常在細菌たらしめる要因の解明を試みた。第二章のヒト皮膚細菌叢を用いた個人特定の研究では、ヒト皮膚細菌叢の中には、その被験者を特徴づける細菌が存在しており、それらの細菌がヒト皮膚ではない環境から由来する可能性を示唆した [14]。しかしながら、これらの細菌は、あるときから定着し常在し続けている可能性や、外環境からヒトの皮膚へと流入し続けている可能性などが考えられる。実際に、あらゆる環境には様々な系統の細菌が分布しており、その分布もまた系統によって異なることが知られている [15]。細菌の環境への分布を調査した研究では、細菌が単離された環境データをデータベースとして保持する方法 [16–18] や、培養を介さない 16S rRNA 遺伝子配列解析やメタゲノムショットガンシーケシングによって細菌の相対存在量を明らかにし、細菌叢サンプルの取得環境データをデータベースとして保持する方法 [18–20] によって各系統が生息する環境がある程度示されている。さらに、細菌は群集を形成しており、細菌によって構成される系統組成は環境によっても異なる。それらの細菌は群集単位で環境を往来することも報告されており、ある環境の細菌叢がどのような環境の細菌叢をどの程度含むのかについて確率モデルから推定する手法も開発された [21, 22]。特に、ヒトとヒトを取り巻く環境の細菌叢の研究がいくつかなされている。例えば、家の至るところ (キッチンの床やドアノブ) に存在する細菌叢は、そこに住む家族に特異的なものへと変化することや、犬を飼っている家族は飼っていない家族に比べて多くの系統を互いにシェアしていることが報告された [23, 24]。ヒト以外でも、飼育されている動物の腸内細菌は、野生に生息して

いる動物と大きく異なることや、その皮膚表面の細菌叢は生息する環境の細菌叢と深く関わっていることが明らかになっている [25,26]。以上の先行研究より、ヒト常在細菌がどのような環境に分布しているのかについて明らかにすることも個人特異的な細菌がヒトで観測される要因の解明として重要であると考えられる。この背景をもとに、公開されている大規模コホートデータを用いて、ヒト皮膚 (額, 手)、腸、舌にて観測される個人特異的な細菌と被験者の詳細なメタデータ (性別や日常的行動など) の関連、個人特異的な細菌の生息環境を調査した。各体部位におけるヒトで観測された細菌について系統組成から各細菌の属を、一般的に広い被験者で観測されるヒト常在細菌である General、特定の被験者で観測される個人特異的な細菌である Specific、一時的に観測された非ヒト常在細菌である Drifter、の3タイプへと分類した。結果、ヒト皮膚に存在する個人特異的な細菌は、被験者の振る舞いや状態に依存することを示唆することができた。個人特異的な細菌が生息する環境を調査した結果、ヒトでは無い環境に主に生息する細菌であることが明らかになった。

さらに、General と Drifter を比較解析し、細菌のヒトへの常在性に関連する因子の推定も試みた。ヒト皮膚では *Propionibacterium acnes* という種が系統組成の 90%以上を占めており、第二章で我々が採取したサンプルでも、アメリカで行われた研究でも、同じような結果が得られている [13,27,28]。ヒト腸内についても、北欧、ヨーロッパ、アメリカなど様々なヒトから採取された腸内細菌叢データから *Bacteroides* 属や *Prevotella* 属が決まって系統組成中で多く観測される [29–31]。このように、ヒト細菌叢は体部位によって決まった種類の系統が蔓延ることが明らかとなっている。しかしながら、何故これらの細菌がヒト常在細菌として系統が固定されているのかについては明らかになっていない。決まった細菌の系統が、各ヒト体部位で固定されている因子の一つとして細菌が持つ遺伝子が挙げられる。細菌が持つ遺伝子は、その細菌の生存や環境中における化合物代謝反応など多くの表現型に寄与していることが知られている。それらの遺伝子は具体的にどのような表現型として機能するのかについてデータベース化もされており、細菌が持つ遺伝子からヒト常在性の因子として働いている機能の推定も可能である [32]。近年の研究では、植物の根から採取した細菌の株と、根以外の箇所から採取された細菌の株を比較ゲノム解析することで植物との共生に寄与する遺伝子 (クラスター) の存在がいくつか明らかとなった [33]。このように、細菌が持つ遺伝子は環境にて常在するために非常に重要である。そこで、第二章で用いた大規模コホートから各体部位における一般的に観測される属と、一時的にしか観測されない属が持つ機能遺伝子の違いを比較解析から明らかにした。結果、ヒト皮膚などの環境に生息する細菌は、生息できない細菌と比較して異なる機能遺伝子をいくつか持っており、それらの機能遺伝子がヒトに生息することに重要であることも考え

られた。

第2章 ヒト皮膚細菌叢を用いた個人特定

2.1 緒言

ヒト皮膚細菌叢は個人を特定することに利用が可能なツールとして脚光を浴び、20 人以下の被験者でヒト皮膚細菌叢を用いた個人特定を試みた研究がいくつか報告された。しかしながら、ヒト皮膚細菌叢を用いた個人特定を 80 人以上の被験者に対して行った結果、30%にも満たない低い精度でしか個人を特定できないことが報告された。

本研究では、高精度な個人特定手法の開発を試みた。開発した個人特定手法では、個人を分類することができる要因として、ヒト皮膚細菌叢の系統組成の中でも低い相対存在量の細菌が重要であると考え、それらの細菌の差を考慮した Canberra 距離という指標を用いた。さらに、学習データとして一人の被験者由来のサンプルを複数用意した。実際に本研究では、1 年目 (first year) と 2 年目 (second year) に 11 人の被験者から各年度に 3 サンプルずつサンプリングを行った。本手法を用いることで、一年前に取得した 33 サンプルを学習データ (reference) とすることで、新たに取得した 33 サンプルの個人特定に用いるテストデータ (query) を 85%の精度で正しい被験者へと分類することができた。さらに本手法によって、公開済みの 89 人の被験者から 837 サンプル取得したヒト皮膚細菌叢データに対しても、leave-one-out 法による検証で 78%の精度で個人を特定することができた。

結果、i) 本研究の個人特定手法は 2 つのヒト皮膚細菌叢データに対して高精度に個人を分類することができること (Our dataset: 11 individual; Public dataset: 89 individual)、ii) 個人特定を行うには一人の被験者から複数サンプリングする必要があること、iii) 個人特定では系統組成中の少数の相対存在量の系統を用いることで精度を上げることができると、の三点が明らかとなった。

2.2 材料と方法

2.2.1 細菌叢サンプル採取

20代の男性10名と30代の女性1名の額の中央からサンプリングを行った。本サンプリングは、各被験者から1年目 (first year) に3回、2年目 (second year) に3回の計6回ずつ行った (表 2.S1)。サンプリングに際しては東京工業大学疫学研究等倫理審査委員会 (approval number: 2013003) の規定に基づき被験者に同意を得た。サンプリング手法を以下に示す。

1. 各被験者の手をアルコール消毒し、グローブを装着した。
2. 各被験者自身が無菌綿棒 (AS ONE) を用いて、額の中央 2cm×2cm を 50 往復こすりサンプリングを行った。
3. 1.5 µl チューブにサンプリングに使用した綿棒の先を折り入れた。
4. -20 °C で冷凍保存した。

2.2.2 メタデータ採取

滅菌綿棒を用いてサンプル採取を行った後、ヒト皮膚環境を示すものとしてヒト皮膚メタデータである水分量 (moisture)、酸性度 (pH)、油分量 (sebum) の計測を行った。計測機器は MULTI DISPLAY DEVICE MDD 4 (Courage+Khanzaka) を用いた。以下に詳細な計測手法を示す。

水分量の計測

MULTI DISPLAY DEVICE MDD 4 (Courage+Khanzaka) に付属のプローブ Corneometer ® CM815 (Courage+Khanzaka) を用いてサンプリング箇所近傍の水分量を計測した。付属のプローブの計測部である先端をキムタオルで十分に拭き取り、計測部を被験者の皮膚に垂直に押し当て、測定を行った。計測は3回試行し、その中央値を結果として扱った。

油分量の計測

MULTI DISPLAY DEVICE MDD 4 (Courage+Khanzaka) に付属のプローブ Sebumeter ® SM815 (Courage+Khanzaka) を用いてサンプリング箇所近傍の皮脂量を計測した。付属のカセットのプラスチックテープ部分を被験者の皮膚に一定力で5秒間押し当て、測定を行った。計測は3回試行し、その中央値を結果として扱った。

酸性度の計測

MULTI DISPLAY DEVICE MDD 4 (Courage+Khanzaka) に付属のプローブ Skin-pH-Meter ® PH815 (Courage+Khanzaka) を用いてサンプリング箇所近傍の酸性度を計測した。付属のプローブの計測部である先端を超純水で十分に洗浄し、計測部を被験者の皮膚に垂直に押し当て、測定を行った。計測は3回試行し、その中央値を結果として扱った。

2.2.3 16S rRNA 遺伝子の PCR 増幅

得られた抽出済みの DNA 溶液から真正細菌が持つ 16S rRNA 遺伝子の特異的に polymerase chain reaction (PCR) によって増幅させた。プライマーは 27F/338R (Forward 5' -AGAGTTTGGATCCTGGCTCAG-3' , Reverse 5' -TGCTGCCTCCCGTAGGAGT-3') を用いて、16S rRNA 遺伝子の V1-V2 領域を特異的に増幅させた。本プライマーはシーケンシングの際にサンプルを分類するのに必要なバーコードとタグが 5' 側に連結している。PCR の工程を以下に示す。

1. 0.2 ml PCR 用チューブに以下の組成で試薬をアプライした。

- Nuclease Free Water (Ambion)	18.75	µl
- 抽出済み DNA 溶液	20	µl
- 10X Taq Buffer (Takara)	5	µl
- dNTP mix (Takara)	4	µl
- Forward Primer (27F)	1	µl
- Reverse Primer (338R)	1	µl
- Ex Taq HS polymerase	0.25	µl

2. 各被験者自身が無菌綿棒 (AS ONE) を用いて、額の中央 2cm × 2cm を 50 往復こすりサンプリングを行った。

3. 1.5 μ l チューブにサンプリングに使用した綿棒の先を折り入れた。
4. -20 $^{\circ}$ C で冷凍保存した。

2.2.4 PCR 産物の精製

得られた反応後 PCR 溶液には、反応に使われなかったプライマーや、プライマー同士で増幅が生じてしまったプライマーダイマーなどの不純物が存在している。そのため、目的とする PCR 産物の精製が必要である。精製には E-gel Size Select 2% (Thermo Fisher Scientific) を用いており、以下にその工程を示す。本工程では、同時に Negative control で PCR 産物が作成されていないことを確認し、以降の操作を行った。

1. E-gel Size Select 2% (Thermo Fisher Scientific) の上段ウェルに反応後 PCR 溶液を 22.5 μ l アプライした。
2. マーカー用レーンに 50 塩基 DNA ladder (Thermo Fisher Scientific) を 10 μ l (25 ng/ μ l) アプライした。
3. 回収用ウェルには Nuclease free water (Ambion) を 25 μ l アプライした。
4. iBase (Thermo Fisher Scientific) に E-gel Size Select 2% (Thermo Fisher Scientific) をセットし、電気泳動を開始した。
5. 目的 PCR 産物のバンド (約 400 塩基) を確認し、目的産物のバンドが回収用ウェルに入る直前で、回収用ウェルに Nuclease free water (Ambion) を回収用ウェル一杯に補填した。
6. 下段の回収用ウェルから生成済み PCR 産物として 20 μ l 回収した。

2.2.5 PCR 産物の定量

得られた精製済み PCR 産物はシーケンサーに投入する際に、詳細な濃度が必要となってくる。今回の定量では、Qubit 2.0 Fluorometer (Thermo Fisher Scientific) を用いて定量化を行った後、BioAnalyzer 2100 (Agilent Technologies) を用いて定量・定性分析を行った。

Qubit 2.0 Fluorometer (Thermo Fisher Scientific) の使用方法については「QubitTM Assays QUICK REFERENCE CARD」に従った。Working Solution は 199 μ l、User Sample は 1 μ l で行った。

BioAnalyzer 2100 (Agilent Technologies) の使用方法については「バイオアナライザ単マニュアル 2010 年 8 月作成 ver.01.02」を参照した。目的 PCR 産物である 16S rRNA 遺伝子アンプリコンのリード長 400 塩基付近にピークが確認できたら、そのサンプルを 26 pM に Nuclease Free Water (Ambion) を用いて希釈した。

2.2.6 16S rRNA 遺伝子アンプリコンシーケンス

シーケシングには Ion PGMTM (Thermo Fisher Scientific) を用いた。今回行ったシーケシングは、配列を 400 塩基ほどの長さで読むことができる Ion PGMTM Sequencing 400Kit (Thermo Fisher Scientific) と Ion PGMTM Template OT2 400Kit (Thermo Fisher Scientific) を用いた。さらに、配列を出力サイズが異なる 3 種類のチップのうち Ion 316 Chip Kit v2 (Thermo Fisher Scientific) を用いた。シーケシングを行う際の操作については、プロトコル (PN:MAN0007274 Rev. 2.0 Data16 May 2013) に従った。

シーケシングでは、26 pM に調整されたサンプルを等量混合して作成された DNA 溶液を 25 μ l 用いた。

2.2.7 16S rRNA 遺伝子系統解析

Ion PGMTM (Thermo Fisher Scientific) によって取得できた 16S rRNA 遺伝子アンプリコンデータを用いて各サンプルに存在する 16S rRNA 遺伝子の系統組成推定を行った。さらに、public database の 16S rRNA 遺伝子アンプリコンデータに対しても 16S rRNA 遺伝子の系統組成推定を行った (図 2.S1)。

2.2.8 97% OTU の系統組成テーブル作成

Our dataset について

1. Trimmomatic version 0.33 [34] を用いて、パラメーターを ‘LEADING:17 TRAILING:17 AVGQUAL:25 MINLEN:200’ で配列をクリーニングした。
2. tagcleaner0.12 [35] を用いて、パラメーターを ‘-mm5 3 -mm3 3 -nomatch 3’ で、プライマー配列のミスマッチ 3 塩基以下で所持するリードのみ採用し、そのプライマー配列を除去した。
3. USEARCH [36] のコマンドからパラメーターを ‘-minseqlength 200’ でプライマー除去後の長さが 200 bp 未満の配列を除去した。

4. 1つにまとめたプライマー除去済みのリードを、usearch7.0.1001.i86linux32の sort-bylength オプションを用いてリード長で降順に並び替えた。
5. ソート済みのリードを usearch7.0.1001.i86linux32 の UCLUST を用いて Identity $\geq$ 98%, Coverage $\geq$ 90%でクラスタリングし、UCLUST による OTU (Operational taxonomic unit) である 98% UCLUST OTU を作成した。
6. 98% OTU の代表配列を usearch7.0.1001.i86linux32 の UCHIME [37] を用いて denovo mode と reference mode でキメラ判定を行った。reference mode で判定を行う際、キメラ配列検出用データベースである gold.fa を参照した。
7. キメラ除去済みの 98% OTU の代表配列同士を all-against-all で (全配列をクエリ、全配列をデータベースとして)、BLASTN [38] を e-value $\leq 1e-8$, b = 74,955, v = 1 で行いアライメントした。
8. アライメントされた 98% OTU の代表配列同士を Identity $\geq$ 97%, Coverage $\geq$ 0.8 で mcl-14-137 [39] を用いてクラスタリングし、97% OTU を作成した。
9. 各サンプル中のリードが何本各 97% OTU にアノテーションされたかを示す 97% OTU 組成テーブルとして出力を行った。なお、得られた 97% OTU 組成テーブルは各サンプルの総リード数で、各サンプルに存在する 97% OTU の数値を割った相対存在量として扱った。
10. 作成された 97% OTU の中で最もサイズの大きい 98% OTU の代表配列を、97% OTU の代表配列とした。
11. 97% OTU の代表配列を、BLASTN (Identity $\geq$ 97%, Coverage $\geq$ 80%) を用いて、The All-Species Living Tree Project (LTP) が提供する 16S/23S rRNA 遺伝子配列データベース (https://www.arb-silva.de/fileadmin/silva_databases/living_tree/LTP_release_128/LTPs128_SSU/LTPs128_SSU_unaligned.fasta) [40] に登録されている Type strain の配列へとアノテーションした。閾値以下であった OTU については Unclassified とした。

Public dataset について

1. European Nucleotide Archive (<https://www.ebi.ac.uk/ena>) から、ヒト額由来 16S rRNA 遺伝子配列データの public dataset である ERP00512 [7] を fastq 形式でダウンロードした。

2. ダウンロードしたデータから各サンプルに紐づくメタデータ情報 (SampleID, BodySite, PersonalID, WeeksSinceStart) [7] の情報から 6 タイムポイント以上のサンプルが存在する被験者 89 人を選択した。
3. Trimmomatic version 0.33 [34] を用いて、パラメーターを ‘LEADING:17 TRAILING:17 AVGQUAL:25 MINLEN:75’ で配列をクリーニングした。
4. 97% OTU を作成するため、USEARCH 10.0.240 にインストールされている UPARSE [41] をデフォルトパラメータで全配列に対して行った。
5. 各サンプル中のリードが何本各 97% OTU にアノテーションされたかを示す 97% OTU 組成テーブルとして出力を行った。なお、得られた 97% OTU 組成テーブルは各サンプルの総リード数で、各サンプルに存在する 97% OTU の数値を割った相対存在量として扱った。

2.2.9 多様性指数の計算

得られたサンプルの 97% OTU 組成の多様性について評価を行う。多様性とは、すなわち、ある細菌叢を占める細菌の種がどれだけ多く均等に存在しているかを数値として示すものである。今回、多様性指数として Shannon diversity index (SDI) を用いた [42]。SDI は、R のパッケージとして公開されている VEGAN 2.4-2 [43] を用いて計算した。

2.2.10 距離法を用いたサンプル間比較

得られたサンプル間の 97% OTU 組成の個人内の変動と個人間の違いを明らかにする手法として、系統組成間の類似度を距離として計算した。しかしながら、ヒトの額では *Propionibacterium* 属、*Corynebacterium* 属、*Staphylococcus* 属が優占しており、ユークリッド距離のような優占種の相対存在量に影響を受けやすい距離指標は適していないと考えられる。そこで、系統組成中で小さい相対存在量である細菌の差も考慮することができる Canberra 距離を用いてサンプル間の比較を行った。Canberra 距離の計算では、R の dist 関数を用いた。ここで、サンプル X の系統組成を $(X_1, \dots, X_n)$ 、サンプル Y の系統組成を $(Y_1, \dots, Y_n)$ とするときの Canberra 距離の式は次の形で与えられる：

$$D_{canberra}(X, Y) = \sum_{i=1}^n \frac{|X_i - Y_i|}{|X_i| + |Y_i|} \quad (1)$$

2.2.11 個人特定手法について

ヒト皮膚細菌叢を用いた個人特定手法を考案した。今回行った個人特定とは、既知の個人由来ヒト皮膚細菌叢の系統組成を reference とし、未知の個人由来ヒト皮膚細菌叢の系統組成である query が由来する個人の特定を行うものである (図 2.1)。手法の概要を以下に示す。

1. 未知の個人由来サンプルとしてサンプル X を query として用意する。
2. 既知の個人由来サンプルを reference として用意する。
3. query と、各個人の reference との平均 Canberra 距離を計算する：

$$D_{individual}(q, r_{individual}) = \frac{\sum_{i=1}^n D_{individual}(q, r_i)}{n} \quad (2)$$

$D_{canberra}(q, r_i)$: query と、ある個人の i 番目のサンプルとの Canberra 距離
 q : query の系統組成
 r_i : ある個人の i 番目の系統組成
 n : ある個人の個人由来サンプル数

4. query が最も近い平均距離を示した個人を、query の由来する個人とする。
5. query が、その個人由来であった場合を True Positive (TP) とし、その被験者由来でなかった場合を False Positive (FP) とした
6. 個人特定の精度を測るためのスコアとして *Accuracy of personal identification* を計算した：

$$Accuracy\ of\ personal\ identification = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (3)$$

N_{TP} : TP と判断された query の数
 N_{FN} : FN と判断された query の数

2.2.12 拡張版個人特定手法について

開発した個人特定手法が、reference に query が由来する被験者が含まれていない場合にも対応できる拡張版も開発した。Canberra 距離に閾値を設けることで、True Negative (TN) と False Positive (FP) を決定することができる。すなわち、由来する被験者が reference 中に含まれるか否かを答えとして出すことができる。拡張版個人特定手法の精度検証のた

め、サンプル全体からひとつ取り出し query とする leave-one-out 法を用いた。検証では主に二通りの方法、1) query が由来する被験者の reference と閾値で定められた距離よりも近い場合を TP とし、それ以外であれば FN もしくは FP する方法、2) query の由来する被験者が含まれない reference を作成し個人特定を行い、いずれの被験者とも閾値以上の距離が計算された場合を TN とし、それ以外を FP とする方法、を用いた。この検証を public dataset に対して、Canberra 距離の閾値 (14000 から 18000 まで 100 ずつ増加) で行った。

2.2.13 個人特定に関連したパラメータについて

個人特定における reference 数の影響

個人特定における reference 数 (number of reference: 1 to 5 sample) が、個人特定精度に与える影響を精査した。reference 数の精査では、同じ被験者から複数時系列採取している public dataset を用いた。以下のテストを各 reference 数ごとに 100 回行った。

1. 各被験者から、reference としてランダムに任意の数だけサンプルを取得し、残ったサンプルから query をランダムにサンプルを一つ選択する。
2. 個人特定精度を計算する。

個人特定における時系列の影響

個人特定における reference と query 間の時系列が、個人特定精度に与える影響を精査した。時系列の精査では、同じ被験者から年度内・年度間で複数サンプル採取している our dataset を用いた。時系列としては、reference と query が同じ年度 (1st $\rightarrow$ 1st, 2nd $\rightarrow$ 2nd)、異なる年度 (1st $\rightarrow$ 2nd, 2st $\rightarrow$ 1st) の4つの組み合わせを作成した。以下のテストを各時系列の組み合わせごとに 100 回試行した。

1. 各被験者から、reference としてランダムにサンプルを一つ取得し、残ったサンプルから query をランダムにサンプルを一つ選択する。
2. 個人特定精度を計算する。

個人特定における少数種の影響

細菌叢において相対存在量が著しく低い細菌すなわち少数種は、コンタミネーションやシーケシングエラーとしてしばしば除かれることがある。これらの少数種も個人特定において重要であることを考え、少数種の個人特定に対する影響を明らかにした。以下の2つの検証を行い、少数種の個人特定精度に与える影響を精査した。

1. 少数種の存在の有無を指定する閾値 (1e-01, 1e-02, 1e-03, 1e-04, 1e-05, 1e-06, and 0) において、それ以下の細菌をゼロとして考え、個人特定精度を検証した。
2. シーケンス深度 (100, 1000, 10000, 20000) における OTU の系統組成を元の系統組成から GUniFrac [44] を用いて各シーケンス深度で取得した系統組成を用いて個人特定精度を検証した。GUniFrac による各シーケンス深度へのダウンサンプリング (rarefy) はランダムなため、各シーケンス深度について 10 回ずつ試行した。

2.2.14 個人特定に重要な系統 personal OTU の探索

各被験者の特異性に寄与する OTU として personal OTU を定義した。personal OTU の定義として、1) ある被験者にて全ての時系列で観測されること、2) OTU の相対存在量の中央値が、その被験者由来サンプル (6 sample) が、他の被験者由来サンプル (60 sample) よりも二倍以上高いこと、の2つの条件を満たす OTU とした。さらに定義した personal OTU が由来する環境の推定をおこなった。personal OTU が分類された type strain が単離された環境を調査し、9 個のカテゴリ (human skin, human-related, air, soil, plant, water, other, unknown, and unclassified) へと単離環境を分類した [14]。

2.2.15 ヒト皮膚メタデータと OTU0001 の相対存在量の相関解析

ヒト皮膚メタデータ (moisture, pH, sebum) と最も相対存在量が高かった OTU0001 の相関を計算した (図 2.S2)。相関解析では、Spearman の相関係数を計算した。今回、OTU0001 のみを相関解析に用いた理由として、OTU0001 以外の OTU は相対存在量が非常に低いため、OTU0001 の変動に大きく左右されることを考慮したからである。

2.3 結果

2.3.1 ヒト皮膚細菌叢の系統組成について

今回、11人の被験者から6タイムポイントずつ二年間に渡りサンプリングし、系統組成を明らかにした(図2.2)。66サンプル全体で、最終的に7,005,342本のハイクオリティ配列を取得できた(表2.S2)。皮膚の種レベルでの系統組成を明らかにするために、全ての配列を相同性97%で4,155 OTUsへとクラスタリングした。659 OTUsはいずれかのtype strainへと、492 OTUsは複数のtype strainへと、残りの3,004 OTUsはいずれのtype strainにも分類されなかった。特に、OTU0001 (*Propionibacterium acnes*), OTU0002 (*Staphylococcus*), OTU0003 (*Corynebacterium tuberclostericum / accolens*), OTU0004 (*Staphylococcus*) が全てのサンプルで優占する系統として観測された。

さらに、系統組成の多様性を示すことができる Shannon diversity index の計算を行った。ほとんどの被験者において、全てのタイムポイントにて低い Shannon diversity index の値が計算された(図2.2A)。すなわち、ほとんどのヒト皮膚細菌叢における系統組成は *P. acnes* のような優占種により多様性が低く計算されていた。しかしながら、被験者BとCは他の被験者よりも *P. acnes* に分類される OTU0001 の相対存在量が低いことから、他の被験者に比べて高い値の Shannon diversity index が計算された(図2.2B)。

ヒト皮膚細菌叢の経時的な安定性と、個人間の差異を明らかにするためサンプル間の差異を数値化することができる Canberra 距離を用いた。比較は、被験者それぞれにおける、異なる被験者由来のサンプル同士 (inter-individual)、同じ被験者由来のサンプルで異なる年に採取されたサンプル同士 (inter-year)、同じ被験者由来のサンプルで同じ年に採取されたサンプル同士 (intra-year) の三群で行った(図2.2C)。結果、全ての被験者において inter-individual のほうが、inter-year と intra-year に比べて高い Canberra 距離の値を示すことが明らかとなった (Wilcoxon の順位和検定, $p \leq 0.05$)。

2.3.2 ヒト皮膚細菌叢を用いた個人特定精度

今回取得したヒト皮膚細菌叢サンプル (66 sample from 11 individual) を用いた個人特定手法の開発を試みた(表2.1)。個人特定手法の精査には leave-one-out 法を用いた。leave-one-out 法は、個人特定に用いるテストデータとしてサンプルを一つだけ取り出し、残りのサンプルを学習データとし、テストデータが由来する個人を特定する手法である。結果、95% (63/66) の精度で個人特定を行うことができた。これは、66サンプル中63サンプルを正確に分類することができたことを示す。さらに、first year のサンプルを用いて、

second year のサンプルを 85% (28/33) の精度で分類することができた。本個人特定手法の汎用性を試すため、独自に取得したヒト皮膚細菌叢データではない、公開済みの大規模ヒト皮膚細菌叢データ (837 sample from 89 individual) についても個人特定を行った。結果、78% (663/837) の精度で個人特定が可能であることがわかった。

さらに、reference としたサンプルと、query としたサンプルの時系列の個人特定精度に対する影響を明らかにするため、同じ年度同士で個人特定した場合と異なる年度同士で個人特定した場合の個人特定を行った。結果、同じ年度同士の場合の方が異なる年度同士よりも高い精度で個人特定可能であることが明らかとなった (図 2.3A)。

次に、reference の数が個人特定の精度に与える影響について精査した。結果、個人特定に用いる reference の数が多いほど個人特定精度が高くなることが明らかとなった (図 2.3B)。

個人特定に用いるサンプルは細菌叢データであり、その細菌叢を明らかにするために用いられる配列の数や、サンプル内に存在する細菌の中でも極端に少なく存在する細菌を解析から除外する閾値の設定を考慮することが重要である。その配列の数や、系統組成内に存在する細菌の割合の閾値の影響についても議論をおこなった。実験では、各配列の数 (100, 1000, 10000, 20000) を各サンプルにおいてランダムサンプリングを行った後、個人特定精度を算出した。結果、配列の数が多いときの方が個人特定精度が高くなる傾向にあることが明らかとなった (図 2.3C)。系統組成における細菌の割合の閾値については、閾値が低いときまたは閾値を設けない方が高い個人特定精度を示した (図 2.3D)。

個人特定手法で用いている距離法は Canberra 距離であるが、その他の距離手法との比較も行った。結果、Bray-Curtis 距離と Jaccard 距離よりも Canberra 距離の方が高い精度で個人特定が可能であることが明らかとなった (図 2.S3)。

一方、本研究の個人特定手法は reference の中に query と同じ被験者が含まれていない場合には用いることができない。拡張版として、個人特定において計算される Canberra 距離の閾値を設けることで、query の被験者が含まれているか否かを判定するアルゴリズムを設けた。本手法では、public data を用いており、大体 Canberra 距離の閾値を 16,000 で用いることで、TN と TP を 60% の精度で分類することが可能であることがわかった (図 2.S4)。

2.3.3 OTU0001 とヒト皮膚環境メタデータの相関解析

ヒト皮膚細菌叢がヒト皮膚の moisture, pH, sebum に何らかの影響を与えると考えた。ヒト皮膚に影響を与える因子として、今回のヒト皮膚叢で最も優占していた OTU0001 (*P.*

acnes) に着目した (図 2.S2)。OTU0001 の相対存在量と、ヒト皮膚環境メタデータ (油分量, 水分量, 酸性度) の相関関係を Spearman の相関係数を用いて明らかにした。結果、水分量について、被験者 A, G, K では 0.5 以上の相関係数が計算された。一方、被験者 B, E では -0.5 以下の相関係数が計算された。油分量については、ほとんどの被験者で負の相関関係を示した。特に、被験者 C, D については強い負の相関関係を示した。被験者 C については、pH についても強い負の相関関係が確認された。

2.4 考察

本研究では被験者 11 人の額から 2 年間に渡ってヒト皮膚細菌叢サンプルを取得した。さらに、公共データベースから 89 人の額から取得された公開済みヒト皮膚細菌叢サンプルの 16S rRNA 遺伝子配列データも取得した。これらのデータを用いて、ヒト皮膚細菌叢を用いた個人特定手法の開発を行った。

取得した 66 サンプルの系統組成では、*Propionibacterium* 属が最も優占していた。過去の先行研究では、ヒト皮膚に存在する細菌の属として、*Propionibacterium* 属, *Corynebacterium* 属, *Staphylococcus* 属が報告されている [13, 28, 45–47]。同様に、これらの 3 属は本研究で取得したサンプルでも観察された。しかしながら、ヒト皮膚細菌叢解析に用いるプライマーの一種である 515F/806R は、*P. acnes* の相対存在量を低く推定することが報告されている [48]。本プライマーは公共データベースから取得した 16S rRNA 遺伝子解析データで用いられており、本データからヒト皮膚細菌叢の系統組成を明らかにした研究において、*Propionibacteriaceae* 科が低い相対存在量で推定されている。以上の理由より、本研究で取得した 66 サンプルと統合することは不可能であると考えられる。

取得した 66 サンプルのヒト皮膚細菌叢データから、被験者それぞれにおける、異なる被験者由来のサンプル同士 (inter-individual)、同じ被験者由来のサンプルで異なる年に採取されたサンプル同士 (inter-year)、同じ被験者由来のサンプルで同じ年に採取されたサンプル同士 (intra-year) の三群を用いた比較を行った。inter-year と比較して、intra-year (2, 3 ヶ月程度) であれば安定していることが明らかとなった。そして、inter-year (一年間以上) であれば少し変動することが観察された。さらに、ヒト皮膚細菌叢の経時的な安定性と個人間の差異については先行研究でも報告されている。このように、ヒト皮膚細菌叢は個人特定に用いることが可能であると考えられる。実際に、ヒト皮膚細菌叢を用いた個人特定手法の開発を行った。

結果、1) 10 人程度であれば、ヒト皮膚細菌叢は年をまたいだ個人特定に用いることが可能であること、2) 2, 3 ヶ月程度の時系列であれば、約 90 人の被験者を分類することも

可能であること、の二点が明らかとなった。しかしながら、100 万人規模の人数から被験者を特定することは、*FP* の極端な増加により正確性が落ちることが予想される。

サンプルが由来する被験者を正確に分類することが可能な個人特定手法の開発過程において、1) 低い相対存在量を示す系統の情報が個人特定に重要であること、2) 同じ被験者由来の複数のヒト皮膚細菌叢サンプルを学習データに用いることが重要であること、の二点が明らかとなった。個人特定手法ではヒト皮膚細菌叢の系統組成に含まれる低い相対存在量の系統を解析から排除するための閾値を設定していない。いくつかの研究において、低い相対存在量の系統はコンタミネーション等の影響であると定義され、解析から排除されることがある [49–51]。しかしながら、本研究の解析では低い相対存在量を示す系統を漏れなく用いることで高い精度の個人特定が可能になることを示した。同様に、シーケンスの深度についても深い方が高い精度を示すことが明らかとなった。これらの結果から、ヒト皮膚細菌叢を用いた個人特定では低い相対存在量の系統の存在が重要であることが示唆された。このように、低い相対存在量の系統を除く閾値設定する場合は慎重に議論を行わなければならない。さらに、個人を定義するために用いるサンプルは複数サンプル取得する必要があることが示唆された。しかしながら、法医学の分野等で複数サンプルを取得することは困難である。複数サンプルの取得が困難な場合、取得する 16S rRNA 遺伝子配列データのシーケンス深度を深くすること、不用意な閾値を設けないことが重要である。さらに、ヒト皮膚細菌叢のサンプリングの際に、場所をずらして複数サンプル取得することも解決策の一つとして挙げられる。

本研究の個人特定手法の拡張版として、サンプルを由来する被験者に特定したい場合において、学習データにその被験者が含まれているか否かを判断するアルゴリズムを追加した。本手法では、大体 60% の精度で *TN* と *TP* を分類することができた。オリジナルの個人特定手法と比べて、低い精度の *TP* を示した。

ヒト皮膚細菌叢を用いた個人特定が可能であることは、個人のヒト皮膚細菌叢が個人特異的な細菌を有していることを示唆する。先行研究において、いくつかの系統はヒト皮膚細菌叢と、それを取り巻く環境の細菌叢との間で共有されていることが報告された [23,24,52]。今回、個人特有であると考えられる系統を personal OTU として定義し、それらの系統が単離される環境を調査した。これらの personal OTU はヒト皮膚のみではなく、様々な環境 (土壌, ヒト腸, 植物等) から単離されるような細菌であることが明らかとなった (表 2.S3)。すなわち、ヒト皮膚細菌叢は、それを有する被験者を日常的に取り巻く環境から細菌を取り込んでいる、もしくは、過去に一度取り込んだ細菌がヒト皮膚常在細菌として定着していることが考えられる。例えば、いくつかの personal OTU は水環境と関連する環境から単離されるような細菌であった。近年の報告でも、活発に屋外で活動する被験者

は個人特異性が高い OTU (indicator OTU) を多く有することが明らかとなった [45]。本観察結果は、これからヒト細菌叢に含まれる個人特異的な細菌の源を明らかにする一助となると考えられる。

さらに、本研究では、ヒト皮膚細菌叢で最も多く観測された OTU0001 (*P. acnes*) の相対存在量と、ヒト皮膚が持つ三つの環境メタデータ (油分量, 水分量, 酸性度) との関連性を調査した。結果、正の相関関係を示す被験者と負の相関関係を示す被験者が存在した。このように、OTU0001 の相対存在量と、ヒト皮膚が持つ三つの環境メタデータの相関関係は被験者ごとに異なることが明らかとなった。先行研究では、頬における *Propionibacterium* 属の量の増加は、油分量の増加と相関があることが明らかとなっている [53]。一方、本研究では、ヒト皮膚に存在する *Propionibacterium* 属の遺伝子は個人間で異なると考えられる。実際に、先行研究では、個人で有する *P. acnes* の遺伝子は大きく異なっており、その違いが痤瘡の発症に影響することが示唆されている [27,54]。しかしながら、16S rRNA 遺伝子解析では *P. acnes* を株レベルで分類することが困難である。今後、*P. acnes* とヒト皮膚環境メタデータの相関関係を明らかにするために、株レベルで解析可能な手法を用いる必要がある。

本研究では一年間という時系列、さらには大規模な被験者数において高い精度で個人特定可能な手法を開発することができた。しかしながら、本研究で独自に採取したデータと、公共データを用いて別々に議論を行っているため、時系列が長く、性別が男女共に多く、被験者が多い場合において高精度な個人特定が可能であるか否かについては結論付けることができない。法医学などの分野での応用のためにも、これら時系列, 性別, 被験者数がより拡張されたデータに対して精度検証を続けていく必要がある。さらに、本研究の個人特定手法は、新しいサンプルを取得した際に、再び配列のクラスタリングを行う必要があり、この点も応用にむけて改善すべきである。考えられる改善案として、配列を直接系統へと分類するデータベースを用いる方法がある。現在、細菌と古細菌の 16S rRNA 遺伝子配列データベース (SILVA ribosomal RNA gene database, Greengenes Database, and Ribosomal Database Project) が開発・公開されている [55–57]。これらを用いることで再び配列をクラスタリングする必要性はなくなる。しかしながら、データベースに存在しない配列は分類することができないため、この案も妥協案でしかない。

このように、ヒト皮膚細菌叢解析では経時的な変動と相対存在量が低い細菌も考慮すべきであることが結論付けられる。特に、本研究の個人特定手法においては相対存在量が低い細菌は無視することができない情報であった。しかしながら、本研究の個人特定手法も完全な精度で個人特定できる手法では無いため、手法の拡張もしくは全く異なった手法と組み合わせながら応用に向けて改善していく必要があるだろう。本研究では非常に小さい

コホートで議論をしているが、個人のヒト皮膚細菌叢に常に存在する環境由来であると考えられる細菌の存在を確認できたことから、ヒト皮膚細菌叢が被験者が訪れたことのある場所のログとして扱える可能性としても期待できる。次の研究課題として、これらの環境由来であると考えられる系統が、どのような環境に分布しており、被験者のどのような行動に起因して、個人特異的な細菌として定着しているのかについて明らかにしていく必要がある。

2.5 図表

表 2.1. 個人特定の精度

Data	Query	Reference	Number of individuals	Number of samples	Accuracy
Our data	all samples	all samples	11	66	0.95
Our data	first year	second year	11	33	0.85
Our data	second year	first year	11	33	0.85
Public data	all samples	all samples	89	837	0.78

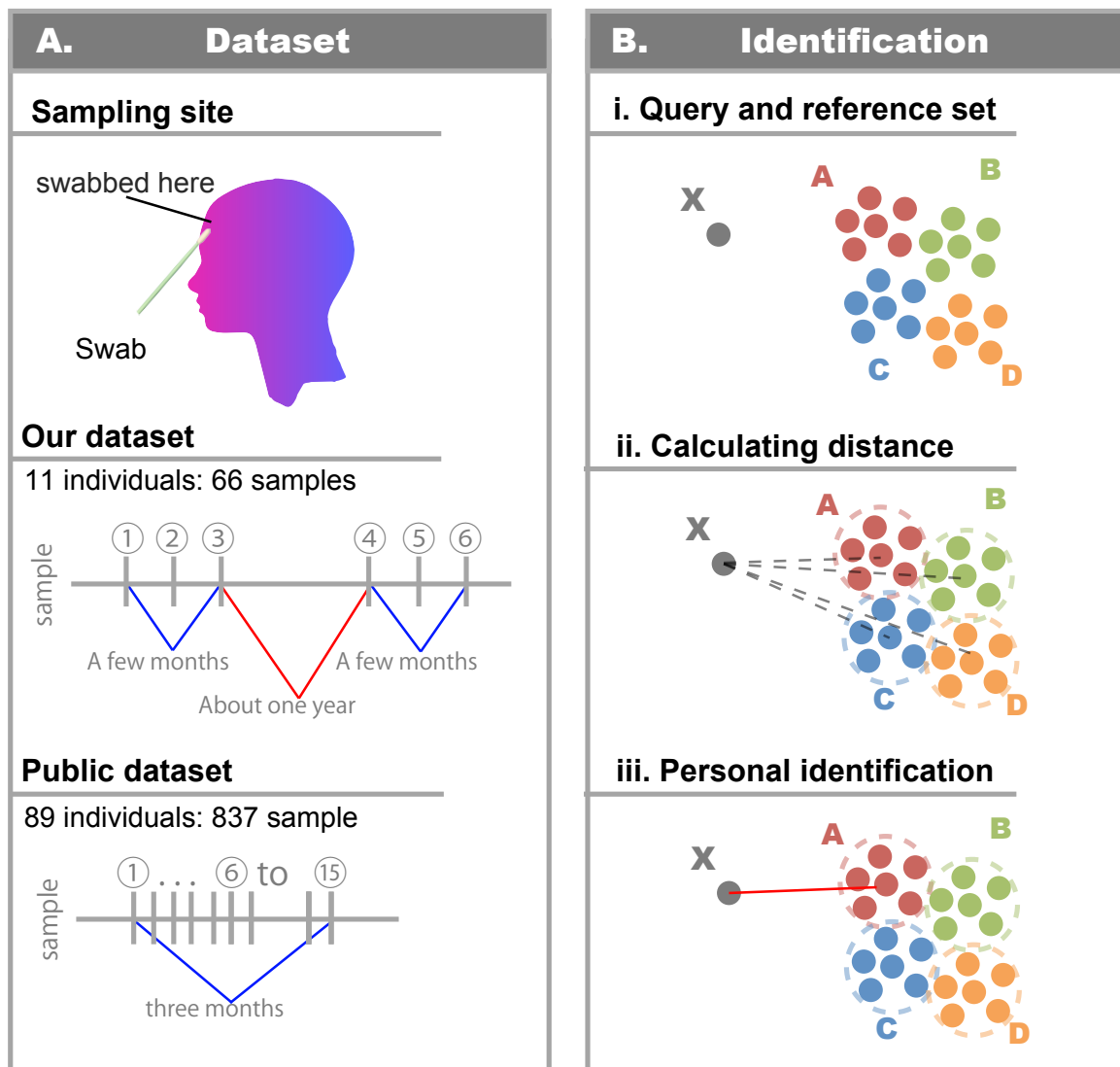


図 2.1. サンプルの概要と個人特定手法の説明

A) 本研究で用いたサンプルの概要を示す。Our dataset は、2 年間の各 3 タイムポイントで 11 人の被験者の額から採取されたサンプル。Public dataset は、3 ヶ月以上かけて 6–15 の時点で額から採取されたサンプル (ERP00512) [7]。各サンプルについて、16S rRNA 遺伝子配列アンプリコンを用いて微生物群集組成を分析した。 B) 個人特定手法の概要を示す。 i) 未知の被験者由来サンプル X が query であり、既知の被験者由来サンプルが reference として定義されている。 ii) query と reference との間の平均の Canberra 距離が計算される。 iii) サンプル X を距離に従って reference に割り当て、由来する被験者がサンプル X の真のオーナーであるか否かを評価する。

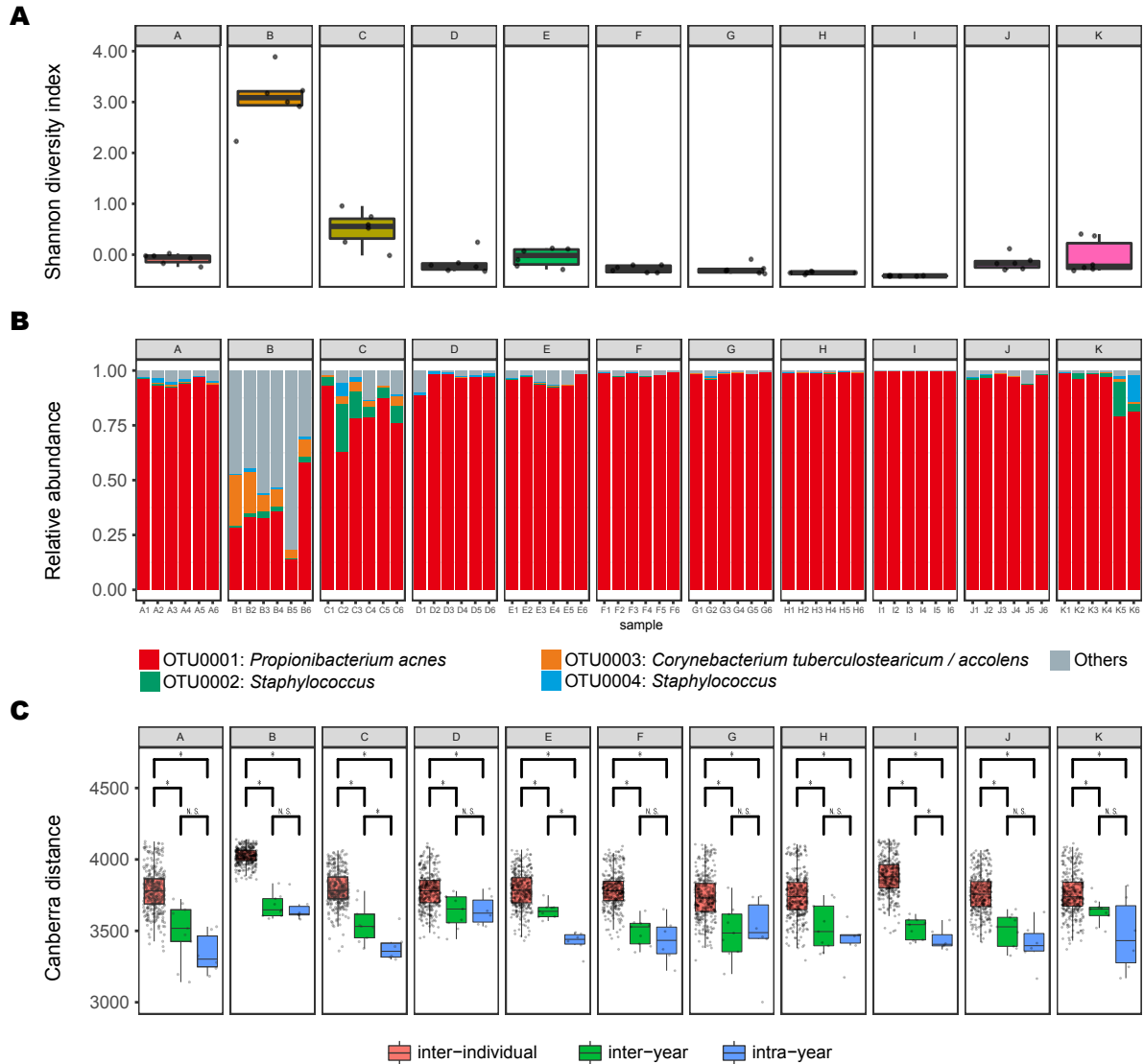


図 2.2. Our dataset における被験者 11 人の細菌叢

A) 被験者 A から K の 6 つのタイムポイントにおける Shannon diversity index を示す。
 B) 被験者 A から K の 6 つのタイムポイントにおけるサンプルの OTU の相対存在量を示す。各被験者について、最初の 3 列は一年目のサンプル、残りの 3 列は二年目のサンプルを示す。
 C) 各点は、3 つの組み合わせによって計算されたサンプル間のキャンベラ距離を表す。

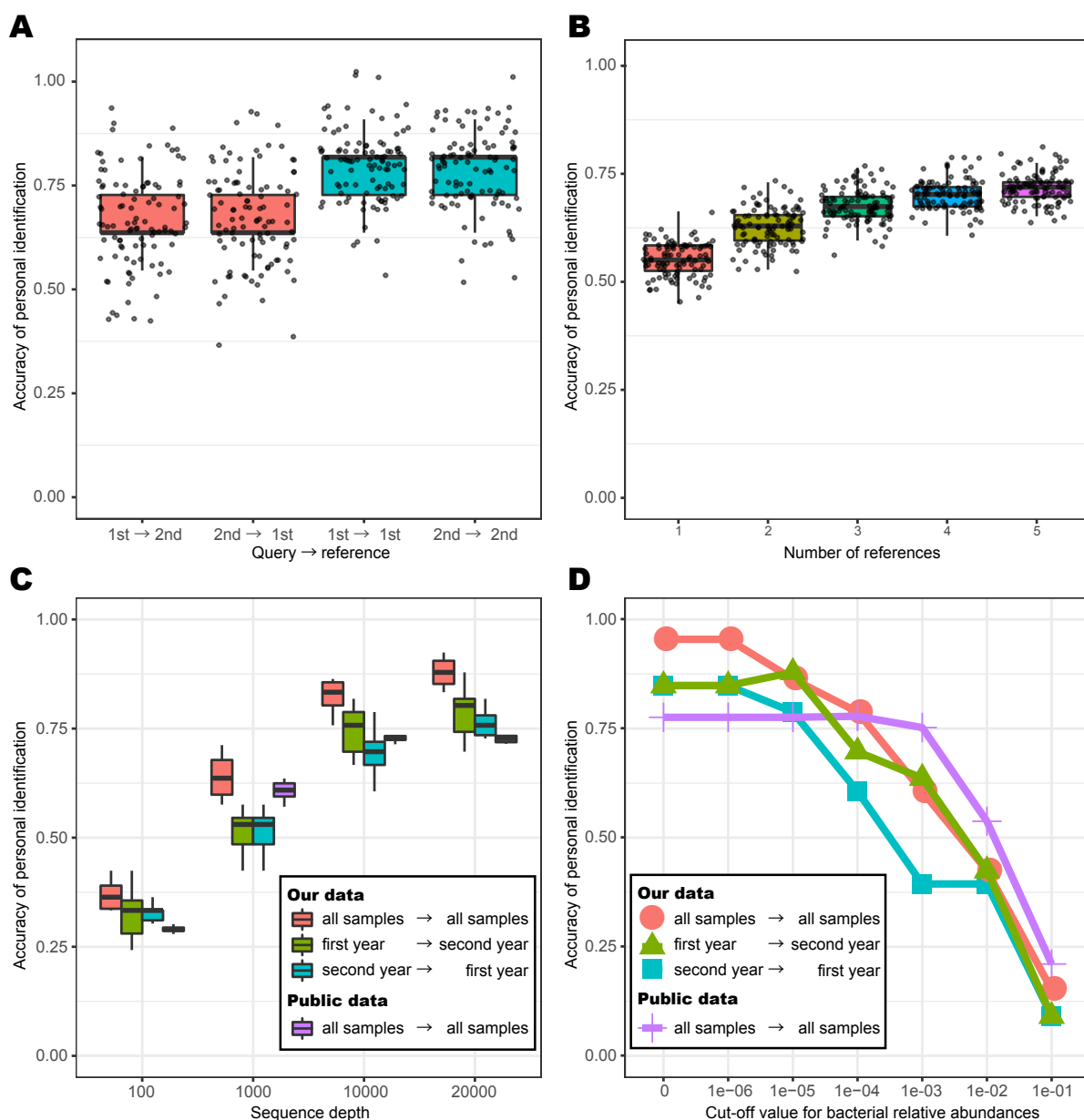


図 2.3. 個人特定精度に影響を及ぼす要因の評価

A) query と reference が同じ年および異なる年の場合における個人特定精度を示す。B) 被験者一人あたりの reference の数 (1–5 samples) の場合における個人特定精度を示す。C) 個人特定精度に対する細菌の相対存在量に対する cut-off 値の影響。D) 個人特定精度に対する細菌の相対的存在量に対する read depth (配列読み取り深度) の影響。

第3章 個人特異的な細菌と、細菌のヒト常在因子の解明

3.1 緒言

前章の研究では、ヒト皮膚に存在する個人に特異的な細菌はヒトではない外環境由来であることを示唆した。しかしながら、個人を特徴づけると定義した細菌の種の基準株の単離環境が外環境であることが示されたのみであり、実際に個人特異的に存在する種がヒト皮膚以外の外環境由来であることを結論付けることは困難である。

本章では、ヒトの額以外の体部位にも着目した。そして、個人を特徴づけるとされる細菌の存在に被験者の日常における振る舞いが関連すると仮定し、個人特異的な細菌と被験者メタデータの関連性について解析を行った。さらに、様々な環境からサンプリングされた 30,000 サンプルに及ぶ細菌叢の系統組成データから、各細菌の属が生息する環境が推定されたデータベース Latent environment allocation (LEA) を用いて、個人を特徴づけるとされる属の生息環境の推定も試みた。本章では、大規模かつ被験者メタデータが充実した公共の 16S rRNA 遺伝子配列データを用いた [7]。本データは、約 90 人近くの大学生の被験者から大体 6–15 タイムポイントを約 3 ヶ月に渡り額 (forehead)、手 (palm)、腸 (gut)、舌 (tongue) の四部位からサンプリングを行い、ヒト細菌叢の 16S rRNA 遺伝子配列を新型シーケンサにより解読したものである。本データから、各被験者の各体部位における個人特異的な細菌を属レベルで定義し、それらの属が生息する環境や関連性のある被験者のメタデータ情報を明らかにした。個人特異的な属を定義する手法として、ヒトで観測される属を以下の 3 つのタイプへと分類した。

1. General: 取得された全サンプルで、被験者を問わず多くのサンプルで観測される。
2. Specific: 取得された全サンプルで、少ないサンプルで観測されるが特定の被験者においては多くのサンプルで観測される。
3. Drifter: 取得された全サンプルで、被験者を問わず少ないサンプルで観測される。

以上の分類を行った後、個人に特異的な細菌すなわち Specific と、被験者のメタデータの関連を明らかにする解析を行った。結果、ヒト皮膚 (forehead, palm) において、被験者が

Specific を数多く持つ要因として「犬を飼うこと」や「年齢」等の被験者の要因が大きく関わってくることが明らかとなった。また、gut においては、被験者が個人特異的な細菌を多く持つことに対して関連のあるメタデータは挙がってこなかったが、tongue においては「hispanic であること」や「犬を飼うこと」が個人特異的な細菌を多く持つことに関連するものとして挙がってきた。

さらに、個人特異的な細菌が生息する環境について調査を行った結果、ヒト皮膚においては、ヒト皮膚だけではなく犬に関連した環境に生息する属が Specific として多く存在していた。ヒト常在細菌であると考えられる General と非ヒト常在細菌であると考えられる Drifter に分類された細菌が持つ機能遺伝子 KEGG ORTHOLOGY (KO) の比較解析を行った。結果、いくつかの機能遺伝子がヒト常在性に関連する因子として挙がってきた。具体的には、ヒト皮膚 (forehead, palm) における Actinobacteria 門や Firmicutes 門はファージに対する免疫機構である clustered regularly interspaced short palindromic repeat (CRISPR) 関連遺伝子を持つこと、gut における Bacteroidetes 門と Actinobacteria 門は抗生物質耐性遺伝子を持つこと等が明らかとなった。

本章では、1) 個人特異的な属の存在は被験者メタデータと強く関連すること、2) ヒト皮膚における Specific に分類された多くの細菌は犬に関連した環境に分布していること、3) ヒト皮膚や腸内にて観測される General に分類された細菌がヒトへの常在性に関連すると考えられる遺伝子をいくつか持っていること、の三点が明らかとなった。

3.2 材料と方法

3.2.1 16S rRNA 遺伝子配列データ取得と系統解析

16S rRNA 遺伝子配列データ (ERP005150–ERP005153) は European Bioinformatics Institute (EBI) の公共データベースより取得した。取得したデータから細菌の属レベルでの系統組成を以下の手順で得た。

1. Trimomatic version 0.33 を用いて、塩基配列のスコアが先頭もしくは後尾で 17 bp 未満が連続する領域を除去、平均クオリティ 25 以下もしくは長さが 90 bp 以下の配列を除去した。
2. N を含む配列を除去した。
3. usearch version 10.0.240 の ‘usearch -filter_phix’ コマンドを用いて PhiX 配列を除去した。

4. usearch version 10.0.240 の ‘usearch -fastx_uniques’ コマンドを用いて重複する配列をクラスタリングした。
5. usearch version 10.0.240 の ‘usearch -cluster_otus’ を用いて配列を 97% の一致度でクラスタリングした。
6. Rdp classifier version 2.12 を用いて、各クラスターの代表配列を属レベルで分類した。
7. 自作 python スクリプトを用いて各細菌のランクである domain, phylum, class, order, family, genus への分類の信頼度である confidence が 0.5 以上であった場合、その系統に分類した。confidence が 0.5 よりも下回った場合、Unclassified とした。

3.2.2 細菌の属レベルでのヒト常在性の定量化

本データにはネガティブコントロールのサンプルで読まれた 16S rRNA 遺伝子配列データも含まれており、細菌叢サンプルが入っていない状態での細菌叢データを取得することが可能である。すなわちキットや試薬に含まれている回避することができないコンタミネーションを細菌叢の系統組成として明らかにできる。全てのサンプルの系統組成を計算した後、ネガティブコントロールのサンプル全体で中央値、平均値のどちらも 0.5% 以上の値で観測された属についてはサンプリングに用いたキットや抽出キット等に含まれるコンタミネーションであるとし、*Novosphingobium* 属, *Aquabacterium* 属, *Idiomarina* 属, *Stenotrophomonas* 属, *Pseudomonas* 属, *Acinetobacter* 属を以降の解析から除外した。コンタミネーションと考えられる属の配列を除去した後、Bacteria もしくは Archaea へと分類することができた配列が 10,000 本以上あるサンプルが 6 タイムポイント以上得られている被験者のサンプルのみを用いた。各サンプルにおいて、各属の配列が少なくとも 10 本以上であれば存在すると定義した。以降の解析は各体部位 forehead, gut, palm, tongue それぞれ独立で行った。各サンプルにおける各属の存在の有無を定義した後、各被験者ごとに各属を General, Specific, Drifter の 3 タイプへと分類した。分類では以下の 2 つの統計値を用いた。

$$\text{inter-sample} = \frac{\text{total number of existence}}{\text{number of all sample}} \quad (1)$$

$$\text{inter-timepoint} = \frac{\text{total number of existence in the individual}}{\text{number of all sample in the individual}} \quad (2)$$

以上の2つの統計値の計算結果をもとに、被験者ごとに General ($\text{inter-sample} \geq 0.5$ and $\text{inter-timepoint} \geq 0.5$), Specific ($\text{inter-sample} < 0.5$ and $\text{inter-timepoint} \geq 0.5$), Drifter ($\text{inter-sample} < 0.5$ and $\text{inter-timepoint} < 0.5$) の3タイプへと、各属を定義した。

さらに、各被験者の inter-timepoint で最も高い値を採用する $\text{inter-timepoint-max}$ も計算し、被験者全体の、General ($\text{inter-sample} \geq 0.5$ and $\text{inter-timepoint-max} \geq 0.5$), Specific ($\text{inter-sample} < 0.5$ and $\text{inter-timepoint-max} \geq 0.5$), Drifter ($\text{inter-sample} < 0.5$ and $\text{inter-timepoint-max} < 0.5$) の3タイプへと、各属を定義した。

また、以降の解析では閾値 0.5 以外にも閾値をいくつか変化させ同じ解析を行っており、閾値 0.5 以外でも同じような結果が得られるか否かを議論した。

3.2.3 被験者メタデータの整形

公共データベースより取得したデータには被験者のメタデータ情報が結びついている。被験者のメタデータ情報には、身長や体重のような連続値の変数、運動の頻度や睡眠時間のような順序カテゴリカル変数、所属する学校や性別のようなカテゴリカル変数が含まれている。公開されている被験者メタデータの情報を統計解析にて扱うことが可能な形へと加工した。具体的には、順序カテゴリカルは連続値へと変換し、カテゴリカル変数はダミー変数へと変換を行った。公開されている被験者メタデータを加工した手順は以下の通りである。最終的に数値化とスケーリングがなされた 49 種類の被験者メタデータを取得した。

1. 被験者メタデータの記述された csv ファイル (https://static-content.springer.com/esm/art%3A10.1186%2Fs13059-014-0531-y/MediaObjects/13059_2014_531_MOESM3_ESM.csv) を取得した。
2. データから negative control を除き、被験者 ID (PersonalID) が紐づくデータのみにした。
3. 異なるデータが2つ以上紐付いていた PersonalID を除いた。
4. アンケートデータ (https://static-content.springer.com/esm/art%3A10.1186%2Fs13059-014-0531-y/MediaObjects/13059_2014_531_MOESM1_ESM.pdf) と被験者メタデータが記述された csv ファイルを元に、各メタデータを連続値変数 (continuous)、順序カテゴリカル変数 (ordinal)、カテゴリカル変数 (categorical)、yes/no 形式カテゴリカル変数 (two_category) へと分類した (表 3.S1)。

5. 各順序カテゴリカル変数の値を 1 から順に数値化した。例えば、never, rarely, occasionally, regularly, daily の 5 つが順序としてあった場合、順に 1, 2, 3, 4, 5 とした。
6. yes/no 形式カテゴリカル変数の場合、yes, no をそれぞれ 1, 0 とした。
7. カテゴリカル変数を R の fastDummies パッケージの dummy_cols 関数を用いてダミー変数化した。例えば、University の項目に A, B, C が存在する場合、University_A, University_B, University_C の 3 つへと分解し、それぞれ当てはまる選択肢に 1 を、当てはまらない選択肢に 0 を置いた。
8. 変換したデータ全体でカラムの合計値が 10 を下回る変数すなわち要素の数が 10 を下回ると考えられる変数は正常な統計解析に向かない変数として除いた。
9. 各変数の値を R の scale 関数を用いてスケーリングした。

3.2.4 各体部位における Specific の数と、被験者メタデータの相関解析

各被験者における Specific の数を計算した。観測された Specific の数を目的変数としたとき、各被験者メタデータ (年齢, 性別等) を説明変数とし Generalized linear model (GLM) を構築した。GLM の構築では、R の関数 glm を用いた。以下に GLM 構築時の R のコマンドを示す。ここで被験者メタデータに欠損値が確認された被験者はモデル構築時に除いた。

```
glm(data = D, y ~ x, family = poisson)
```

- D : 一列目被験者 (PersonalID), 二列目 y, 三列目 x からなるデータフレーム
y : 観測される Specific の数
x : ある被験者メタデータの値 (一種類のみ)

R の glm 関数による GLM の構築では、各説明変数にかかるパラメータの推定値 A と、その推定値の信頼区間に 0 が含まれる確率を計算する Wald 検定の結果が $\Pr(>|z|)$ として得られる。本来、Wald 検定で得られる $\Pr(>|z|)$ は推定値がゼロから十分に離れているかどうかを示すだけの指標であるが、この値を検定で得られる p 値として用いた。すなわち、 $\Pr(>|z|)$ が低いほど、その被験者メタデータが観測される Specific の数の増減に関連しているとした。具体的なプラスもしくはマイナスへの関連の強さは、推定値 A の大き

さから確認することができる。このように、推定値 A に対する Wald 検定の結果を元に Specific の数に関連のありそうな被験者メタデータを、多重検定補正の一つである BH 法を用いて被験者メタデータの数に対して行い列挙した ($q \leq 0.05$, BH 法)。以下に GLM に用いたモデル式を記述する。

$$y = \exp(Ax + B)$$

y : 観測される Specific の数

x : 被験者メタデータ

A : 推定値

B : 推定値

3.2.5 細菌の生息環境データ取得

各細菌の属が生息する環境をまとめた Web データベースとして、Latent Environment Allocation (LEA) が公開されている [21]。現在 LEA では、30,000 サンプル以上の公開済み 16S rRNA 遺伝子配列データ/メタゲノム配列データから VITCOMIC2 [58] を用いて属レベルでの系統組成を推定している。LEA では、各サンプルに存在する属レベルの系統組成と、サンプルに付随する英単語のワード組成を、トピックという環境の単位で推定している。LEA によるトピック数推定法では、80 トピックが属の生息環境を分類すべき数であるとされており、それら 80 トピックにおける属レベルの系統組成と環境を表す英単語のワード組成が公開されている。データとして LEA の API (<http://leamicrobe.jp/>) から 80 トピックにおける細菌 1,675 属の系統組成の二次元マトリックスを取得した。各属が生息する環境すなわちトピックは 80 トピック中で、その属の相対存在量が最も高い値を示すトピックと定義した。取得したデータから、各トピックに自身で理解しやすい形でラベリングした。例えば、topic 00 では、“topic 00 → marine | mat | hydrothermal | vent | water” という形で、ワード組成において割合が高かった上位 5 ワードを左から並べている。

3.2.6 General と Drifter が持つ機能遺伝子データの比較解析

一般的に観測される General は、一般的に観測されない Drifter と比較して、どのような機能的特徴を持っているのかについて推定を行った。各属が持つ機能の推定を行う方法と

して KEGG に登録されている細菌の遺伝子を機能へと分類した機能遺伝子 (KO) のデータを用いた [32]。KEGG API (2019-01-12) より各細菌の株が持つ KO をリストとして取得した。今回、細菌の属レベルでの解析のため、ある属に所属する株が特定の KO を持っていた場合、その属が特定の KO を持っていることとする Pan-genomics の概念を用いた。ヒト常在細菌が持つ機能遺伝子を推定するため、General と Drifter の間で有無に関連のある KO を Fisher の正確確率検定を用いて推定した。これらの KO の推定では、ある程度類似した細菌同士の比較を行うため Actinobacteria 門, Bacteroidetes 門, Firmicutes 門, Proteobacteria 門へと細菌の属を分類し、それぞれの門について行った。さらに、比較解析によって得られた KO 群から、それらの KO が所属するモジュールを明らかにした。各モジュールに所属する KO の偏りを Fisher の正確確率検定から統計的に算出し、General と Drifter にて差のあるモジュールを明らかにした。General と Drifter の間の機能遺伝子の比較解析手法は、以下の通りである。

1. General と Drifter の間で、取得されている全ての KO について Fisher の正確確率検定を行った。Fisher の正確確率検定は、二群で 3 つ以上の要素がある場合に行った。また、比較に用いる属の全てがその KO を持っている場合と持っていない場合は Fisher の正確確率検定を行わなかった。検定を行うときに用いる分割表の例を以下に示す。例では、比較を行う General と Drifter がそれぞれ KO00001 を持つか持たないかを集計した結果を示す。

例) K00001 の有無	General	Drifter
あり	10	0
なし	0	10

2. Fisher の正確確率検定により有意差 ($p \leq 0.05$) が確認された KO 群から、さらに Fisher の正確確率検定を用いて KO の各モジュールへの偏りを明らかにした ($q \leq 0.05$, BH 法)。General と Drifter を定義する閾値 0.5, 0.4, 0.3, 0.2 について行った。結果で得られた KO 群から、各モジュールに所属する KO の数を数える手法を用いた場合、そのモジュールに属する KO の数の違いによってバイアスが生じる。このバイアスを回避するため、どのモジュールに KO が偏って得られているのかについて以下の表で示す分割表にし、Fisher の正確確率検定を行った。例では、全体でモジュールに紐づく KO が 18,000 個、M00001 に所属する KO が 13 個のとき、結果で得られた KO 群が 100 個、その中で 10 個が M00001 に所属していた場合を示す。

例) M00001	所属しない KO 群	所属する KO 群
全体の KO 群	17,987	13
結果の KO 群	90	10

3.3 結果

3.3.1 各体部位にて観測される General, Specific, Drifter の属の数

各属のヒトへの常在性を定量化するため、各体部位から採取された細菌叢の 16S rRNA 遺伝子配列解析サンプルにおける各属の存在の有無を明らかにした。サンプル中に存在するか否かの定義を行う配列数の閾値は、細菌の存在の有無を議論した研究が報告されており、同じく 10 本という閾値を用いて議論を進めていった [59]。ヒトへの常在性の定量化では各被験者について、1) 全てのサンプルからどの程度の割合で観測されたのかを示す inter-sample、2) 被験者内でどの程度の割合で観測されたのかを示す inter-timepoint、の 2 つの統計値を用いた。

今回は、inter-timepoint が被験者間で最も高い値を示した inter-timepoint-max を用いて General, Specific, Drifter へと各属を分類し集計した。Actinobacteria 門, Bacteroidetes 門, Firmicutes 門, Proteobacteria 門の 4 門の観測される属の数を General, Specific, Drifter の間で比較した結果、ヒト皮膚 (forehead, palm) のほうが、gut と tongue に比べて多いことが明らかとなった (図 3.1)。しかしながら、gut の Firmicutes 門における General は、他の三部位よりも圧倒的に多くの属が観測された。実際の属について深く見てみると、ヒト腸内には Firmicutes 門の中でも Clostridia 綱の属が General として多く存在していた。

3.3.2 Specific と被験者メタデータの関連性と環境への分布

Specific と被験者メタデータ間の関連性を明らかにするため、GLM を用いた解析を行った。具体的には、それぞれの被験者において観測された Specific の数を目的変数としたとき、どのような被験者メタデータが説明変数として適しているのかについて GLM を用いて探索を行った。結果、ヒト皮膚 (forehead, palm) については、被験者の年齢の増加 (Age)、水疱瘡の歴 (Chikinpox)、犬の飼育 (Dog) が Specific の数の増加を説明する変数として推定された (図 3.2)。一方、サプリメントの服用 (nutritionalSupple)、皮膚が健常であること (SkinCondition_none)、水泳や入浴の頻度 (Swimming) が個人特有の細菌の減少を説明する変数として挙がってきた。tongue については、犬の飼育 (Dog) と人種が hispanic (Ethnicity_hispanic) であることが Specific の数の増加を説明する変数として推定

された。一方、BMI の高さ (BMIclass) と体重 (weight_kg) が Specific の数の減少に関わるを説明変数として推定された。gut については、個人特異的な属の増減に関わる説明変数として被験者メタデータは推定されなかった。

今回の解析では、Specific を inter-sample と inter-timepoint の値に対して 0.5 という閾値を用いて定義している (inter-sample < 0.5 and inter-timepoint $\geq$ 0.5)。さらに、0.1—0.9 の間でも閾値を設定し、同じように GLM を構築し、推定値 A に対する Wald 検定 ($q \leq 0.05$, BH 法) を行った。結果、閾値が異なっても、ある程度同じ傾向がいくつかの閾値で確認された (図 3.S1)。

さらに、各被験者において Specific として定義された属の環境への分布を集計し、どのような環境由来の属が各被験者の Specific な細菌として観測されたのかについて明らかにした (図 3.6)。結果、ヒト皮膚 (forehead, palm) では、topic 27 $\rightarrow$ skin | forearm | volar | surface | hand が最も多く、同じヒト皮膚に関連する環境の細菌が多かった。次に多かった環境は topic 45 $\rightarrow$ america | arable | dog | organismal | door であり、犬に関連する環境であった。forehead と palm のどちらにおいても、階層的クラスタリングを行った際に大きく別れたクラスター的一方では被験者 C237, A042, C214, B110 が所属しており、もう一方のクラスターと比較して topic 45 $\rightarrow$ america | arable | dog | organismal | door と Others に生息する属が多かった。gut では特に嫌気性環境すなわち動物の糞便や腸内に関連した環境すなわち gut のワードを多く含むトピックが主であった。tongue では最も多い環境として、topic 12 $\rightarrow$ tonsil | diseased | oral | canal | root が挙がってきていた。次に多かった環境は topic 48 $\rightarrow$ america | skin | dog | organismal | pet であった。

3.3.3 ヒト常在細菌と非ヒト常在細菌の機能遺伝子の比較解析

ヒトで観測された属について被験者全体で考慮したときの、一般的にヒトで観測される General と、一時的にしか観測されない Drifter について着目した。今回の解析では、General と Drifter をそれぞれヒト常在細菌、非ヒト常在細菌として定義した。各属が持つ KO の推定には、KEGG 上に登録されている株の機能遺伝子を属へと分類し、Pan-genomic に考慮した。推定の結果、各被験者において General と Drifter の間で有無に関連のある KO がいくつか確認された (Fisher の正確確率検定, $p \leq 0.05$)。付録の表では、Fisher の正確確率検定で特に低い p 値を示した結果のみを記載するため $p \leq 0.005$ 以下で帰無仮説が棄却された KO のみを示す (表 3.S3)。今回はヒト常在性因子として General が Drifter と比較して保有する確率が高い KO に着目し、細菌のヒト常在性に関与するとされる因子として挙げた。本結果は検定回数すなわち比較を行った KO の数だけ各体部位と各門の中

では検定が行われており、比較に用いた KO の数について BH 法を用いて多重検定補正を行った結果も q value.BH として表に記載した。得られた KO を Fisher の正確確率検定による p 値が低い上位から確認した。

結果、ヒト皮膚 (forehead, palm) では、Actinobacteria 門では、CRISPR 関連遺伝子である K19131: csb1; CRISPR-associated protein Csb1 と K19132: csb2; CRISPR - associated protein Csb2 や、K14645: K14645; serine protease [EC:3.4.21. -] や K11065: tpx; thiol peroxidase, atypical 2 - Cys peroxiredoxin [EC:1.11.1.15] が、Firmicutes 門では CRISPR 関連遺伝子である K09952: csn1, cas9; CRISPR - associated endonuclease Csn1 [EC:3.1. -. -]、転移因子である K07482: K07482; transposase, IS30 family, K05873: cyaB; adenylate cyclase, class 2 [EC:4.6.1.1] が挙げられてきた。palm では、Proteobacteria 門では、K06908: K06908; uncharacterized protein と K06906: K06906; uncharacterized protein が挙げられてきており uncharacterized protein となっていたが UniProt における同遺伝子のアノテーションにおいてはファージの体を構成するタンパク質に関連した遺伝子であった [60, 61]。さらに palm では、Bacteroidetes 門では K01834: PGAM, gpmA; 2,3 - bisphosphoglycerate - dependent phosphoglycerate mutase のような抗酸化機能が挙げられてきた。gut では多くの KO がヒト常在性の因子として挙げられてきた。中でも特に有意であった KO は、Firmicutes 門では K03737: por, nifJ; pyruvate-ferredoxin/flavodoxin oxidoreductase [EC:1.2.7.1 1.2.7.-]、バンコマイシン耐性遺伝子である K18349: vanRC, vanRE, vanRG; two-component system, OmpR family, response regulator VanR、Actinobacteria 門ではマクロライド耐性遺伝子である K08217: mef; MFS transporter, DHA3 family, macrolide efflux protein、テトラサイクリン耐性遺伝子である K18220: tetM, tetO; ribosomal protection tetracycline resistance protein、Bacteroidetes 門では多くのフェレドキシン機能遺伝子等が同じ p 値でトップに並んでいた。tongue では、多重検定補正後に有意水準以下 ($q \leq 0.05$, BH 法) を示す KO が得られなかったが、 p 値が低い上位の KO に着目すると K01226: treC; trehalose-6-phosphate hydrolase [EC:3.2.1.93] や K02793, K02794, K02795, K02796 のようなマンノース感受性の PTS system を持つことが明らかとなった。

さらに KO が担う役割を理解しやすいように、各 KO が所属するモジュールの情報からヒト常在に関わる機能を明らかにした。結果、各体部位におけるヒト常在性に寄与すると考えられるモジュールが推定できた (図 3.7)。推定では、General と Drifter を定義する閾値 0.5, 0.4, 0.3, 0.2 で行い、3 つ以上の閾値で推定されたモジュールに着目した ($q \leq 0.05$, BH 法)。

ヒト皮膚 (forehead, palm) においては、Actinobacteria 門の General は Drifter に比べて M00378: F420 biosynthesis と M00237: Branched - chain amino acid transport system

に関連した KO を持たないこと、が明らかとなった。

gut においては、Firmicutes 門の General は Drifter に比べて M00159: V/A - type ATPase, prokaryotes に関連した KO を持つこと、Bacteroidetes 門の General は Drifter に比べて M00159: V/A - type ATPase, prokaryotes を持つこと、が明らかとなった。

tongue においては、Actinobacteria 門の General は Drifter に比べて M00276: PTS system, mannose - specific II component, M00269: PTS system, sucrose - specific II component に関連した KO を持つが M00222: Phosphate transport system を持たないこと、Firmicutes 門の General は Drifter に比べて M00222: Phosphate transport system を持たないこと、Proteobacteria 門の General は Drifter に比べて M00317: Manganese/iron transport system を持つが M00323: Urea transport system と M00122: Cobalamin biosynthesis, cobinamide => cobalamin を持たないこと、が明らかとなった。

3.4 考察

ヒト腸内をはじめ、ヒト細菌叢と宿主のメタデータの間で観測される相関関係や、ヒトの振る舞いがヒト細菌叢に及ぼす影響を明らかにした研究が多くなされてきた [7, 62, 63]。これらの研究では、ヒトが服用する薬やヒトの状態が細菌叢の系統組成に影響を与えること、細菌叢の系統組成の多様性などに影響を与えることが主に報告されており、ヒト細菌叢とヒトの疾患の関連性を明らかにする研究ではサイドエフェクトとして注意深く考えられるべき事項として注目されている。本章の研究より、個人特定に寄与する細菌の存在がヒトの生活習慣やメタデータによって説明できることが示唆された。細菌叢に含まれる個人特異的な細菌そのものを明らかにすることで、その被験者がどのような環境と日頃から接触しているのかについての行動ログとして扱うことも期待される。実際に、本章で行った研究では、被験者からいくつか個人特異的であると考えられる属を定義し、被験者が持つ個人特異的な属の数と被験者メタデータの関連性を調査した。ヒト皮膚である forehead と palm に共通して得られた傾向として、被験者が持つ個人特異的な属の数は、犬を飼育していることや被験者の年齢、その被験者が水疱瘡の歴があるかに強く関連しうることが明らかとなった。先行研究より、犬を飼育している家族は飼育していない家族に比べて多くの細菌を皮膚上でシェアすることが報告されている [24]。この犬の飼育という行動によって、個人特異的であると考えられる細菌を犬を媒介して多く取得していると考えられる。実際に、ヒト皮膚における個人特異的な細菌が生息する環境を調査したところ、犬に関連したトピックに存在する細菌が上位に挙がっており、犬こそ個人特異的な細菌をヒト皮膚に流入させる要因の一つとして考えられる。特に、被験者 A042, C214, B110,

C237 は犬に関連した環境に生息する個人特異的な細菌を多く持っていた。さらに、犬に関連した環境の細菌を個人特異的な属として多く持っている被験者は、他にも様々な環境由来である属を持っていることが明らかとなっており、動物の飼育はヒト皮膚に様々な細菌を取りこむ現象を起こすと考えられる。水疱瘡の歴も大きな要因として確認されていたが、被験者メタデータ間の相関関係を調査した結果、犬の飼育が水疱瘡の歴、と正の相関を示していた (図 3.S2)。さらに、犬を飼育するという被験者メタデータに比べて推定値が小さいことから、水疱瘡が強く関連していた要因は偽相関であると考えられる。一方で、水泳を日課とすることや、サプリメントの服用はヒト皮膚の個人特異的な属を少なくする傾向にあることが確認された。水泳に関しては入浴の意味も含まれており、入浴することはヒト皮膚に存在する細菌を洗い流す役目があると考えられる。実際に、手を洗うという行為自体が、手の表面の細菌叢の系統組成に変化を与えるという報告もある [64]。入浴というヒト皮膚への擾乱はヒト皮膚における個人特異的な細菌の排除機能の効果もあると考えられる。サプリメントの服用については先行研究からも考えられる理由を発見することができなかった。サプリメントの服用という行為が被験者の体調管理のための振る舞いだと考えれば、他にも被験者が体調管理のために他の振る舞いを多く行っていることが予想される。例えば、被験者が皮膚に何かしらの保湿剤やスキンケア商品を扱っている可能性もあり今回のような結果が得られたと想像することもできる。ヒト腸については個人特異的な細菌の数と大きく関連したメタデータは確認されなかった。理由として、ヒト腸という環境自体が、ヒトの体部位の中でも外環境から比較的奥まった体部位であることが考えられる。ヒト腸に外環境由来の細菌が到達するためには、被験者の口、食道、胃、小腸、などを通り最終的に大腸へと到達する必要がある。これを踏まえると、ヒト腸に到達できる細菌は非常に限られるため、個人特異的な細菌の存在を確認できない可能性もある。しかしながら、属をより細かい系統である種や株で分類すると、各被験者に存在する属や種は SNP レベルや遺伝子レベルで異なることが報告されている [12,65]。より細かいレベルで系統を分類することで、外環境由来の細菌と個人特異的な細菌の存在を確認できることが期待される。ヒト舌については人種が hispanic であることや、犬を飼っていることが関連していることが明らかとなった。人種によって食文化などが異なる影響等が考えられたが、推定値を見ると、ヒト皮膚における犬の飼育などの推定値に比べて小さいことから、その影響は特に大きなものではないと考えられる。一方、体重や BMI についても、ヒト舌における個人特異的な細菌を減らす傾向に寄与することが推定された。BMI の上昇は唾液の分泌量と相関しており、この唾液の分泌が舌の個人特異的な細菌を減らす役割があると考えられる [66]。

以上のように、個人特異的に存在する細菌は何かしらの被験者の振る舞いや状態によっ

て生じることが示唆できた。しかしながら、少なくとも特定の被験者にて高頻度で観測される細菌であることから、これらが外環境由来であると断定することはできず、特定の被験者に存在し続けるヒト常在細菌である可能性もある。ヒト常在細菌と非ヒト常在細菌の共通点や異なる点の議論を行う場合、個人特異的な細菌はどのようなヒトに対しても常在正を持つと定義することができないため、個人特異的な細菌を除いた上で議論する必要がある。そこで、個人特異的な細菌すなわち Specific を除き、ヒト常在細菌と定義した General と非ヒト常在細菌と定義した Drifter の機能遺伝子の比較解析を行った。本比較解析によって、細菌のヒト常在性に関連した機能遺伝子を推定し、それらの機能遺伝子の役割からヒトへの常在に重要な細菌の機能を議論した。

ヒト常在性に関連する機能遺伝子ということで、各体部位における各門の比較で General が Drifter に比べて高い確率で保有しているとされた KO を Fisher の正確確率検定によって明らかにした。その保有率が高い KO として、ヒト皮膚における Actinobacteria 門と Firmicutes 門では CRISPR 関連遺伝子である K19131: csb1; CRISPR – associated protein Csb1、K19132: csb2; CRISPR – associated protein Csb2、K09952: csn1, cas9; CRISPR – associated endonuclease Csn1 [EC:3.1. –. –] を挙げることもできた。CRISPR の働きとして、その領域中にファージやプラスミドなどの外来の塩基配列をコードしており、それらがファージの感染を防ぐ機構として働くことが知られている [67]。実際に、ヒト皮膚における *Propionibacterium* 属や *Corynebacterium* 属や *Staphylococcus* 属は CRISPR を持っており、ヒト皮膚上のファージの塩基配列をコードした領域の存在まで確認されている [50]。このファージとの関連性がヒト皮膚上に存在するための何らかの因子として働いている可能性もある。さらに、ヒト皮膚に存在する Actinobacteria 門は K11065: tpx; thiol peroxidase, atypical 2 – Cys peroxiredoxin [EC:1.11.1.15] のようなペルオキシダーゼに関連する遺伝子を持っていた。thiol peroxidase は酸化ストレス耐性を担うことが報告されている [68]。本機能が宿主のマクロファージによる酸化を抑え細菌の生存に寄与する可能性も考えられ、細菌がヒト皮膚に常在するためにはヒトの免疫機構であるマクロファージに対抗する術が必要であることが予想される。gut では抗生物質耐性に関連した KO がいくつか推定できた。Firmicutes ではバンコマイシン耐性遺伝子である K18349: vanRC, vanRE, vanRG; two-component system, Odmpr family, response regulator VanR、Actinobacteria 門では、マクロライド耐性遺伝子である K08217: mef; MFS transporter, DHA3 family, macrolide efflux protein、テトラサイクリン耐性遺伝子である K18220: tetM, tetO; ribosomal protection tetracycline resistance protein が挙がってきた。これらは抗生物質から細菌が身を守るための機能である [69, 70]。抗生物質が使われている現代では、ヒトが服用する抗生物質に対して耐性を持つために、ヒト腸内に常在するうえで必須な

因子となっていることが考えられる。tongue では、Actinobacteria 門において糖の代謝に関連する遺伝子がいくつか推定された。K01226: treC; trehalose-6-phosphate hydrolase [EC:3.2.1.93] はトレハロースをグルコースへと分解する機能を持っており、ヒトの口腔内における糖の代謝を担っていると考えられる。tongue の Actinobacteria 門は他にも K02793, K02794, K02795, K02796 といったマンノース感受性の PTS system を持っており、ヒトの舌に生息する Actinobacteria 門が糖代謝に特化している可能性も示唆でき、この糖代謝も舌で常在する因子としても考えられる。この Actinobacteria 門で確認された KO 群であるが、p 値では 0.005 を下回っているものの多重検定補正をかけると q 値が 0.05 を上回ってしまい有意差が確認されないことから、この糖代謝に関連した機能は検定による偽陽性の確率が高い。しかしながら、tongue における General な Actinobacteria 門は、*Actinomyces* 属、*Rothia* 属、*Atopobium* 属の 3 属であり、いずれも inter-sample の値が 0.9 以上と非常に高く、これらの属が持つ機能を重要な因子として見逃すことはできない。

さらに、本解析では各体部位の各門で取得された KO 群から、それらの KO が属するモジュールについても明らかにした。ヒト皮膚 (forehead, palm) に存在する Actinobacteria は F420 biosynthesis を持たない傾向にあった。コエンザイム F420 は、メタン生成におけるその役割でよく知られているデトックス活性コファクターである。コエンザイム F420 は、methanoarchaeal 代謝において重要な役割を果たすが、テトラサイクリンおよびリンコマイシンのような抗生物質の合成を助ける *Streptomyces* 属、*Rhodococcus* 属、*Nocardioidea* 属、*Mycobacterium* 属などの Actinobacteria 門の様々な真正細菌においても見出されていることから、ヒト皮膚に生息する Actinobacteria 門は抗生物質を作らない傾向にあると考えることができる [71]。ヒト皮膚がディフェンシン等の抗細菌性のペプチドを分泌するという報告もあり、ヒト皮膚に生息する Actinobacteria 門は、抗生物質を用いたヒト皮膚上での細菌同士の繁栄争いをする必要がなくなったため、この機能を持たなくなった可能性もある [72]。

ヒト腸における Bacteroidetes 門と Firmicutes 門は、V/A - type ATPase を構成する機能遺伝子を多く持っていた。先行研究より、*Enterococcus* 属が持つ V-type ATPase はアルカリ環境下でナトリウム恒常性を維持する働きがあることで知られている [73]。真核生物では pH の恒常性など多くの機能を担うことで知られている [74]。このようにヒト腸内に存在するためには、外環境の化学的な影響に耐えうる機能が必須であることが考えられる。

ヒト舌については、多くの化合物輸送系のモジュールの有無が結果として確認されており、これらの化合物の取り込みや排出が重要な因子となっている可能性も考えられた。

本研究で用いた手法は属レベルでの比較ゲノム解析かつ 16S rRNA 遺伝子配列系統解

析で明らかとなった系統の機能遺伝子を公開されている同じ系統の株の情報から推定したものである。より詳細な比較ゲノム手法として、環境中の細菌のゲノムを正確に読むことができるロングリードシーケンスを用いたメタゲノム解析手法や、メタゲノム配列から特定の種の遺伝子を推定する手法などを用いて、非常に近縁な細菌同士のゲノムを比較する方法を採用することも重要である。一方、本研究で行った比較方法でも、仮説を立てるに値する機能遺伝子がいくつか挙がっており、ここで挙がってきた遺伝子の有無がヒトへの常在性に関連するか否かを検証実験により証明することや、これらの機能の発生源を明らかにすることで、ヒト常在細菌の由来を議論することも十分に可能であると考えられる。

3.5 図表

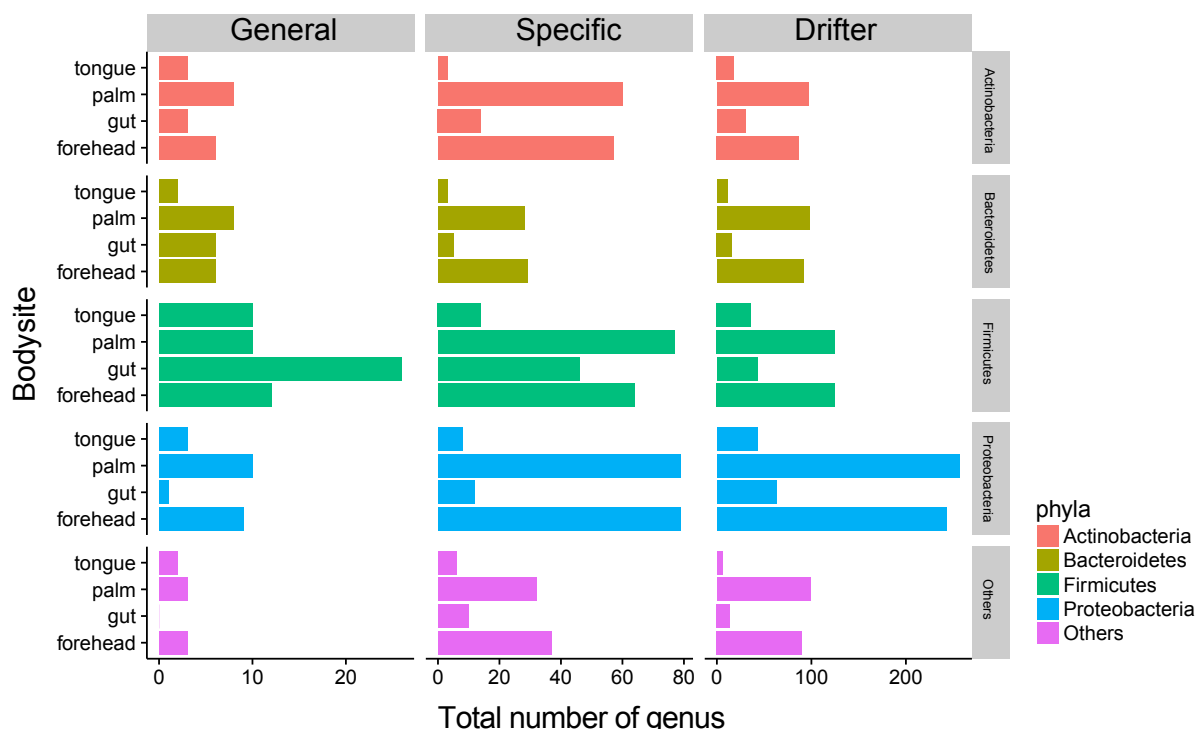


図 3.1. General, Specific, Drifter の属の数

各体部位ごとに各門の General, Specific, Drifter の数を集計した。列は General, Specific, Drifter を示しており、行は各細菌の門を示す。

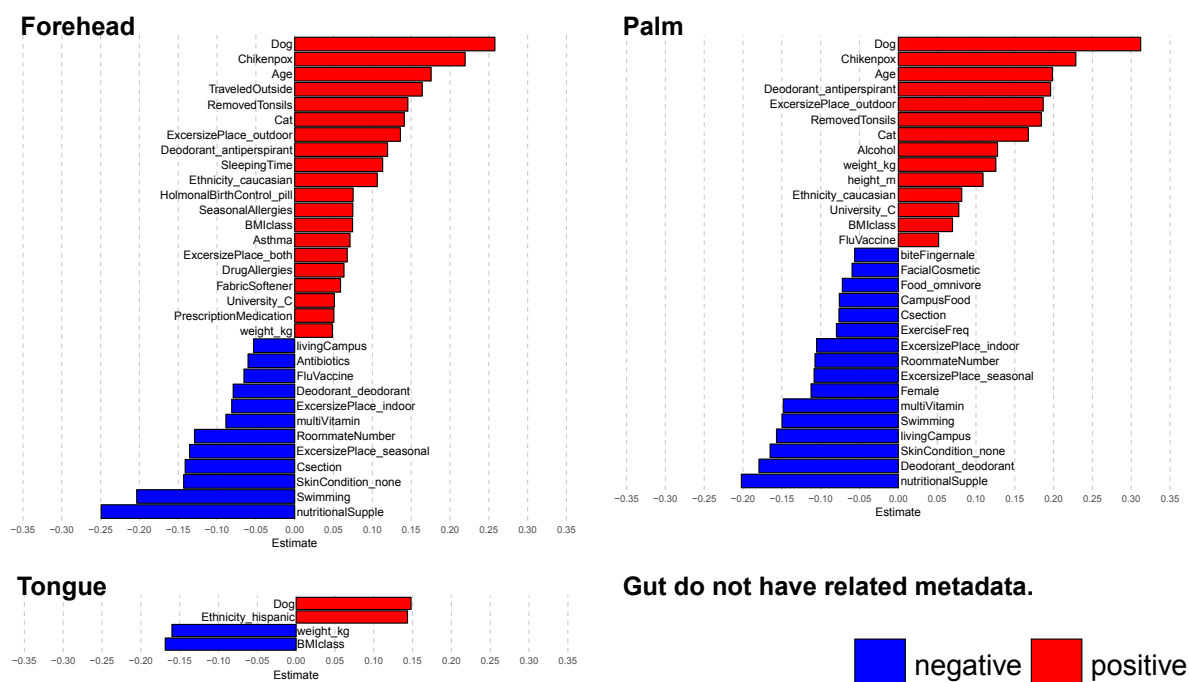


図 3.2. 各体部位における Specific の数と関連のあるメタデータ

各対部位ごとに、各被験者の Specific の数と、それに関連があるとされた被験者メタデータを明らかにした。Specific の数に対して正 (赤)、負 (青) に関連するメタデータの、その推定値 (Estimate) の大きさを x 軸にあらわしている。

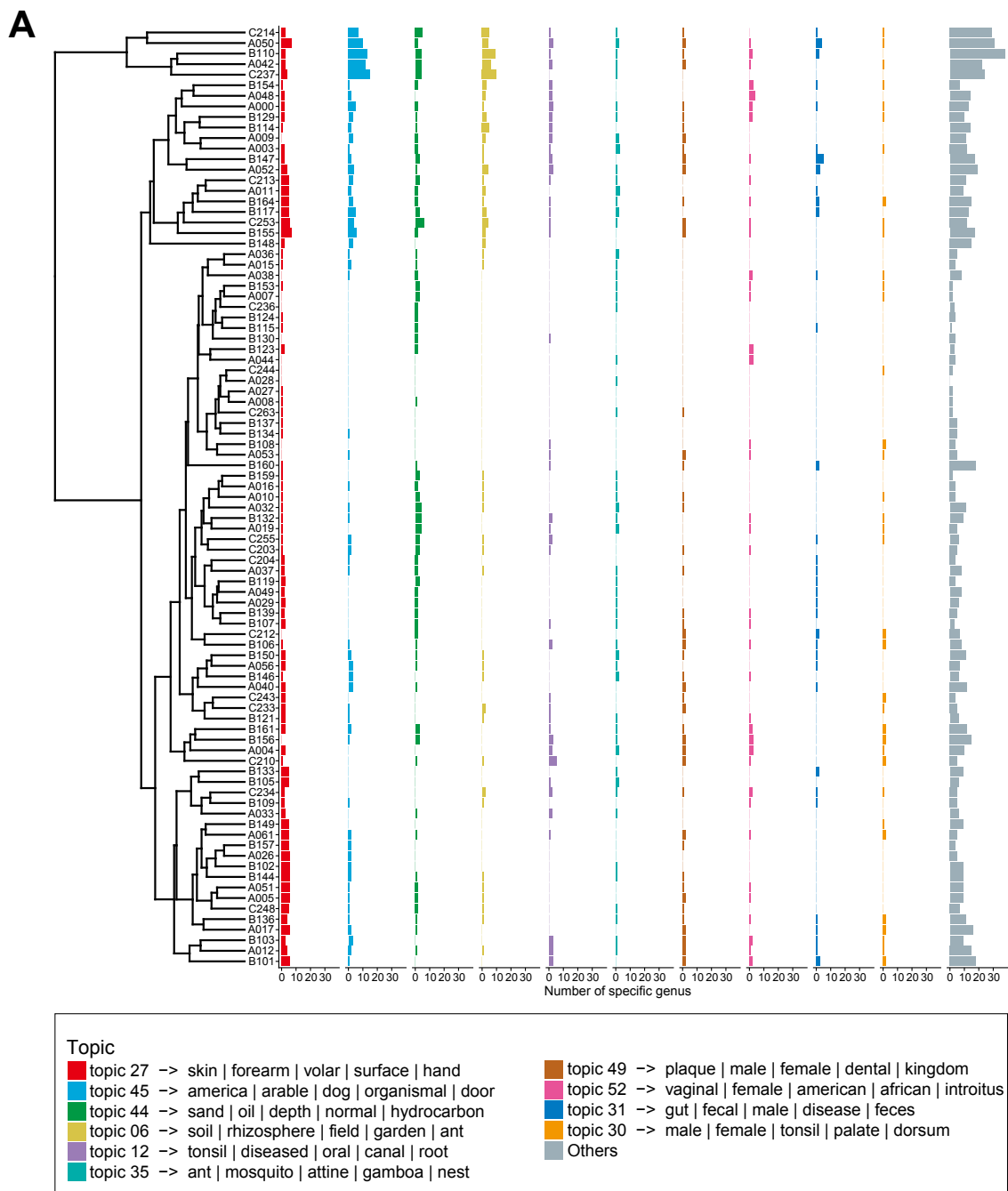


図 3.3. 各体部位における各被験者の Specific の生息環境の集計結果

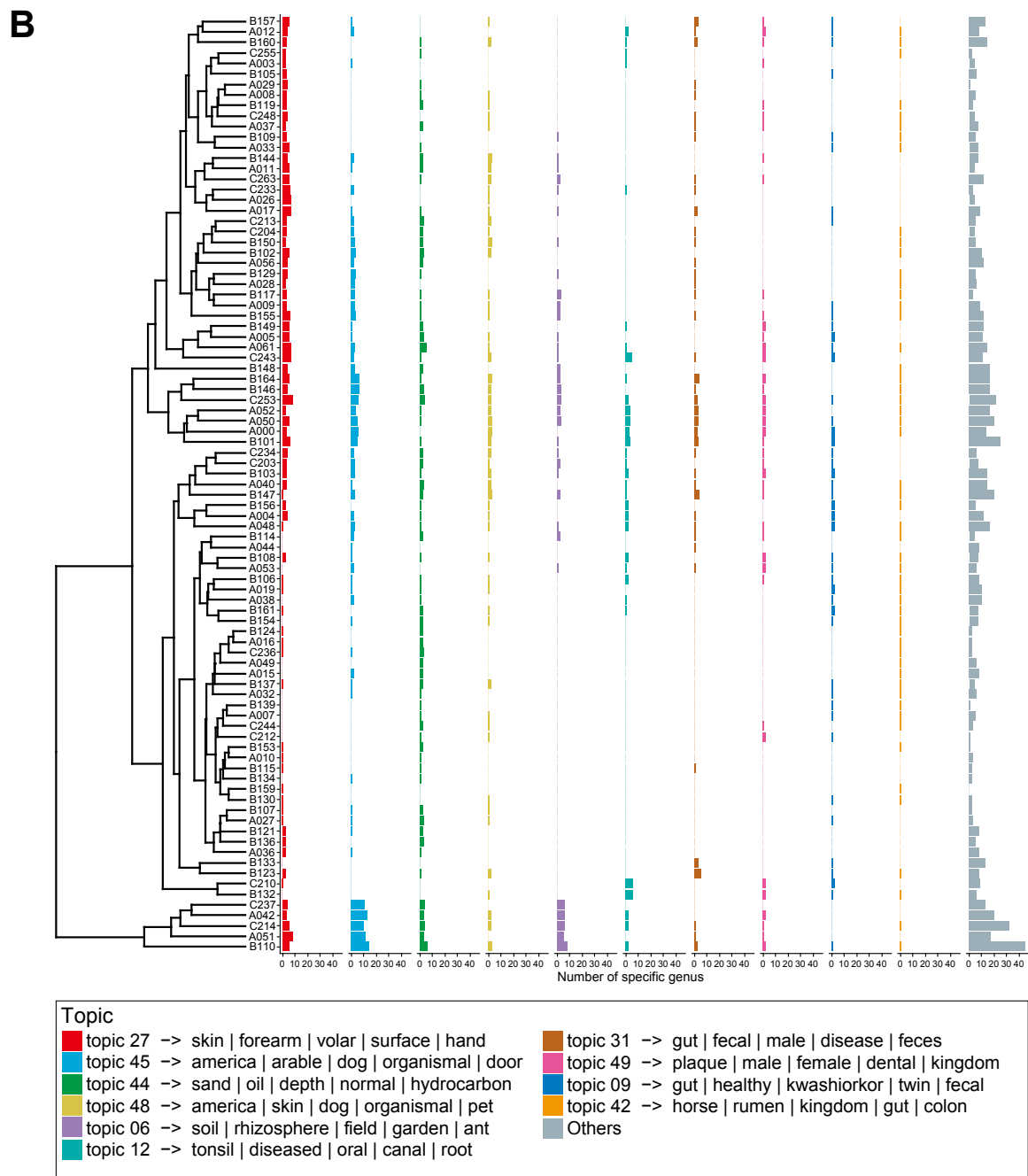


図 3.4. 各体部位における各被験者の Specific の生息環境の集計結果

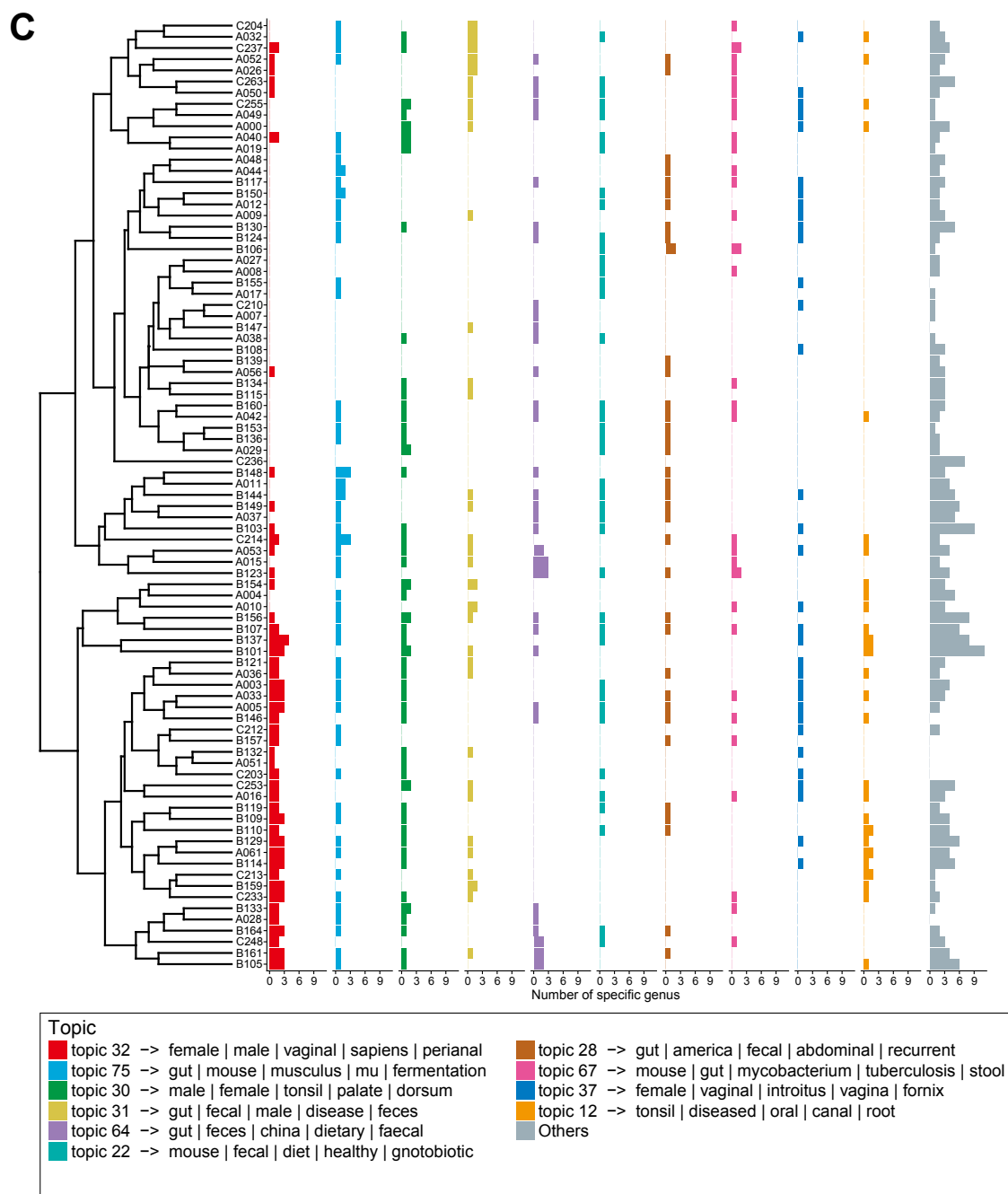


図 3.5. 各体部位における各被験者の Specific の生息環境の集計結果

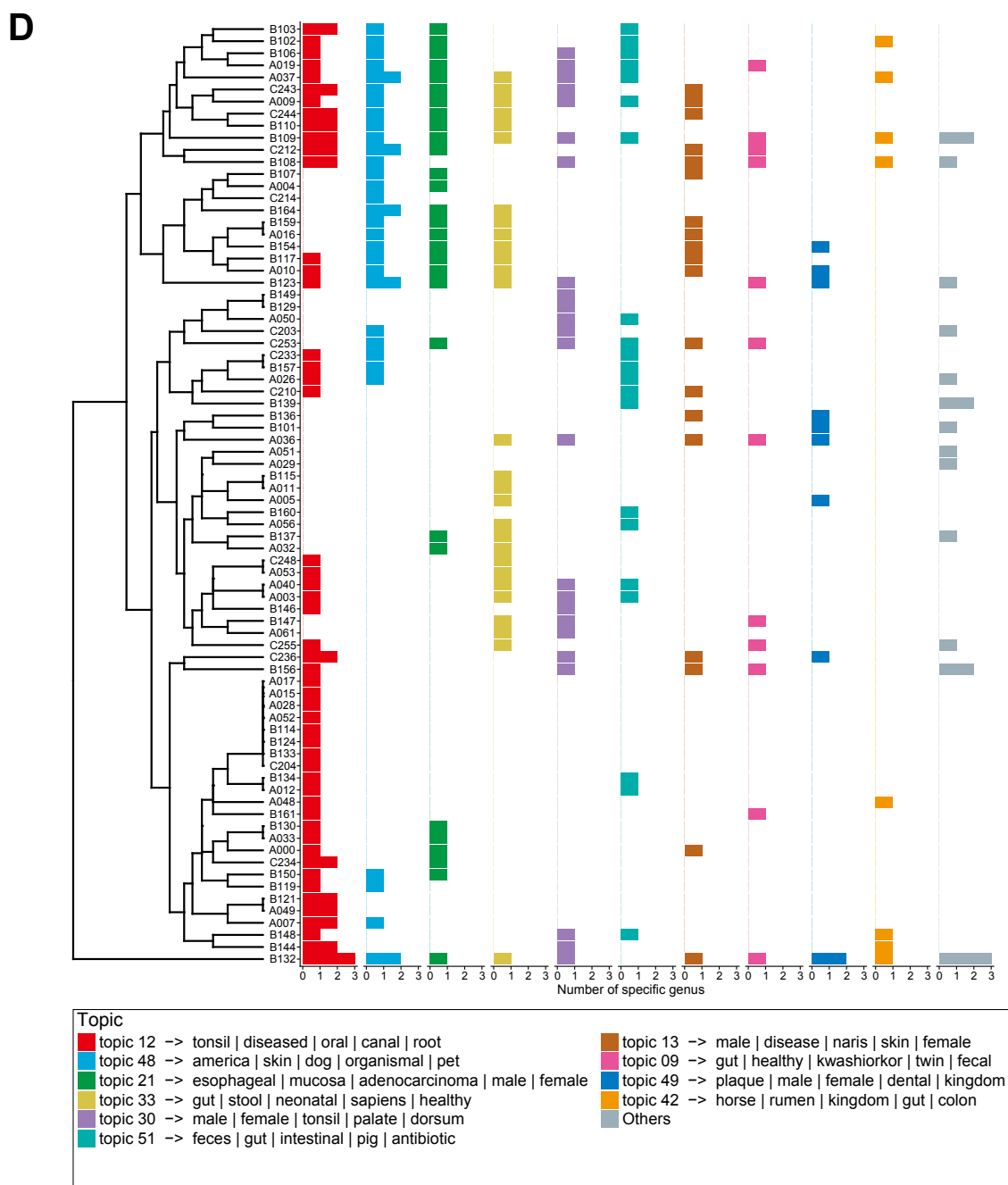


図 3.6. 各体部位における各被験者の Specific の生息環境の集計結果

各被験者で観測された Specific の生息地を LEA の結果から推定し、どのような環境に生息する属が多いのかを集計した。列は、左から集計した合計が高かったトピックのトップ 10 を示す。x 軸は各トピックに生息している属の数を示し、y 軸は PersonalID を示しており、Euclidean 距離を用いた階層的クラスタリングによって並べている。それぞれ、A) forehead, B) gut, C) palm, D) tongue の体部位を示す。

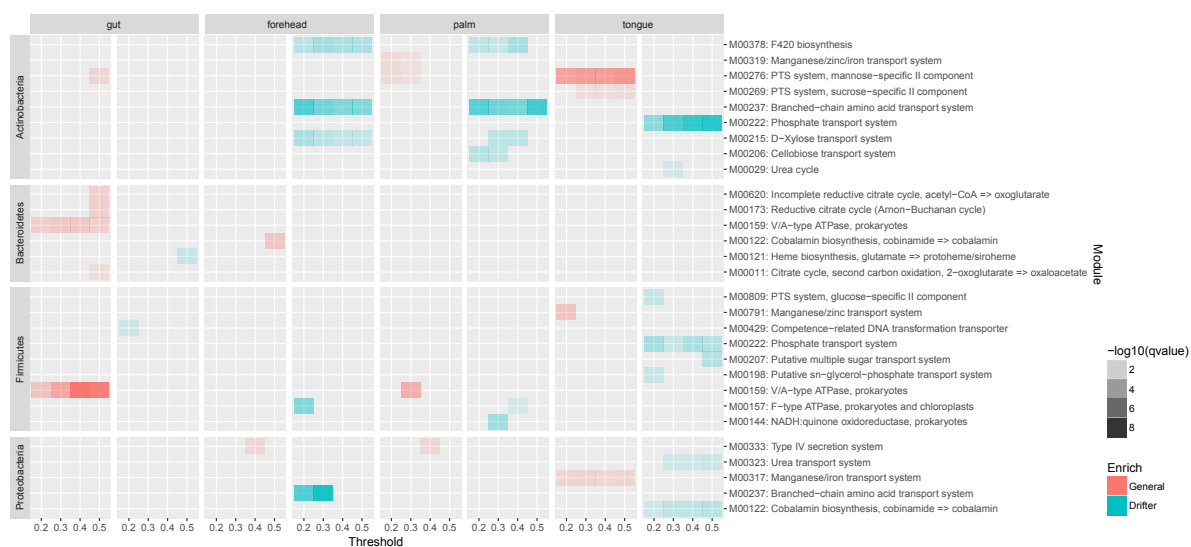


図 3.7. 各体部位の各門における General と Drifter で多く保持されたモジュール

各体部位の各門について General と Drifter で多く保持されたモジュールを示す。列は各体部位を示し、行は各門を示す。ヒートマップの色は General と Drifter のどちらで多く保有されていたかを示し、濃さは多重検定補正を行ったときの q 値の $-\log_{10}$ スケールを示す。x 軸は General と Drifter を決定するときの閾値 0.2, 0.3, 0.4, 0.5、y 軸はモジュールを示す。

第4章 総括と展望

これまで環境中の細菌叢の研究が多くなされてきており、それらの細菌叢の中に含まれている細菌の系統や、遺伝子といった情報が解析・公開され続けている。このように様々な環境の細菌の系統や遺伝子情報が明らかになることは、細菌そのものが何故、その環境で、その系統で、そのような遺伝子を持っているのかを明らかにすることへと繋がる。

その一つとして、二章の研究では同じ環境であると考えられたヒトの額という環境の細菌叢を採取し、解析することで、各被験者間で異なる細菌叢を持っており、その情報が個人特定に用いることが可能なレベルで被験者間で異なっており、維持され続けていることが明らかとなった。この要因として、ヒト皮膚に存在する *Propionibacterium acnes* という優占種に埋もれて確認するのが困難であった少数種の存在があった。これらの少数種は決まった被験者で多く観測されることから、各被験者のアイデンティティの一つである。何故、決まった被験者でのみ長期的に観測されるのかについて考えると、被験者が生活の中で常に外環境から取得しているもしくは、過去に取得し共生を果たしている可能性が挙げられた。

それらを明らかにするため、第三章の研究ではどのような被験者メタデータが個人特異的な細菌の数を増加させるのかについての解析を行った。この結果から、ヒト皮膚では、犬を飼育すること、年齢の増加など、いくつかの被験者メタデータが個人特異的な細菌の数の増加に関連していることを示す結果が得られた。一方、入浴や栄養剤の摂取などは、個人特異的な細菌の数の減少に関与している可能性を示した。ヒト皮膚だけでなく、ヒト腸とヒト舌でも同じ解析を行ったが、ヒト皮膚のように多くの関連する被験者メタデータを得ることができなかった。確かに、ヒト皮膚についてはいくつかの被験者メタデータから個人特異的な細菌の保持を説明することができると期待されるが、個人特異的な細菌の存在そのものが全て被験者の振る舞いから由来するものであると結論付けることは困難である。これらの被験者特異的に存在する細菌の中には、被験者が幼少期に獲得し常在しているものや、先祖代々受け継がれているものも存在している。現在、どのような細菌がヒトに定着し常在し続けているのかについては明らかになっていない点も多い。三章の研究では、引き続き個人特異的な細菌を定義した際に同じく定義したヒト常在細菌と非ヒト常在細菌が持つ機能遺伝子を比較し、ヒト常在細菌がヒトに常在するために必要な細菌自

身が持つべき機能の推定を行った。結果、ヒトの額、腸、手、舌について、いくつかの機能遺伝子の有無がヒト常在細菌と非ヒト常在細菌の間で差があることが明らかとなった。これらの機能遺伝子を調べると、何故ヒト常在細菌に必要な機能であるのかを議論することが可能なものもあった。もちろん、これらの機能を持つがゆえにヒトへと常在を果たしたという考え方以外に、ヒトに常在するがゆえに獲得した機能もあるだろう。しかしながら、ヒト常在細菌が持つ機能を次々に明らかにすることは、細菌をヒト常在細菌たらしめる要因から、ヒト共生細菌が発端となった環境や、細菌が経たヒトとの共生までの過程を解明することへと繋げるために重要な研究一つである。最終的には、ヒト常在細菌叢改変を用いたヒトの健康コントロールなどの技術革新も期待される。

謝辞

本研究を進めるにあたり、指導教員として数多くの御指導、御助言を頂きました山田 拓司 准教授に心から感謝致します。また、サンプリングからシーケシングまで様々な実験を御指導を国立遺伝学研究所 黒川 裕美子 博士に心から感謝致します。学部から修士までの研究の御指導を国立遺伝学研究所 黒川 顕 教授、16S rRNA 遺伝子配列解析から研究論文投稿までのご指導を国立遺伝学研究所 森 宙史 助教、論文投稿に関する御指導を山田研究室 水谷 紗弥佳 博士、統計解析と研究内容へのアドバイスについては国立遺伝学研究所 東 光一 博士、実験操作については産業技術総合研究所 中島 信孝 博士、山田研究室 山本 希 博士、元 黒川・中島・山田研究室 研究員 西山 依里 博士、に手厚く御指導を頂き、心から感謝致します。また、サンプル採取からシーケシングまで共同で行って頂いた元 黒川・中島・山田研究室所属 中村 一成 修士、高松 みゆき 修士、伊東 泰雄 修士、安川 玲央 学士に心から感謝致します。特に、山田准教授には、私が学部生の頃より、研究の議論や論文の投稿まで多くの時間をかけて御指導いただきました。重ねて御礼申し上げます。

大学院に入学してから経済的なご支援だけでなく数多くの機会や温かい応援をしてくださった東京工業大学情報生命博士教育院の皆様には心から感謝申し上げます。

また、山田研究室 渡辺 夕紀子 様をはじめとする山田研究室の皆様には、毎日様々な面でお世話になりました。

最後になりましたが、私の長い学生生活を、経済的な支援にとどまらず様々な局面で支えてくださった、福岡の父、母、妹、祖母に心から感謝致します。

付録

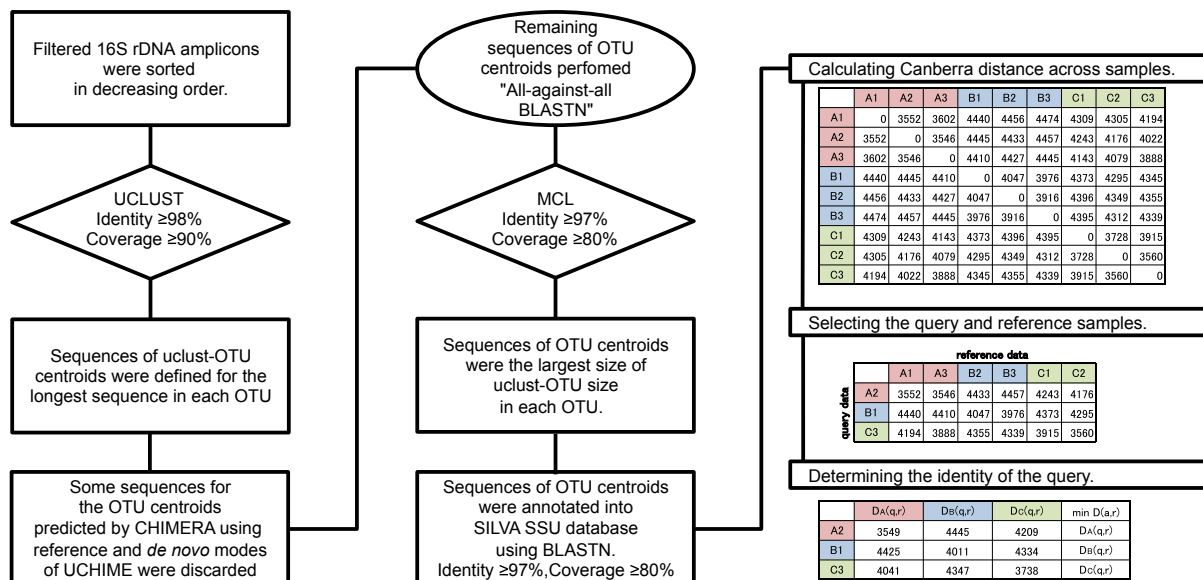


図 2.S1. 16S rRNA 遺伝子配列系統解析から個人特定までの流れ

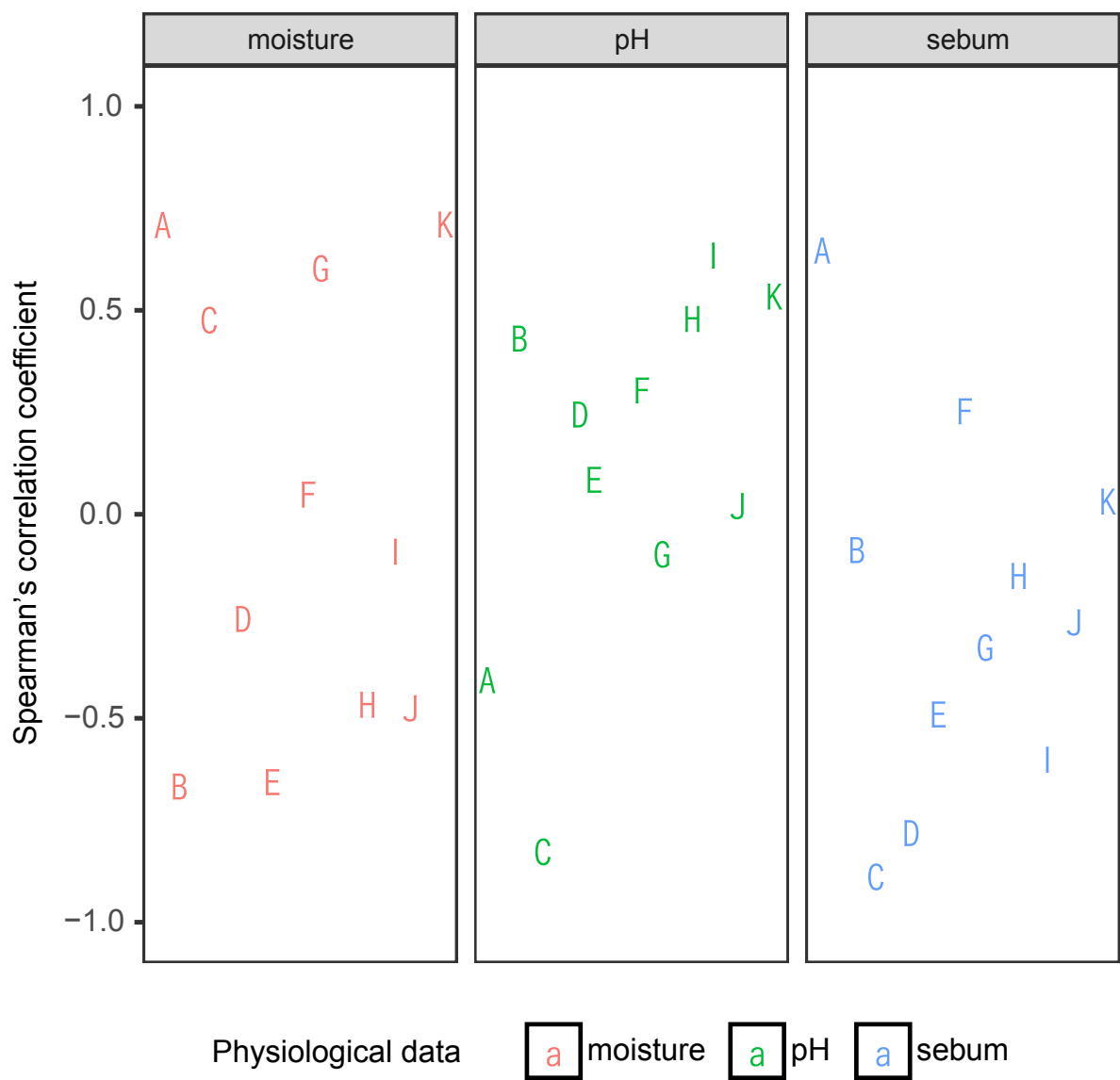


図 2.S2. OTU0001 の相対存在量とヒト皮膚メタデータの相関解析結果

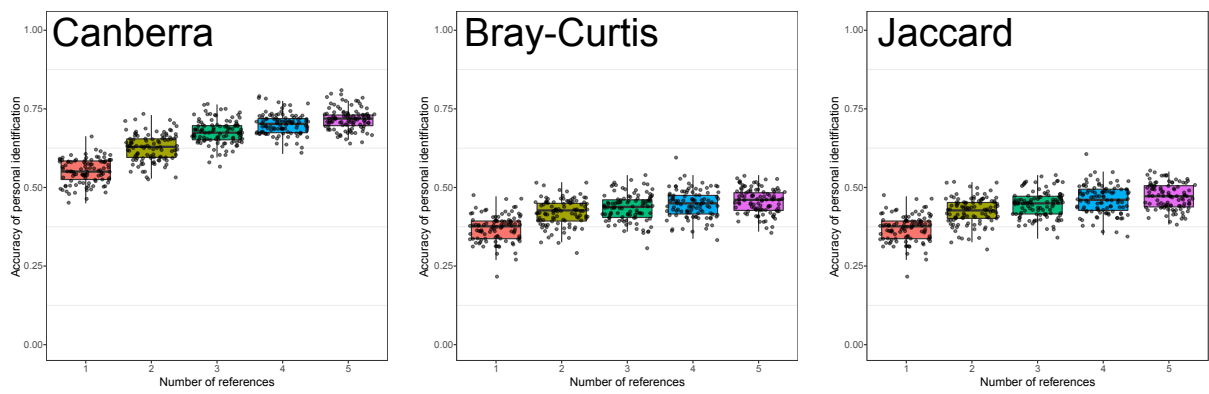
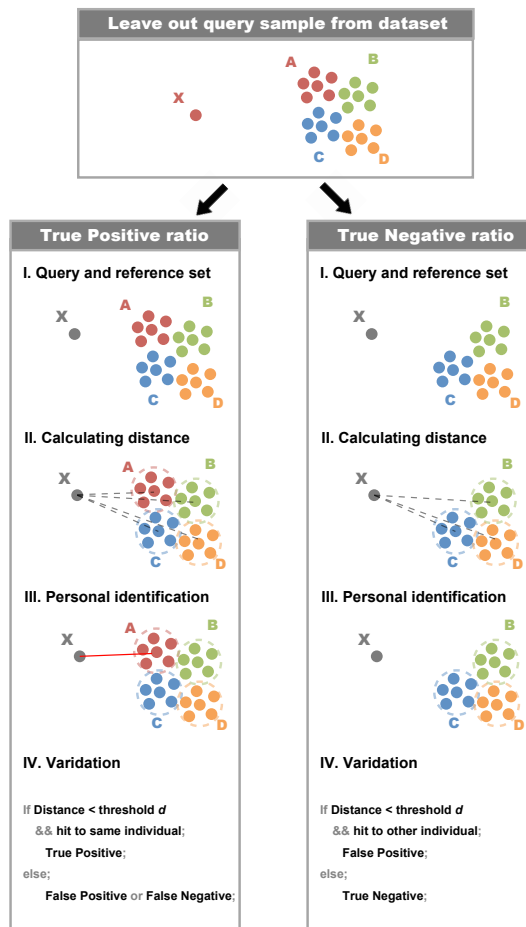


図 2.S3. Canberra 距離, Bray-Curtis 距離, Jaccard 距離を用いた場合の個人特定精度比較

A



B

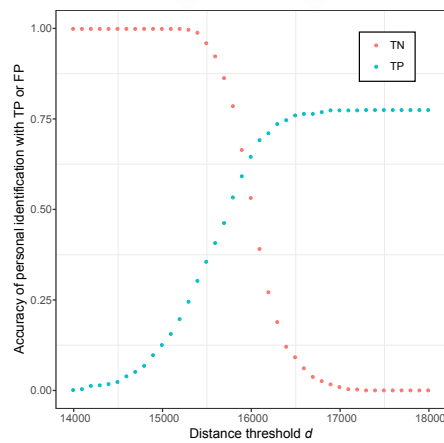


図 2.S4. Public dataset を用いた個人特定手法の拡張版について

表 2.S1. Our data におけるサンプル概要

SampleID	gender	Individual	SampleName	Sampling date (YYYYM- MDD)
N001-Cfh-20130903-0000	male	A	A1	20130903
N001-Cfh-20131001-0000	male	A	A2	20131001
N001-Cfh-20131028-0000	male	A	A3	20131028
N001-Cfh-20140730-1329	male	A	A4	20140730
N001-Cfh-20140813-1100	male	A	A5	20140813
N001-Cfh-20140901-1400	male	A	A6	20140901
N002-Cfh-20130903-0000	male	B	B1	20130903
N002-Cfh-20131001-0000	male	B	B2	20131001
N002-Cfh-20131028-0000	male	B	B3	20131028
N002-Cfh-20140806-1400	male	B	B4	20140806
N002-Cfh-20140822-1430	male	B	B5	20140822
N002-Cfh-20140901-1400	male	B	B6	20140901
N003-Cfh-20130903-0000	male	C	C1	20130903
N003-Cfh-20131001-0000	male	C	C2	20131001
N003-Cfh-20131022-0000	male	C	C3	20131022
N003-Cfh-20140806-1400	male	C	C4	20140806
N003-Cfh-20140813-1100	male	C	C5	20140813
N003-Cfh-20140901-1400	male	C	C6	20140901
N004-Cfh-20130909-0000	male	D	D1	20130909
N004-Cfh-20131001-0000	male	D	D2	20131001
N004-Cfh-20131021-0000	male	D	D3	20131021
N004-Cfh-20140805-1800	male	D	D4	20140805
N004-Cfh-20140813-1100	male	D	D5	20140813
N004-Cfh-20140901-1400	male	D	D6	20140901
N005-Cfh-20131001-0000	male	E	E1	20131001
N005-Cfh-20131021-0000	male	E	E2	20131021
N005-Cfh-20131028-0000	male	E	E3	20131028

表 2.S1. Our data におけるサンプル概要

N005-Cfh-20140806-1400	male	E	E4	20140806
N005-Cfh-20140822-1120	male	E	E5	20140822
N005-Cfh-20140902-1200	male	E	E6	20140902
N006-Cfh-20131001-0000	male	F	F1	20131001
N006-Cfh-20131021-0000	male	F	F2	20131021
N006-Cfh-20131029-0000	male	F	F3	20131029
N006-Cfh-20140807-1200	male	F	F4	20140807
N006-Cfh-20140813-1100	male	F	F5	20140813
N006-Cfh-20140902-1100	male	F	F6	20140902
N007-Cfh-20131001-0000	male	G	G1	20131001
N007-Cfh-20131022-0000	male	G	G2	20131022
N007-Cfh-20131028-0000	male	G	G3	20131028
N007-Cfh-20140806-1500	male	G	G4	20140806
N007-Cfh-20140815-1400	male	G	G5	20140815
N007-Cfh-20140930-0930	male	G	G6	20140930
N008-Cfh-20131001-0000	male	H	H1	20131001
N008-Cfh-20131021-0000	male	H	H2	20131021
N008-Cfh-20131028-0000	male	H	H3	20131028
N008-Cfh-20140730-1335	male	H	H4	20140730
N008-Cfh-20140813-1100	male	H	H5	20140813
N008-Cfh-20140901-1400	male	H	H6	20140901
N012-Cfh-20131001-0000	female	I	I1	20131001
N012-Cfh-20131021-0000	female	I	I2	20131021
N012-Cfh-20131028-0000	female	I	I3	20131028
N012-Cfh-20140807-1100	female	I	I4	20140807
N012-Cfh-20140813-1100	female	I	I5	20140813
N012-Cfh-20140901-1400	female	I	I6	20140901
N013-Cfh-20131001-0000	male	J	J1	20131001
N013-Cfh-20131021-0000	male	J	J2	20131021
N013-Cfh-20131029-0000	male	J	J3	20131029
N013-Cfh-20140730-1354	male	J	J4	20140730

表 2.S1. Our data におけるサンプル概要

N013-Cfh-20140806-1500	male	J	J5	20140806
N013-Cfh-20140813-1100	male	J	J6	20140813
N014-Cfh-20131001-0000	male	K	K1	20131001
N014-Cfh-20131021-0000	male	K	K2	20131021
N014-Cfh-20131029-0000	male	K	K3	20131029
N014-Cfh-20140730-1259	male	K	K4	20140730
N014-Cfh-20140929-1630	male	K	K5	20140929
N014-Cfh-20141024-1300	male	K	K6	20141024

表 2.S2. 各サンプルにおける生リード数と最終リード数

Sample	Number of raw reads	Number of final reads
A1	89045	45638
A2	104342	80349
A3	138690	95589
A4	157460	103630
A5	200863	155798
A6	148109	113871
B1	106196	41088
B2	151489	97198
B3	124566	77526
B4	172515	93318
B5	161292	105973
B6	61663	40974
C1	1147197	526568
C2	160303	112901
C3	139831	96337
C4	149214	84879
C5	135568	103897
C6	136646	97242
D1	247895	169808
D2	138397	109357
D3	106400	82883
D4	161968	106104
D5	169837	135591
D6	153351	116960
E1	133359	103434
E2	130161	97853
E3	109680	82178
E4	163288	100927
E5	133558	104783
E6	127734	102783

表 2.S2. 各サンプルにおける生リード数と最終リード数

F1	161091	122505
F2	80352	62319
F3	198042	138735
F4	196179	121698
F5	140609	111792
F6	117006	91349
G1	144382	111703
G2	111479	82988
G3	131918	99642
G4	154239	98764
G5	144249	112159
G6	154042	122379
H1	132058	101486
H2	169151	120161
H3	59558	45318
H4	166483	105563
H5	179522	143294
H6	169826	133504
I1	125972	99665
I2	90299	67229
I3	157373	120094
I4	185244	116543
I5	179680	145381
I6	111109	88123
J1	117420	89629
J2	128530	98676
J3	180109	132972
J4	119890	76188
J5	165020	103327
J6	141975	114135
K1	145616	110579

表 2.S2. 各サンプルにおける生リード数と最終リード数

K2	72335	55116
K3	148296	111027
K4	136856	89687
K5	135980	32532
K6	185377	43643

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

Individual	OTU	Environment	Isolation site	Citation	Type strain
A	OTU0418	Unclassified	Unclassified	Unclassified	Unclassified
A	OTU1033	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0012	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0013	Unclassified	Unclassified	Unclassified	Rhodococcus corynebacterioides / Rhodococcus en- clensis
B	OTU0016	human skin	human skin	BacDive	Corynebacterium amycolatum
B	OTU0018	human re- lated	human se- men	BacDive	Corynebacterium sin- gulare
B	OTU0020	Unclassified	Unclassified	Unclassified	Mycobacterium neoaurum / Mycobac- terium parafortuitum / Mycobacterium for- tuitum subsp. fortui- tium / Mycobacterium fortuitum subsp. acetamidolyticum
B	OTU0025	Unclassified	Unclassified	Unclassified	Corynebacterium aurimucosum / Corynebacterium minutissimum
B	OTU0031	Unclassified	Unclassified	Unclassified	Neisseria flavescens / Neisseria subflava
B	OTU0036	Unclassified	Unclassified	Unclassified	Gordonia lacunae / Gordonia terrae
B	OTU0053	human re- lated	throat	BacDive	Rothia mucilaginosa

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0054	Unclassified	Unclassified	Unclassified	Acinetobacter junii / Acinetobacter bohemicus / Acinetobacter modestus / Acinetobacter calcoaceticus / Acinetobacter pakistanensis
B	OTU0057	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0059	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0061	Unclassified	Unclassified	Unclassified	Stenotrophomonas maltophilia / Stenotrophomonas pavanii / Pseudomonas beteli / Stenotrophomonas daejeonensis
B	OTU0062	UNKNOWN	UNKNOWN	UNKNOWN	Paracoccus yeei
B	OTU0064	Unclassified	Unclassified	Unclassified	Streptococcus salivarius subsp. thermophilus / Streptococcus salivarius subsp. salivarius / Streptococcus vestibularis
B	OTU0066	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0067	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0071	human related	ear from patient with Otitis media	BacDive	Acinetobacter parvus
B	OTU0072	Unclassified	Unclassified	Unclassified	Unclassified

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0073	Unclassified	Unclassified	Unclassified	Mycobacterium tokaiense / My- cobacterium obuense / Mycobacterium murale / Mycobac- terium senegalense / Mycobacterium hou- stonense / Mycobac- terium farcinogenes / Mycobacterium conceptionense / Mycobacterium pho- caicum / Mycobac- terium mucogenicum / Mycobacterium brisbanense / My- cobacterium engbaekii / Mycobacterium in- subricum / Mycobac- terium fortuitum subsp. fortuitum / Mycobacterium fortuitum subsp. acetamidolyticum
B	OTU0074	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0076	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0079	Unclassified	Unclassified	Unclassified	Chryseobacterium ho- minis / Chryseobac- terium hagamense / Chryseobacterium flavum

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0085	Unclassified	Unclassified	Unclassified	Acinetobacter seifer- tii / Acinetobacter calcoaceticus / Acine- tobacter ursingii / Acinetobacter pittii
B	OTU0087	Unclassified	Unclassified	Unclassified	Sphingomonas desic- cabilis / Sphingobium abikonense / Sphin- gomonas molluscorum / Sphingobium vermi- composti / Sphingob- ium indicum / Sphin- gobium baderi
B	OTU0091	Unclassified	Unclassified	Unclassified	Streptococcus mi- tis / Streptococcus pseudopneumoniae
B	OTU0092	others	cheddar cheese	BacDive	Brevibacterium casei
B	OTU0097	human re- lated	human vaginal fluid	BacDive	Dermabacter vaginalis
B	OTU0102	human re- lated	septic fin- ger	BacDive	Haemophilus parain- fluenzae
B	OTU0104	human re- lated	human saliva	”Sakamoto M, Int J Syst Evol Microbiol, 2015”	Porphyromonas pas- teri

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0109	Unclassified	Unclassified	Unclassified	Yokenella regens- burgei / Enterobacter ludwigii / Klebsiella michiganensis / En- terobacter cloacae subsp. dissolvens / Kosakonia sacchari / Enterobacter cloacae subsp. cloacae / Enterobacter cantero- genus / Kosakonia oryzae / Klebsiella pneumoniae subsp. pneumoniae / Kleb- siella pneumoniae subsp. ozaenae / Klebsiella oxytoca / Klebsiella quasip- neumoniae subsp. similipneumoniae etc.
B	OTU0111	Unclassified	Unclassified	Unclassified	Sphingobium rhi- zovicinum / Sphingo- bium lucknowense / Sphingobium indicum
B	OTU0117	Unclassified	Unclassified	Unclassified	Massilia brevitalea / Massilia jejuensis / Massilia umbonata
B	OTU0118	Unclassified	Unclassified	Unclassified	Unclassified

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0120	Unclassified	Unclassified	Unclassified	Bosea lupini / Bosea vestrisii / Bosea mas-siliensis / Bosea eneae / Bosea thiooxidans / Bosea robiniae
B	OTU0121	Unclassified	Unclassified	Unclassified	Kocuria marina / Kocuria indica
B	OTU0122	Unclassified	Unclassified	Unclassified	Prevotella melanino-genica / Prevotella scopos / Prevotella je-juni
B	OTU0127	human re-lated	human blood culture	BacDive	Corynebacterium coyleae
B	OTU0131	Unclassified	Unclassified	Unclassified	Veillonella parvula / Veillonella denti-cariosi / Veillonella dispar / Veillonella atypica
B	OTU0136	Unclassified	Unclassified	Unclassified	Neisseria elongata subsp. glycolytica / Neisseria elongata subsp. elongata
B	OTU0142	air	air in the Russian space lab. Mir	BacDive	Rothia aeria

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0143	Unclassified	Unclassified	Unclassified	Staphylococcus cohnii subsp. urealyticus / Staphylococcus cohnii subsp. cohnii / Staphylococcus arlettae / Staphy- lococcus gallinarum / Staphylococcus nepalensis
B	OTU0146	Unclassified	Unclassified	Unclassified	Aeromonas taiwanen- sis / Aeromonas sanarellii / Aeromonas punc- tata subsp. punctata / Aeromonas punc- tata subsp. caviae / Aeromonas hy- drophila subsp. hy- drophila / Aeromonas hydrophila subsp. anaerogenes / Aeromonas aquar- iorum / Aeromonas dhakensis / Aeromonas rivipol- lensis / Aeromonas media / Aeromonas enteropelogenes / Aeromonas hy- drophila subsp. ranae

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0155	Unclassified	Unclassified	Unclassified	Chryseobacterium culicis / Chryseobac- terium nakagawai / Chryseobacterium arthrosphaerae
B	OTU0159	water	untreated municipal wastewater	BacDive	Cloacibacterium nor- manense
B	OTU0161	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0163	human re- lated	human sputum	BacDive	Corynebacterium du- rum
B	OTU0169	Unclassified	Unclassified	Unclassified	Gordonia sputi / Gor- donia aichiensis / Gor- donia otitidis
B	OTU0177	Unclassified	Unclassified	Unclassified	Kluyvera georgiana / Kluyvera cryocrescens / Kluyvera intermedia
B	OTU0178	water	the al- kaline wash- waters of edible olives	BacDive	Alkalibacterium olivapovliticus
B	OTU0179	others	fish of the South Atlantic Ocean off South Africa	"de Beer H, Int J Syst Evol Microbiol, 2016"	Chryseobacterium piscium

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0180	Unclassified	Unclassified	Unclassified	Janthinobacterium agaricidamnorum / Janthinobacterium lividum / Aquaspiril- lum arcticum / Her- miniimonas saxobsi- dens / Herminiimonas fonticola / Duganella phyllosphaerae
B	OTU0181	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0185	Unclassified	Unclassified	Unclassified	Staphylococcus warneri / Staphy- lococcus pasteurii
B	OTU0195	human re- lated	human sputum	BacDive	Corynebacterium du- rum
B	OTU0200	human re- lated	oral calcu- lus	BacDive	Corynebacterium ma- truchotii
B	OTU0203	Unclassified	Unclassified	Unclassified	Propionibacterium propionicum / Propi- onibacterium avidum
B	OTU0212	water	drinking water	BacDive	Massilia aurea
B	OTU0213	human re- lated	perianal abscess	BacDive	Actinomyces turicen- sis
B	OTU0221	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0231	human re- lated	subacute bacterial endocardi- tis	BacDive	Streptococcus sangui- nis
B	OTU0236	human skin	human skin	BacDive	Dermabacter hominis

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0248	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0260	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0275	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0279	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0296	UNKNOWN	UNKNOWN	UNKNOWN	Rhodanobacter glycinis
B	OTU0305	Unclassified	Unclassified	Unclassified	Kocuria rhizophila / Kocuria arsenatis
B	OTU0338	human related	human sputum of a 82-year-old woman with pulmonary disease	BacDive	Corynebacterium kroppenstedtii
B	OTU0355	Unclassified	Unclassified	Unclassified	Cloacibacterium normanense / Cloacibacterium rupense

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0357	Unclassified	Unclassified	Unclassified	<i>Buttiauxella noackiae</i> / <i>Buttiauxella bren-</i> <i>nerae</i> / <i>Buttiauxella</i> <i>agrestis</i> / <i>Buttiauxella</i> <i>izardii</i> / <i>Buttiauxella</i> <i>gaviniae</i> / <i>Butti-</i> <i>auxella ferragutiae</i> / <i>Pectobacterium</i> <i>aroidearum</i> / <i>Butti-</i> <i>auxella warmboldiae</i> / <i>Serratia quinivorans</i> / <i>Serratia protea-</i> <i>maculans</i> / <i>Serratia</i> <i>grimesii</i> / <i>Enterobac-</i> <i>ter soli</i> / <i>Dickeya</i> <i>chrysanthemi</i> / <i>Ser-</i> <i>ratia plymuthica</i> / <i>Serratia myotis</i> / <i>Serratia ficaria</i> / <i>Serratia entomophila</i> / <i>Pectobacterium</i> <i>atrosepticum</i> / <i>Dick-</i> <i>eya zeae</i> / <i>Dickeya</i> <i>dadantii</i> subsp. <i>dief-</i> <i>fenbachiae</i> / <i>Serratia</i> <i>fonticola</i> / <i>Enter-</i> <i>obacter asburiae</i> / <i>Pectobacterium</i> <i>wasabiae</i> / <i>Cronobac-</i> <i>ter dublinensis</i> subsp. <i>lausannensis</i> / <i>Dick-</i> <i>eya aquatica</i> / <i>Serra-</i> <i>tia liquefaciens</i>
---	---------	--------------	--------------	--------------	--

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

B	OTU0373	Unclassified	Unclassified	Unclassified	Pseudomonas nitroreducens / Pseudomonas nitritireducens
B	OTU0433	Unclassified	Unclassified	Unclassified	Unclassified
B	OTU0489	human related	human bone marrow (leprosy patient)	BacDive	Corynebacterium tuberculostearicum
C	OTU0083	water	Antarctica	"Liu H, Int J Syst Evol Microbiol, 2000"	Micrococcus antarcticus
C	OTU0176	plant	root nodules of Sphaerophysa salsula	"Deng ZS, Int J Syst Evol Microbiol, 2011"	Paracoccus sphaerophysae
C	OTU0193	human skin	human skin	"Dekio I, Int J Syst Evol Microbiol, 2015"	Propionibacterium acnes subsp. acnes
C	OTU0198	Unclassified	Unclassified	Unclassified	Unclassified
C	OTU0337	Unclassified	Unclassified	Unclassified	Unclassified
C	OTU0341	human related	human bone marrow (leprosy patient)	BacDive	Corynebacterium tuberculostearicum

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

C	OTU0377	water	lake sedi- ment	BacDive	Arsenicicoccus boli- densis
C	OTU0384	human skin	human skin	”Dekio I, Int J Syst Evol Microbiol, 2015”	Propionibacterium ac- nes subsp. acnes
C	OTU0601	Unclassified	Unclassified	Unclassified	Unclassified
C	OTU0748	Unclassified	Unclassified	Unclassified	Unclassified
D	OTU0112	Unclassified	Unclassified	Unclassified	Unclassified
D	OTU0130	soil	soil crusts	BacDive	Sphingomonas muco- sissima
D	OTU0229	Unclassified	Unclassified	Unclassified	Propionibacterium ac- nes subsp. acnes / Propionibacterium ac- nes subsp. elongatum
D	OTU0243	Unclassified	Unclassified	Unclassified	Pseudomonas psy- chrotolerans / Pseu- domonas oleovorans subsp. oleovorans
E	OTU0039	others	pig heart	BacDive	Corynebacterium suicordis
E	OTU0119	Unclassified	Unclassified	Unclassified	Unclassified
E	OTU0128	Unclassified	Unclassified	Unclassified	Dietzia timorensis / Dietzia cinnamea
E	OTU0187	Unclassified	Unclassified	Unclassified	Corynebacterium pil- barensis / Corynebac- terium afermentans subsp. afermentans
E	OTU0189	Unclassified	Unclassified	Unclassified	Unclassified

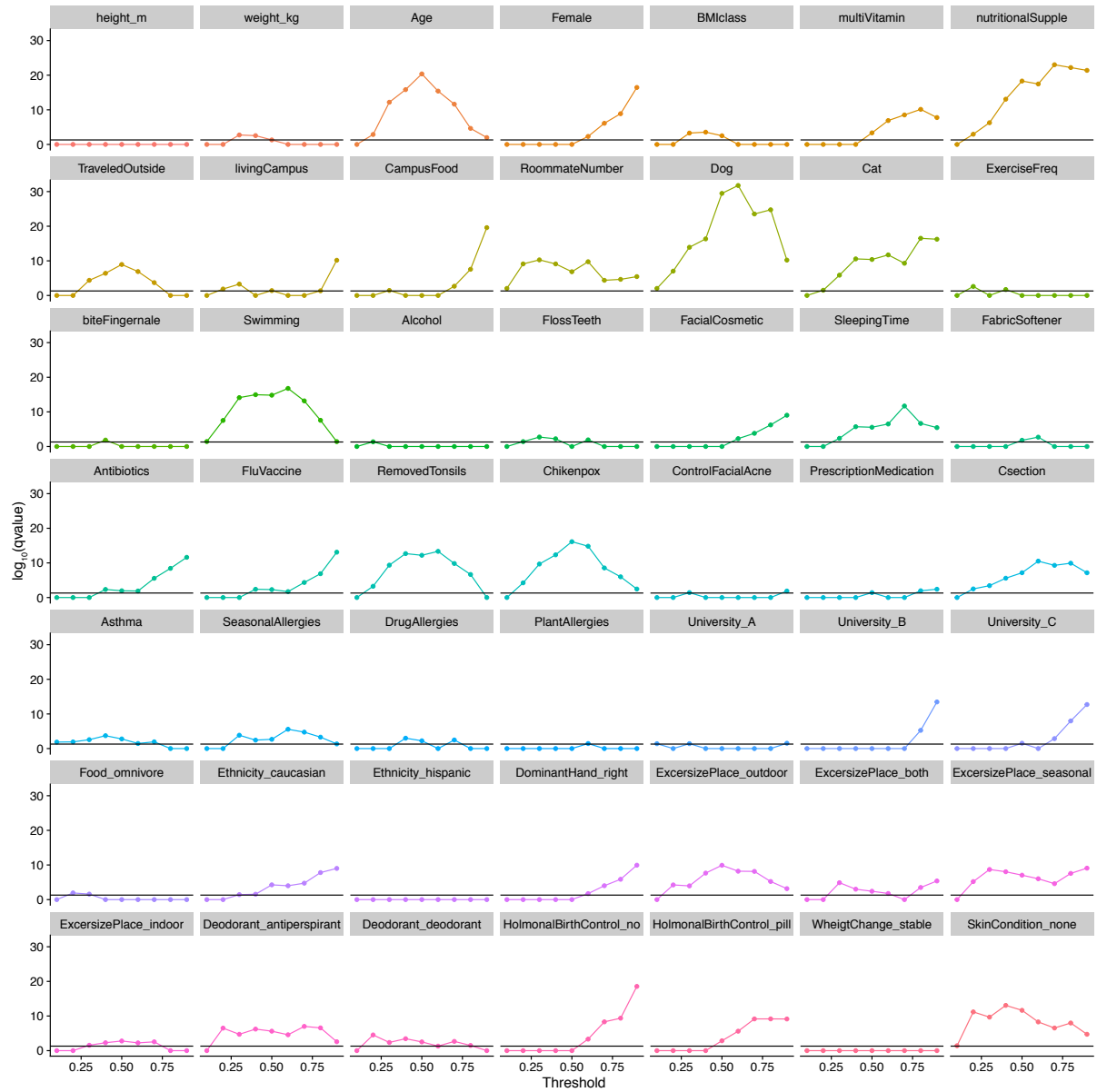
表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

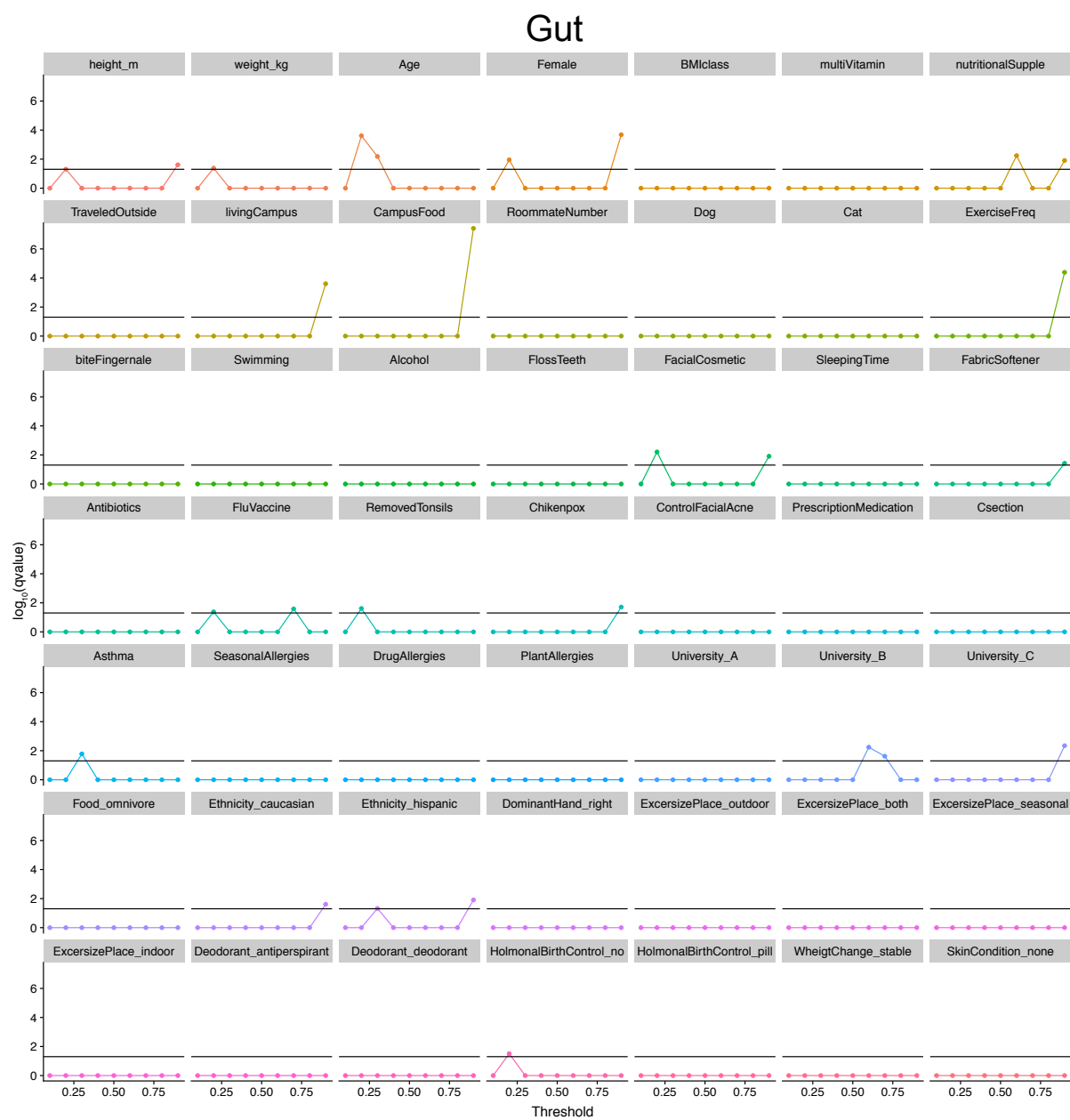
E	OTU0217	Unclassified	Unclassified	Unclassified	Staphylococcus haemolyticus / Staphylococcus de- vriesei / Staphylococ- cus hominis subsp. hominis / Staphylo- coccus saprophyticus subsp. saprophyticus
E	OTU0615	Unclassified	Unclassified	Unclassified	Unclassified
E	OTU0704	Unclassified	Unclassified	Unclassified	Unclassified
F	OTU0186	Unclassified	Unclassified	Unclassified	Microvirga aerilata / Microvirga zambiensis / Microvirga floccu- lans
F	OTU0371	human skin	human skin	”Dekio I, Int J Syst Evol Microbiol, 2015”	Propionibacterium ac- nes subsp. elongatum
F	OTU0438	Unclassified	Unclassified	Unclassified	Unclassified
F	OTU0530	Unclassified	Unclassified	Unclassified	Unclassified
F	OTU1037	Unclassified	Unclassified	Unclassified	Unclassified
G	OTU0228	Unclassified	Unclassified	Unclassified	Staphylococcus schleiferi subsp. schleiferi / Staphy- lococcus schleiferi subsp. coagulans
J	OTU0147	Unclassified	Unclassified	Unclassified	Paracoccus sangui- nis / Paracoccus panacisoli

表 2.S3. ある被験者において全てのタイムポイントで観測された OTU のリスト

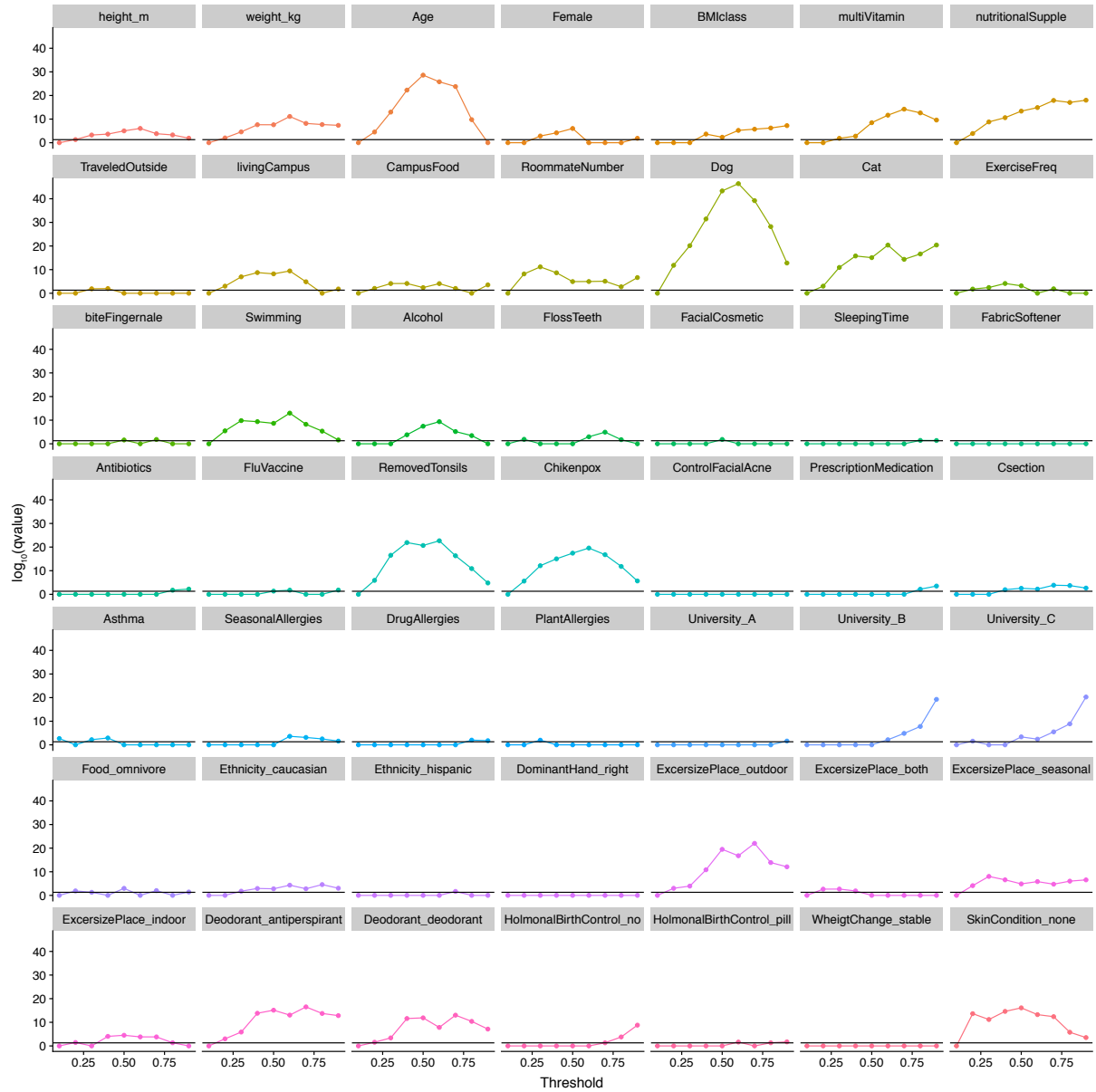
J	OTU0211	Unclassified	Unclassified	Unclassified	Corynebacterium macginleyi / Corynebacterium accolens
J	OTU0222	Unclassified	Unclassified	Unclassified	Unclassified
J	OTU0307	Unclassified	Unclassified	Unclassified	Kytococcus seden- tarius / Kytococcus schroeteri
J	OTU0324	others	activated sludge	BacDive	Micrococcus flavus
J	OTU0352	human re- lated	dental plaque	BacDive	Prevotella buccalis
J	OTU0415	Unclassified	Unclassified	Unclassified	Unclassified
J	OTU0420	Unclassified	Unclassified	Unclassified	Unclassified
J	OTU0588	human re- lated	human blood culture	BacDive	Corynebacterium mu- cifaciens
J	OTU0796	Unclassified	Unclassified	Unclassified	Unclassified
J	OTU0881	Unclassified	Unclassified	Unclassified	Unclassified

Forehead





Palm



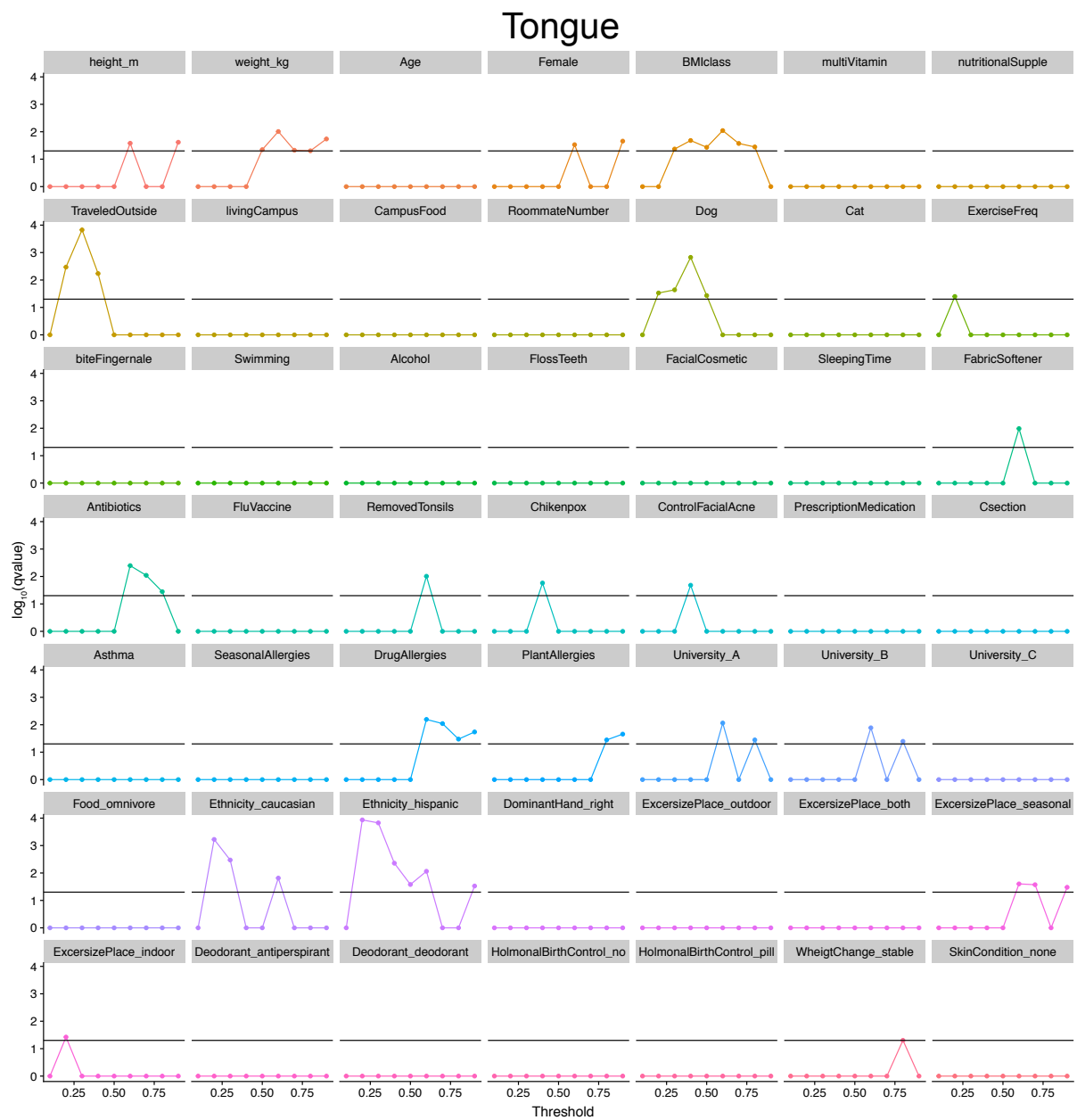


図 3.S1. GLM による各閾値での各メタデータの推定値の Wald 検定の結果

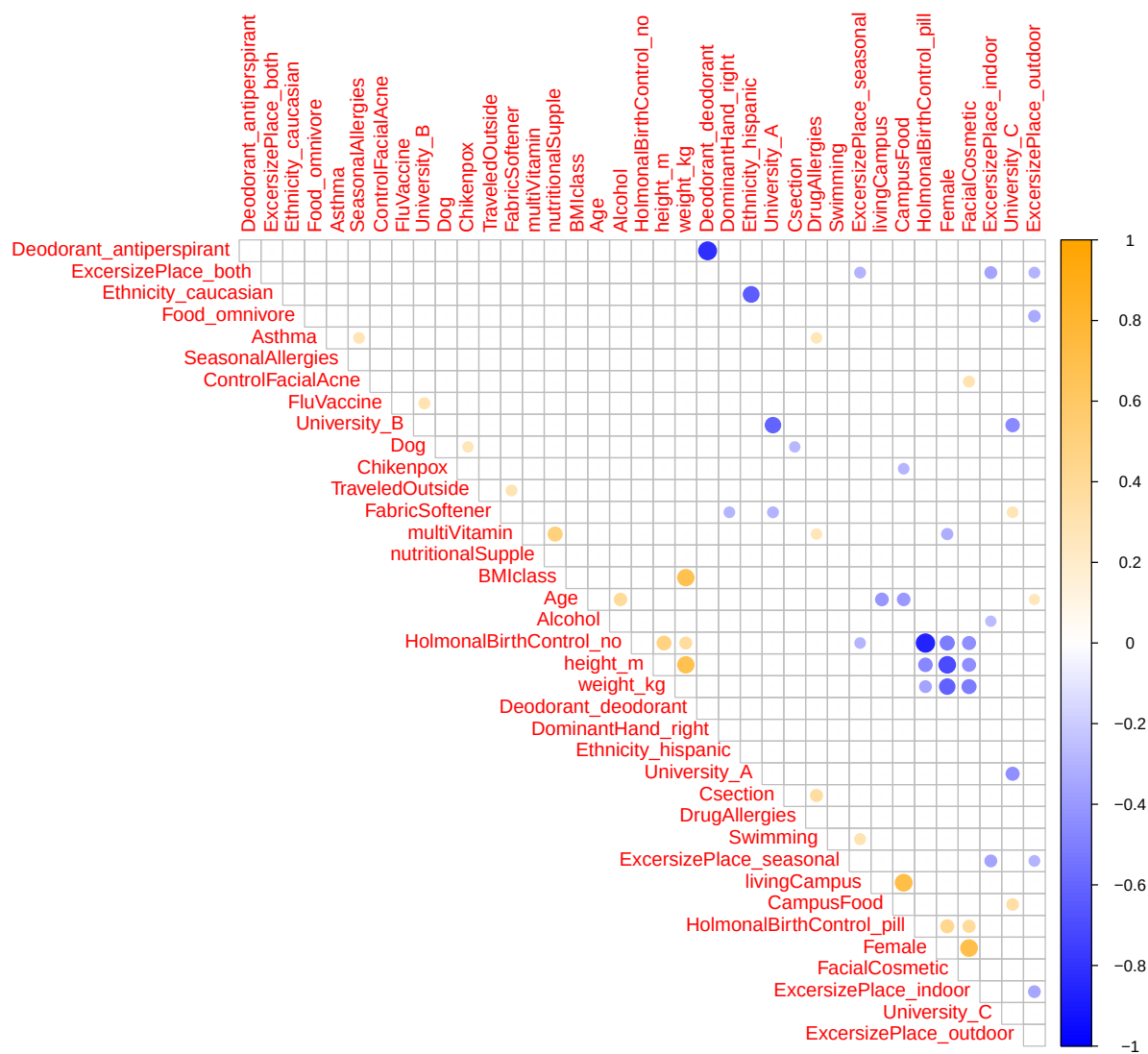


図 3.S2. 被験者メタデータ間の Spearman 順位相関係数 (色付きは $p \leq 0.05$)

表 3.S1. 各被験者メタデータの情報

metadata	meaning.word	data.type
Age	Age	continuous
height_m	height_m	continuous
weight_kg	weight_kg	continuous
PersonalID	PersonalID	PersonalID
Gender	Female	two_category
BMIclass	BMIclass	ordinal
University	University	categorical
Question1	Food	categorical
Question2	multiVitamin	two_category
Question3	nutritionalSupple	two_category
Question4	LactoseIntolerant	two_category
Question5	GlutenIntolerant	two_category
Question6a	PeanutsAllergy	two_category
Question6b	TreenutsAllergy	two_category
Question6c	ShellfishAllergy	two_category
Question6d	OtherFoodAllergy	two_category
Question7	Diabete	two_category
Question8	SpecialDiet	two_category
Question9	Ethnicity	categorical
Question10	Moving	ordinal
Question11	TraveledOutside	ordinal
Question12	livingCampus	two_category
Question13	livingSorority	two_category
Question14	CampusFood	two_category
Question15	RoommateNumber	ordinal
Question16	Dog	two_category
Question17	Cat	two_category
Question18	DominantHand	categorical
Question19	ArtificialTnning	two_category
Question20	ExerciseFreq	ordinal

表 3.S1. 各被験者メタデータの情報

Question21	ExcercisePlace	categorical
Question22	biteFingernale	two_category
Question23	Swimming	ordinal
Question24	SmokeCigarettes	ordinal
Question25	Alcohol	ordinal
Question26	BlushTooth	ordinal
Question27	FlossTeeth	ordinal
Question28	FacialCosmetic	ordinal
Question29	Deodorant	categorical
Question30	SleepingTime	ordinal
Question31	FabricSoftener	two_category
Question32	Antibiotics	ordinal
Question33	FluVaccine	ordinal
Question34	HolmonalBirthControl	categorical
Question35	Pregnant	two_category
Question36	WheightChange	categorical
Question37	RemovedTonsils	two_category
Question38	RemovedAppendix	two_category
Question39	Chickenpox	two_category
Question40	MedicationFacialAcne	two_category
Question41	ControlFacialAcne	two_category
Question42	PrescriptionMedication	two_category
Question43	Csection	two_category
Question44	Phenylketonuria	two_category
Question45	Asthma	two_category
Question46	SeasonalAllergies	two_category
Question47a	DrugAllergies	two_category
Question47c	BeestingsAllergies	two_category
Question47d	PlantAllergies	two_category
Question47e	SunAllergies	two_category
Question48	BowelDisease	categorical

表 3.S1. 各被験者メタデータの情報

Question49	SkinCondition	categorical
------------	---------------	-------------

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

body site	phyla	pvalue	qvalue.BH	ko_definition
forehead	Actinobacteria	0	0.0182	K11733: lysP; lysine-specific permease
forehead	Actinobacteria	0	0.0182	K14645: K14645; serine protease [EC:3.4.21.-]
forehead	Actinobacteria	0	0.0261	K11065: tpx; thiol peroxidase, atypical 2-Cys peroxiredoxin [EC:1.11.1.15]
forehead	Actinobacteria	0	0.0261	K02445: glpT; MFS transporter, OPA family, glycerol-3-phosphate transporter
forehead	Actinobacteria	0.0001	0.0339	K03319: TC.DASS; divalent anion:Na ⁺ symporter, DASS family
forehead	Actinobacteria	0.0001	0.0339	K01989: ABC.X4.S; putative ABC transport system substrate-binding protein
forehead	Actinobacteria	0.0001	0.0339	K05832: ABC.X4.P; putative ABC transport system permease protein
forehead	Actinobacteria	0.0001	0.0339	K05833: ABC.X4.A; putative ABC transport system ATP-binding protein
forehead	Actinobacteria	0.0001	0.0339	K08996: yagU; putative membrane protein
forehead	Actinobacteria	0.0001	0.0339	K19131: csb1; CRISPR-associated protein Csb1
forehead	Actinobacteria	0.0001	0.0339	K19132: csb2; CRISPR-associated protein Csb2

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Actinobacteria	0.0002	0.0717	K03311: TC.LIVCS; branched-chain amino acid:cation transporter, LIVCS family
forehead	Actinobacteria	0.0002	0.0717	K04561: norB; nitric oxide reductase subunit B [EC:1.7.2.5]
forehead	Actinobacteria	0.0003	0.0866	K09125: K09125; uncharacterized protein
forehead	Actinobacteria	0.0003	0.0866	K19158: yoeB; toxin YoeB [EC:3.1.-.-]
forehead	Actinobacteria	0.0004	0.1054	K01575: alsD, budA, aldC; acetolactate decarboxylase [EC:4.1.1.5]
forehead	Actinobacteria	0.0004	0.1054	K03488: licT, bglG; beta-glucoside operon transcriptional antiterminator
forehead	Actinobacteria	0.0005	0.1203	K00773: tgt, QTRT1; queuine tRNA-ribosyltransferase [EC:2.4.2.29]
forehead	Actinobacteria	0.0013	0.1622	K01223: E3.2.1.86B, bglA; 6-phospho-beta-glucosidase [EC:3.2.1.86]
forehead	Actinobacteria	0.0013	0.1622	K01623: ALDO; fructose-bisphosphate aldolase, class I [EC:4.1.2.13]
forehead	Actinobacteria	0.0013	0.1622	K02755: PTS-Bgl-EIIA, bglF, bglP; PTS system, beta-glucoside-specific IIA component [EC:2.7.1.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Actinobacteria	0.0013	0.1622	K02756: PTS-Bgl-EIIB, bglF, bglP; PTS system, beta-glucoside-specific IIB component [EC:2.7.1.-]
forehead	Actinobacteria	0.0013	0.1622	K02757: PTS-Bgl-EIIC, bglF, bglP; PTS system, beta-glucoside-specific IIC component
forehead	Actinobacteria	0.0013	0.1622	K03304: tehA; tellurite resistance protein
forehead	Actinobacteria	0.0013	0.1622	K00150: gap2; glyceraldehyde-3-phosphate dehydrogenase (NAD(P)) [EC:1.2.1.59]
forehead	Actinobacteria	0.0013	0.1622	K01226: treC; trehalose-6-phosphate hydrolase [EC:3.2.1.93]
forehead	Actinobacteria	0.0013	0.1622	K02795: PTS-Man-EIIC, manY; PTS system, mannose-specific IIC component
forehead	Actinobacteria	0.0013	0.1622	K02796: PTS-Man-EIID, manZ; PTS system, mannose-specific IID component
forehead	Actinobacteria	0.0013	0.1622	K03779: ttdA; L(+)-tartrate dehydratase alpha subunit [EC:4.2.1.32]
forehead	Actinobacteria	0.0013	0.1622	K03780: ttdB; L(+)-tartrate dehydratase beta subunit [EC:4.2.1.32]
forehead	Actinobacteria	0.0013	0.1622	K19133: csb3; CRISPR-associated protein Csb3

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Actinobacteria	0.0013	0.1622	K13821: putA; RHH-type transcriptional regulator, proline utilization regulon repressor / proline dehydrogenase / delta 1-pyrroline-5-carboxylate dehydrogenase [EC:1.5.5.2 1.2.1.88]
forehead	Actinobacteria	0.002	0.2059	K11928: putP; sodium/proline symporter
forehead	Actinobacteria	0.0029	0.251	K01792: E5.1.3.15; glucose-6-phosphate 1-epimerase [EC:5.1.3.15]
forehead	Actinobacteria	0.0029	0.251	K03475: PTS-Ula-EIIC, ulaA, sgaT; PTS system, ascorbate-specific IIC component
forehead	Actinobacteria	0.0029	0.251	K18118: aarC, cat1; succinyl-CoA:acetate CoA-transferase [EC:2.8.3.18]
forehead	Actinobacteria	0.0029	0.251	K00016: LDH, ldh; L-lactate dehydrogenase [EC:1.1.1.27]
forehead	Actinobacteria	0.003	0.251	K10254: ohyA, sph; oleate hydratase [EC:4.2.1.53]
forehead	Actinobacteria	0.003	0.251	K11737: cycA; D-serine/D-alanine/glycine transporter
forehead	Actinobacteria	0.0047	0.3264	K00980: tagD; glycerol-3-phosphate cytidylyltransferase [EC:2.7.7.39]
forehead	Actinobacteria	0.0047	0.3264	K07146: K07146; UPF0176 protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Actinobacteria	0.005	0.3264	K00112: glpB; glycerol-3-phosphate dehydrogenase subunit B [EC:1.1.5.3]
forehead	Actinobacteria	0.005	0.3264	K00113: glpC; glycerol-3-phosphate dehydrogenase subunit C [EC:1.1.5.3]
forehead	Actinobacteria	0.005	0.3264	K01760: metC; cysteine-S-conjugate beta-lyase [EC:4.4.1.13]
forehead	Actinobacteria	0.005	0.3264	K02793: PTS-Man-EIIA, manX; PTS system, mannose-specific IIA component [EC:2.7.1.191]
forehead	Actinobacteria	0.005	0.3264	K02794: PTS-Man-EIIB, manX; PTS system, mannose-specific IIB component [EC:2.7.1.191]
forehead	Actinobacteria	0.005	0.3264	K02808: PTS-Scr-EIIA, scrA, ptsS; PTS system, sucrose-specific IIA component [EC:2.7.1.211]
forehead	Actinobacteria	0.005	0.3264	K02858: ribB, RIB3; 3,4-dihydroxy 2-butanone 4-phosphate synthase [EC:4.1.99.12]
forehead	Actinobacteria	0.005	0.3264	K13829: aroKB; shikimate kinase / 3-dehydroquinate synthase [EC:2.7.1.71 4.2.3.4]
forehead	Actinobacteria	0.005	0.3264	K21011: pelF; polysaccharide biosynthesis protein PelF

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Actinobacteria	0.005	0.3264	K21739: rclA; probable pyridine nucleotide-disulfide oxidoreductase
forehead	Bacteroidetes	0.0002	0.4557	K01834: PGAM, gpmA; 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase [EC:5.4.2.11]
forehead	Bacteroidetes	0.001	0.4609	K01623: ALDO; fructose-bisphosphate aldolase, class I [EC:4.1.2.13]
forehead	Bacteroidetes	0.001	0.4609	K05791: terZ; tellurium resistance protein TerZ
forehead	Bacteroidetes	0.0011	0.4609	K00077: panE, apbA; 2-dehydropantoate 2-reductase [EC:1.1.1.169]
forehead	Bacteroidetes	0.0011	0.4609	K01744: aspA; aspartate ammonia-lyase [EC:4.3.1.1]
forehead	Bacteroidetes	0.0011	0.4609	K06893: K06893; uncharacterized protein
forehead	Bacteroidetes	0.0014	0.4609	K01185: E3.2.1.17; lysozyme [EC:3.2.1.17]
forehead	Bacteroidetes	0.0014	0.4609	K20459: nukF, mcdF, sboF; lantibiotic transport system ATP-binding protein
forehead	Bacteroidetes	0.0024	0.5869	K01728: pel; pectate lyase [EC:4.2.2.2]
forehead	Bacteroidetes	0.0024	0.5869	K07271: licD; lipopolysaccharide cholinephosphotransferase [EC:2.7.8.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Bacteroidetes	0.0024	0.5869	K13771: nsrR; Rrf2 family transcriptional regulator, nitric oxide-sensitive transcriptional repressor
forehead	Bacteroidetes	0.0028	0.6076	K05795: terD; tellurium resistance protein TerD
forehead	Bacteroidetes	0.0029	0.6076	K20534: gtrB, csbB; polyisoprenyl-phosphate glycosyltransferase [EC:2.4.-.-]
forehead	Bacteroidetes	0.0045	0.7975	K00763: pncB, NAPRT1; nicotinate phosphoribosyltransferase [EC:6.3.4.21]
forehead	Bacteroidetes	0.0047	0.7975	K02230: cobN; cobaltochelatase CobN [EC:6.6.1.2]
forehead	Bacteroidetes	0.0047	0.7975	K04720: cobD; threonine-phosphate decarboxylase [EC:4.1.1.81]
forehead	Firmicutes	0.0001	0.0103	K05873: cyaB; adenylate cyclase, class 2 [EC:4.6.1.1]
forehead	Firmicutes	0.0002	0.0141	K09952: csn1, cas9; CRISPR-associated endonuclease Csn1 [EC:3.1.-.-]
forehead	Firmicutes	0.0006	0.0314	K09765: queH; epoxyqueuosine reductase [EC:1.17.99.6]
forehead	Firmicutes	0.0008	0.0377	K06993: K06993; ribonuclease H-related protein
forehead	Firmicutes	0.001	0.0443	K03756: potE; putrescine:ornithine antiporter
forehead	Firmicutes	0.001	0.0443	K07482: K07482; transposase, IS30 family

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Firmicutes	0.0022	0.0891	K02445: glpT; MFS transporter, OPA family, glycerol-3-phosphate transporter
forehead	Firmicutes	0.0022	0.0891	K19137: csn2; CRISPR-associated protein Csn2
forehead	Firmicutes	0.0023	0.0891	K05363: murM; serine/alanine adding enzyme [EC:2.3.2.10]
forehead	Firmicutes	0.0023	0.0891	K13727: pdc; phenolic acid decarboxylase [EC:4.1.1.-]
forehead	Firmicutes	0.0027	0.0973	K14205: mprF, fmtC; phosphatidylglycerol lysyltransferase [EC:2.3.2.3]
forehead	Firmicutes	0.0044	0.138	K00799: GST, gst; glutathione S-transferase [EC:2.5.1.18]
forehead	Firmicutes	0.0044	0.138	K07694: vraR; two-component system, NarL family, vancomycin resistance associated response regulator VraR
forehead	Firmicutes	0.0044	0.138	K16235: mmuP; S-methylmethionine transporter
forehead	Firmicutes	0.0045	0.138	K02530: lacR; DeoR family transcriptional regulator, lactose phosphotransferase system repressor
forehead	Firmicutes	0.0045	0.138	K13677: dgs, bgsA; 1,2-diacylglycerol-3-alpha-glucose alpha-1,2-glucosyltransferase [EC:2.4.1.208]
forehead	Proteobacteria	0.0001	0.2219	K06906: K06906; uncharacterized protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Proteobacteria	0.0001	0.2219	K02429: fucP; MFS transporter, FHS family, L-fucose permease
forehead	Proteobacteria	0.0001	0.2219	K06908: K06908; uncharacterized protein
forehead	Proteobacteria	0.0001	0.2219	K07504: K07504; predicted type IV restriction endonuclease
forehead	Proteobacteria	0.0002	0.2495	K03300: TC.CITMHS; citrate-Mg ²⁺ :H ⁺ or citrate-Ca ²⁺ :H ⁺ symporter, CitMHS family
forehead	Proteobacteria	0.0002	0.2495	K06905: K06905; uncharacterized protein
forehead	Proteobacteria	0.0003	0.2495	K19355: MAN; mannan endo-1,4-beta-mannosidase [EC:3.2.1.78]
forehead	Proteobacteria	0.0003	0.2495	K07318: K07318; adenine-specific DNA-methyltransferase [EC:2.1.1.72]
forehead	Proteobacteria	0.0003	0.2566	K06148: ABCC-BAC; ATP-binding cassette, subfamily C, bacterial
forehead	Proteobacteria	0.0006	0.3753	K18829: mvpT, vapB; anti-toxin VapB
forehead	Proteobacteria	0.0007	0.4046	K12685: ssp; subtilase-type serine protease [EC:3.4.21.-]
forehead	Proteobacteria	0.0007	0.4046	K15629: CYP152A; fatty-acid peroxygenase [EC:1.11.2.4]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Proteobacteria	0.0009	0.4455	K07031: hddA; D-glycero-alpha-D-manno-heptose-7-phosphate kinase [EC:2.7.1.168]
forehead	Proteobacteria	0.0011	0.4455	K14266: prnA, rebH, ktzQ; tryptophan 7-halogenase [EC:1.14.19.9]
forehead	Proteobacteria	0.0011	0.4455	K16211: malY, malT; maltose/moltooligosaccharide transporter
forehead	Proteobacteria	0.0011	0.4455	K07265: kpsS, lipB; capsular polysaccharide export protein
forehead	Proteobacteria	0.0012	0.4455	K07017: K07017; uncharacterized protein
forehead	Proteobacteria	0.0013	0.4455	K16088: TC.FEV.OM1, fhuE, fpvA, fptA; outer-membrane receptor for ferric coprogen and ferric-rhodotorulic acid
forehead	Proteobacteria	0.0015	0.4455	K01190: lacZ; beta-galactosidase [EC:3.2.1.23]
forehead	Proteobacteria	0.0015	0.4455	K01630: garL; 2-dehydro-3-deoxyglucarate aldolase [EC:4.1.2.20]
forehead	Proteobacteria	0.0015	0.4455	K16568: exoZ; exopolysaccharide production protein ExoZ
forehead	Proteobacteria	0.0015	0.4455	K06903: K06903; uncharacterized protein
forehead	Proteobacteria	0.0016	0.4455	K01235: aguA; alpha-glucuronidase [EC:3.2.1.139]
forehead	Proteobacteria	0.0019	0.4455	K07266: kpsC, lipA; capsular polysaccharide export protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Proteobacteria	0.0021	0.4455	K07326: fhaC; hemolysin activation/secretion protein
forehead	Proteobacteria	0.0021	0.4455	K07741: antB; anti-repressor protein
forehead	Proteobacteria	0.0023	0.4455	K00387: SUOX; sulfite oxidase [EC:1.8.3.1]
forehead	Proteobacteria	0.0023	0.4455	K10039: peb1A, glnH; aspartate/glutamate/glutamine transport system substrate-binding protein
forehead	Proteobacteria	0.0023	0.4455	K11013: cdtA; cytolethal distending toxin subunit A
forehead	Proteobacteria	0.0023	0.4455	K11015: cdtC; cytolethal distending toxin subunit C
forehead	Proteobacteria	0.0023	0.4455	K12247: cstII, cstIII; alpha-2,3 sialyltransferase [EC:2.4.99.-]
forehead	Proteobacteria	0.0023	0.4455	K18197: yesW; rhamnogalacturonan endolyase [EC:4.2.2.23]
forehead	Proteobacteria	0.0023	0.4455	K18833: cmr; MFS transporter, DHA3 family, multidrug efflux protein
forehead	Proteobacteria	0.0023	0.4455	K19239: sapZ; predicted membrane protein
forehead	Proteobacteria	0.0025	0.4881	K07075: K07075; uncharacterized protein
forehead	Proteobacteria	0.0027	0.4919	K18336: K18336; 2,4-diketo-3-deoxy-L-fuconate hydrolase [EC:3.7.1.-]
forehead	Proteobacteria	0.0028	0.4919	K05341: E2.4.1.4; amylosucrase [EC:2.4.1.4]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Proteobacteria	0.0028	0.4919	K09935: K09935; uncharacterized protein
forehead	Proteobacteria	0.003	0.4961	K09688: ABC-2.CPSE.P; capsular polysaccharide transport system permease protein
forehead	Proteobacteria	0.003	0.4961	K18118: aarC, cat1; succinyl-CoA:acetate CoA-transferase [EC:2.8.3.18]
forehead	Proteobacteria	0.0034	0.5514	K06992: K06992; uncharacterized protein
forehead	Proteobacteria	0.0036	0.5529	K01181: E3.2.1.8, xynA; endo-1,4-beta-xylanase [EC:3.2.1.8]
forehead	Proteobacteria	0.0036	0.5529	K01209: abfA; alpha-N-arabinofuranosidase [EC:3.2.1.55]
forehead	Proteobacteria	0.0036	0.5529	K14623: dinD; DNA-damage-inducible protein D
forehead	Proteobacteria	0.0039	0.5673	K02168: betT, betS; choline/glycine/proline betaine transport protein
forehead	Proteobacteria	0.0039	0.5673	K09689: ABC-2.CPSE.A; capsular polysaccharide transport system ATP-binding protein [EC:3.6.3.38]
forehead	Proteobacteria	0.0042	0.5673	K08981: K08981; putative membrane protein
forehead	Proteobacteria	0.0042	0.5673	K14335: pimC; alpha-1,6-mannosyltransferase [EC:2.4.1.-]
forehead	Proteobacteria	0.0045	0.5673	K08363: merT; mercuric ion transport protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

forehead	Proteobacteria	0.0045	0.5673	K06907: K06907; uncharacterized protein
forehead	Proteobacteria	0.0046	0.5673	K00983: neuA; N-acetylneuraminate cytidylyl-transferase [EC:2.7.7.43]
forehead	Proteobacteria	0.0046	0.5673	K07810: cusF; Cu(I)/Ag(I) efflux system periplasmic protein CusF
forehead	Proteobacteria	0.0046	0.5673	K09974: K09974; uncharacterized protein
forehead	Proteobacteria	0.0046	0.5673	K18383: K18383; transferuloyl-CoA hydratase / vanillin synthase [EC:4.2.1.101 4.1.2.41]
forehead	Proteobacteria	0.0046	0.5673	K18991: mtrA; AraC family transcriptional regulator, activator of mtrCDE
gut	Actinobacteria	0.0005	0.9386	K08217: mef; MFS transporter, DHA3 family, macrolide efflux protein
gut	Actinobacteria	0.002	0.9386	K02005: ABC.CD.TX; HlyD family secretion protein
gut	Actinobacteria	0.002	0.9386	K18220: tetM, tetO; ribosomal protection tetracycline resistance protein
gut	Actinobacteria	0.002	0.9386	K20490: nisF, spaF, cprA, epiF; lantibiotic transport system ATP-binding protein
gut	Actinobacteria	0.002	0.9386	K20491: nisE, spaE, cprB, epiE; lantibiotic transport system permease protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Actinobacteria	0.002	0.9386	K21011: pelF; polysaccharide biosynthesis protein PelF
gut	Bacteroidetes	0.0001	0.0026	K00176: korD, oorD; 2-oxoglutarate ferredoxin oxidoreductase subunit delta [EC:1.2.7.3]
gut	Bacteroidetes	0.0001	0.0026	K00177: korC, oorC; 2-oxoglutarate ferredoxin oxidoreductase subunit gamma [EC:1.2.7.3]
gut	Bacteroidetes	0.0001	0.0026	K00179: iorA; indolepyruvate ferredoxin oxidoreductase, alpha subunit [EC:1.2.7.8]
gut	Bacteroidetes	0.0001	0.0026	K00180: iorB; indolepyruvate ferredoxin oxidoreductase, beta subunit [EC:1.2.7.8]
gut	Bacteroidetes	0.0001	0.0026	K01163: K01163; uncharacterized protein
gut	Bacteroidetes	0.0001	0.0026	K01572: oadB; oxaloacetate decarboxylase (Na ⁺ extruding), beta subunit [EC:7.2.4.2]
gut	Bacteroidetes	0.0001	0.0026	K01960: pycB; pyruvate carboxylase subunit B [EC:6.4.1.1]
gut	Bacteroidetes	0.0001	0.0026	K02823: pyrDII; dihydroorotate dehydrogenase electron transfer subunit
gut	Bacteroidetes	0.0001	0.0026	K03612: rnfG; Na ⁺ -translocating ferredoxin:NAD ⁺ oxidoreductase subunit G

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0001	0.0026	K03613: rnfE; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit E
gut	Bacteroidetes	0.0001	0.0026	K03614: rnfD; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit D [EC:7.2.1.2]
gut	Bacteroidetes	0.0001	0.0026	K03615: rnfC; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit C [EC:7.2.1.2]
gut	Bacteroidetes	0.0001	0.0026	K03616: rnfB; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit B [EC:7.2.1.2]
gut	Bacteroidetes	0.0001	0.0026	K03617: rnfA; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit A
gut	Bacteroidetes	0.0001	0.0026	K03803: rseC; sigma-E factor negative regulatory protein RseC
gut	Bacteroidetes	0.0001	0.0026	K03839: fldA, nifF, isiB; flavodoxin I
gut	Bacteroidetes	0.0001	0.0026	K09815: znuA; zinc transport system substrate-binding protein
gut	Bacteroidetes	0.0001	0.0026	K09816: znuB; zinc transport system permease protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0001	0.0026	K09817: znuC; zinc transport system ATP-binding protein [EC:3.6.3.-]
gut	Bacteroidetes	0.0004	0.0124	K00174: korA, oorA, oforA; 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase subunit alpha [EC:1.2.7.3 1.2.7.11]
gut	Bacteroidetes	0.0004	0.0124	K00175: korB, oorB, oforB; 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase subunit beta [EC:1.2.7.3 1.2.7.11]
gut	Bacteroidetes	0.0004	0.0124	K00290: LYS1; saccharopine dehydrogenase (NAD+, L-lysine forming) [EC:1.5.1.7]
gut	Bacteroidetes	0.0004	0.0124	K01006: ppdK; pyruvate, orthophosphate dikinase [EC:2.7.9.1]
gut	Bacteroidetes	0.0004	0.0124	K03737: por, nifJ; pyruvate-ferredoxin/ferredoxin oxidoreductase [EC:1.2.7.1 1.2.7.-]
gut	Bacteroidetes	0.0004	0.0124	K06518: cidA; holin-like protein
gut	Bacteroidetes	0.0004	0.0124	K13747: nspC; carboxynorspermidine decarboxylase [EC:4.1.1.96]
gut	Bacteroidetes	0.0004	0.0124	K17828: pyrDI; dihydroorotate dehydrogenase (NAD+) catalytic subunit [EC:1.3.1.14]
gut	Bacteroidetes	0.0007	0.0162	K00226: pyrD; dihydroorotate dehydrogenase (fumarate) [EC:1.3.98.1]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0007	0.0162	K00610: pyrI; aspartate carbamoyltransferase regulatory subunit
gut	Bacteroidetes	0.0007	0.0162	K00625: E2.3.1.8, pta; phosphate acetyltransferase [EC:2.3.1.8]
gut	Bacteroidetes	0.0007	0.0162	K01119: cpdB; 2,3-cyclic-nucleotide 2-phosphodiesterase / 3-nucleotidase [EC:3.1.4.16 3.1.3.6]
gut	Bacteroidetes	0.0007	0.0162	K01151: nfo; deoxyribonuclease IV [EC:3.1.21.2]
gut	Bacteroidetes	0.0007	0.0162	K01843: kamA; lysine 2,3-aminomutase [EC:5.4.3.2]
gut	Bacteroidetes	0.0007	0.0162	K02117: ATPVA, ntpA, atpA; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit A [EC:7.1.2.2 7.2.2.1]
gut	Bacteroidetes	0.0007	0.0162	K02118: ATPVB, ntpB, atpB; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit B
gut	Bacteroidetes	0.0007	0.0162	K02120: ATPVD, ntpD, atpD; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit D
gut	Bacteroidetes	0.0007	0.0162	K02121: ATPVE, ntpE, atpE; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit E

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0007	0.0162	K02123: ATPVI, ntpI, atpI; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit I
gut	Bacteroidetes	0.0007	0.0162	K02124: ATPVK, ntpK, atpK; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit K
gut	Bacteroidetes	0.0007	0.0162	K02190: cbiK; sirohydrochlorin cobaltochelatase [EC:4.99.1.3]
gut	Bacteroidetes	0.0007	0.0162	K03387: ahpF; alkyl hydroperoxide reductase subunit F [EC:1.6.4.-]
gut	Bacteroidetes	0.0007	0.0162	K03594: bfr; bacterioferritin [EC:1.16.3.1]
gut	Bacteroidetes	0.0007	0.0162	K05601: hcp; hydroxylamine reductase [EC:1.7.99.1]
gut	Bacteroidetes	0.0007	0.0162	K06001: trpB; tryptophan synthase beta chain [EC:4.2.1.20]
gut	Bacteroidetes	0.0007	0.0162	K07089: K07089; uncharacterized protein
gut	Bacteroidetes	0.0007	0.0162	K09181: yfiQ; acetyltransferase
gut	Bacteroidetes	0.0007	0.0162	K11527: K11527; two-component system, sensor histidine kinase and response regulator [EC:2.7.13.3]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0007	0.0162	K21556: malR; CRP/FNR family transcriptional regulator, polysaccharide utilization system transcription regulator
gut	Bacteroidetes	0.0015	0.0262	K00346: nqrA; Na ⁺ -transporting NADH:ubiquinone oxidoreductase subunit A [EC:7.2.1.1]
gut	Bacteroidetes	0.0015	0.0262	K00347: nqrB; Na ⁺ -transporting NADH:ubiquinone oxidoreductase subunit B [EC:7.2.1.1]
gut	Bacteroidetes	0.0015	0.0262	K00348: nqrC; Na ⁺ -transporting NADH:ubiquinone oxidoreductase subunit C [EC:7.2.1.1]
gut	Bacteroidetes	0.0015	0.0262	K00349: nqrD; Na ⁺ -transporting NADH:ubiquinone oxidoreductase subunit D [EC:7.2.1.1]
gut	Bacteroidetes	0.0015	0.0262	K00350: nqrE; Na ⁺ -transporting NADH:ubiquinone oxidoreductase subunit E [EC:7.2.1.1]
gut	Bacteroidetes	0.0015	0.0262	K00351: nqrF; Na ⁺ -transporting NADH:ubiquinone oxidoreductase subunit F [EC:7.2.1.1]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0015	0.0262	K04068: nrdG; anaerobic ribonucleoside-triphosphate reductase activating protein [EC:1.97.1.4]
gut	Bacteroidetes	0.0015	0.0262	K07405: E3.2.1.1A; alpha-amylase [EC:3.2.1.1]
gut	Bacteroidetes	0.0015	0.0262	K11105: cvrA, nhaP2; cell volume regulation protein A
gut	Bacteroidetes	0.0015	0.0262	K18691: mltF; membrane-bound lytic murein transglycosylase F [EC:4.2.2.-]
gut	Bacteroidetes	0.0039	0.0629	K01926: rex; redox-sensing transcriptional repressor
gut	Bacteroidetes	0.0039	0.0629	K03152: thiJ; protein deglycase [EC:3.5.1.124]
gut	Bacteroidetes	0.0039	0.0629	K03308: TC.NSS; neurotransmitter:Na ⁺ symporter, NSS family
gut	Bacteroidetes	0.0039	0.0629	K03340: dapdh; diaminopimelate dehydrogenase [EC:1.4.1.16]
gut	Bacteroidetes	0.0039	0.0629	K03402: argR, ahrC; transcriptional regulator of arginine metabolism
gut	Bacteroidetes	0.0039	0.0629	K06920: queC; 7-cyano-7-deazaguanine synthase [EC:6.3.4.20]
gut	Bacteroidetes	0.0049	0.0629	K00895: pfp, PFP; diphosphate-dependent phosphofructokinase [EC:2.7.1.90]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0049	0.0629	K00929: buk; butyrate kinase [EC:2.7.2.7]
gut	Bacteroidetes	0.0049	0.0629	K01580: E4.1.1.15, gadB, gadA, GAD; glutamate decar- boxylase [EC:4.1.1.15]
gut	Bacteroidetes	0.0049	0.0629	K02169: bioC; malonyl- CoA O-methyltransferase [EC:2.1.1.197]
gut	Bacteroidetes	0.0049	0.0629	K03473: pdxB; erythronate- 4-phosphate dehydrogenase [EC:1.1.1.290]
gut	Bacteroidetes	0.0049	0.0629	K04041: fbp3; fructose- 1,6-bisphosphatase III [EC:3.1.3.11]
gut	Bacteroidetes	0.0049	0.0629	K06926: K06926; uncharacter- ized protein
gut	Bacteroidetes	0.0049	0.0629	K07078: K07078; uncharacter- ized protein
gut	Bacteroidetes	0.0049	0.0629	K07079: K07079; uncharacter- ized protein
gut	Bacteroidetes	0.0049	0.0629	K07138: K07138; uncharacter- ized protein
gut	Bacteroidetes	0.0049	0.0629	K08169: yebQ; MFS trans- porter, DHA2 family, mul- tidrug resistance protein
gut	Bacteroidetes	0.0049	0.0629	K09457: queF; 7-cyano- 7-deazaguanine reductase [EC:1.7.1.13]
gut	Bacteroidetes	0.0049	0.0629	K09790: K09790; uncharacter- ized protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Bacteroidetes	0.0049	0.0629	K12960: mtaD; 5-methylthioadenosine/S-adenosylhomocysteine deaminase [EC:3.5.4.31 3.5.4.28]
gut	Bacteroidetes	0.0049	0.0629	K13378: nuoCD; NADH-quinone oxidoreductase subunit C/D [EC:7.1.1.2]
gut	Bacteroidetes	0.0049	0.0629	K15635: apgM; 2,3-bisphosphoglycerate-independent phosphoglycerate mutase [EC:5.4.2.12]
gut	Bacteroidetes	0.0049	0.0629	K16345: xanP; xanthine permease XanP
gut	Bacteroidetes	0.0049	0.0629	K19117: csd1, cas8c; CRISPR-associated protein Csd1
gut	Bacteroidetes	0.0049	0.0629	K19118: csd2, cas7; CRISPR-associated protein Csd2
gut	Bacteroidetes	0.0049	0.0629	K19119: cas5d; CRISPR-associated protein Cas5d
gut	Bacteroidetes	0.0049	0.0629	K19221: cobA, btuR; cob(I)alamin adenosyltransferase [EC:2.5.1.17]
gut	Bacteroidetes	0.0049	0.0629	K20265: gadC; glutamate:GABA antiporter
gut	Bacteroidetes	0.0049	0.0629	K21063: K21063; 5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase [EC:3.1.3.104]
gut	Firmicutes	0	0.0219	K03737: por, nifJ; pyruvate-ferredoxin/flavodoxin oxidoreductase [EC:1.2.7.1 1.2.7.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Firmicutes	0.0001	0.026	K18349: vanRC, vanRE, vanRG; two-component system, OmpR family, response regulator VanR
gut	Firmicutes	0.0001	0.026	K09765: queH; epoxyqueuosine reductase [EC:1.17.99.6]
gut	Firmicutes	0.0001	0.026	K02107: ATPVG, ahaH, atpH; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit G/H
gut	Firmicutes	0.0001	0.026	K03442: mscS; small conductance mechanosensitive channel
gut	Firmicutes	0.0004	0.035	K07576: K07576; metallo-beta-lactamase family protein
gut	Firmicutes	0.0006	0.0421	K07402: xdhC; xanthine dehydrogenase accessory factor
gut	Firmicutes	0.0007	0.0421	K03612: rnfG; Na ⁺ -translocating ferredoxin:NAD ⁺ oxidoreductase subunit G
gut	Firmicutes	0.0007	0.0421	K03614: rnfD; Na ⁺ -translocating ferredoxin:NAD ⁺ oxidoreductase subunit D [EC:7.2.1.2]
gut	Firmicutes	0.0007	0.0421	K03615: rnfC; Na ⁺ -translocating ferredoxin:NAD ⁺ oxidoreductase subunit C [EC:7.2.1.2]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Firmicutes	0.0007	0.0421	K03616: rnfB; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit B [EC:7.2.1.2]
gut	Firmicutes	0.0007	0.0421	K03617: rnfA; Na+-translocating ferredoxin:NAD+ oxidoreductase subunit A
gut	Firmicutes	0.0007	0.0421	K06179: rluC; 23S rRNA pseudouridine955/2504/2580 synthase [EC:5.4.99.24]
gut	Firmicutes	0.0013	0.0504	K00705: malQ; 4-alpha-glucanotransferase [EC:2.4.1.25]
gut	Firmicutes	0.0013	0.0504	K02118: ATPVB, ntpB, atpB; V/A-type H+/Na+-transporting ATPase subunit B
gut	Firmicutes	0.0013	0.0504	K02120: ATPVD, ntpD, atpD; V/A-type H+/Na+-transporting ATPase subunit D
gut	Firmicutes	0.0013	0.0504	K02122: ATPVF, ntpF, atpF; V/A-type H+/Na+-transporting ATPase subunit F
gut	Firmicutes	0.0013	0.0504	K02123: ATPVI, ntpI, atpI; V/A-type H+/Na+-transporting ATPase subunit I

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Firmicutes	0.0013	0.0504	K02124: ATPVK, ntpK, atpK; V/A-type H ⁺ /Na ⁺ -transporting ATPase subunit K
gut	Firmicutes	0.0013	0.0504	K03613: rnfE; Na ⁺ -translocating ferredoxin:NAD ⁺ oxidoreductase subunit E
gut	Firmicutes	0.0017	0.0621	K00087: ygeS, xdhA; xanthine dehydrogenase molybdenum-binding subunit [EC:1.17.1.4]
gut	Firmicutes	0.0017	0.0621	K07138: K07138; uncharacterized protein
gut	Firmicutes	0.0017	0.0621	K12527: ygfK; putative selenate reductase [EC:1.97.1.9]
gut	Firmicutes	0.0017	0.0621	K13479: ygeT, xdhB; xanthine dehydrogenase FAD-binding subunit [EC:1.17.1.4]
gut	Firmicutes	0.0022	0.0621	K00610: pyrI; aspartate carbamoyltransferase regulatory subunit
gut	Firmicutes	0.0022	0.0621	K01163: K01163; uncharacterized protein
gut	Firmicutes	0.0022	0.0621	K14170: pheA; chorismate mutase / prephenate dehydratase [EC:5.4.99.5 4.2.1.51]
gut	Firmicutes	0.0022	0.0621	K18350: vanSC, vanSE, vanSG; two-component system, OmpR family, sensor histidine kinase VanS [EC:2.7.13.3]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

gut	Firmicutes	0.0024	0.0621	K05808: yhbH; putative sigma-54 modulation protein
gut	Firmicutes	0.0024	0.0621	K07460: yraN; putative endonuclease
gut	Firmicutes	0.0027	0.0621	K01008: selD, SEPHS; selenide, water dikinase [EC:2.7.9.3]
gut	Firmicutes	0.0027	0.0621	K01267: DNPEP; aspartyl aminopeptidase [EC:3.4.11.21]
gut	Firmicutes	0.0027	0.0621	K01950: E6.3.5.1, NADSYN1, QNS1, nadE; NAD ⁺ synthase (glutamine-hydrolysing) [EC:6.3.5.1]
gut	Firmicutes	0.0027	0.0621	K05919: dfx; superoxide reductase [EC:1.15.1.2]
gut	Firmicutes	0.0027	0.0621	K14761: ybcJ; ribosome-associated protein
gut	Firmicutes	0.0027	0.0621	K21071: pfk, pfp; ATP-dependent phosphofructokinase / diphosphate-dependent phosphofructokinase [EC:2.7.1.11 2.7.1.90]
gut	Firmicutes	0.0027	0.0621	K01091: gph; phosphoglycolate phosphatase [EC:3.1.3.18]
gut	Firmicutes	0.0039	0.0825	K01270: pepD; dipeptidase D [EC:3.4.13.-]
gut	Firmicutes	0.0039	0.0825	K01914: asnA; aspartate-ammonia ligase [EC:6.3.1.1]
palm	Actinobacteria	0	0.0143	K11065: tpx; thiol peroxidase, atypical 2-Cys peroxiredoxin [EC:1.11.1.15]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Actinobacteria	0	0.0458	K19131: csb1; CRISPR-associated protein Csb1
palm	Actinobacteria	0	0.0458	K19132: csb2; CRISPR-associated protein Csb2
palm	Actinobacteria	0.0001	0.0616	K04561: norB; nitric oxide reductase subunit B [EC:1.7.2.5]
palm	Actinobacteria	0.0002	0.097	K01623: ALDO; fructose-bisphosphate aldolase, class I [EC:4.1.2.13]
palm	Actinobacteria	0.0002	0.097	K11733: lysP; lysine-specific permease
palm	Actinobacteria	0.0002	0.097	K01575: alsD, budA, aldC; acetolactate decarboxylase [EC:4.1.1.5]
palm	Actinobacteria	0.0002	0.097	K02445: glpT; MFS transporter, OPA family, glycerol-3-phosphate transporter
palm	Actinobacteria	0.0004	0.1259	K01989: ABC.X4.S; putative ABC transport system substrate-binding protein
palm	Actinobacteria	0.0004	0.1259	K05832: ABC.X4.P; putative ABC transport system permease protein
palm	Actinobacteria	0.0004	0.1259	K05833: ABC.X4.A; putative ABC transport system ATP-binding protein
palm	Actinobacteria	0.0004	0.1259	K08996: yagU; putative membrane protein
palm	Actinobacteria	0.0004	0.1303	K14645: K14645; serine protease [EC:3.4.21.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Actinobacteria	0.0006	0.1597	K19158: yoeB; toxin YoeB [EC:3.1.-.-]
palm	Actinobacteria	0.0006	0.1597	K03304: tehA; tellurite resistance protein
palm	Actinobacteria	0.0006	0.1597	K03319: TC.DASS; divalent anion:Na ⁺ symporter, DASS family
palm	Actinobacteria	0.001	0.248	K09125: K09125; uncharacterized protein
palm	Actinobacteria	0.0015	0.3044	K03187: ureE; urease accessory protein
palm	Actinobacteria	0.0015	0.3044	K03311: TC.LIVCS; branched-chain amino acid:cation transporter, LIVCS family
palm	Actinobacteria	0.0016	0.3044	K10254: ohyA, sph; oleate hydratase [EC:4.2.1.53]
palm	Actinobacteria	0.0017	0.3044	K00773: tgt, QTRT1; queuine tRNA-ribosyltransferase [EC:2.4.2.29]
palm	Actinobacteria	0.0017	0.3044	K22552: mmcO; multicopper oxidase [EC:1.16.3.1]
palm	Actinobacteria	0.0017	0.3044	K03488: licT, bglG; beta-glucoside operon transcriptional antiterminator
palm	Actinobacteria	0.0026	0.3998	K00368: nirK; nitrite reductase (NO-forming) [EC:1.7.2.1]
palm	Actinobacteria	0.0026	0.3998	K07452: mcrB; 5-methylcytosine-specific restriction enzyme B [EC:3.1.21.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Actinobacteria	0.0029	0.3998	K00520: merA; mercuric reductase [EC:1.16.1.1]
palm	Actinobacteria	0.0032	0.3998	K00150: gap2; glyceraldehyde-3-phosphate dehydrogenase (NAD(P)) [EC:1.2.1.59]
palm	Actinobacteria	0.0032	0.3998	K01226: treC; trehalose-6-phosphate hydrolase [EC:3.2.1.93]
palm	Actinobacteria	0.0032	0.3998	K02795: PTS-Man-EIIC, manY; PTS system, mannose-specific IIC component
palm	Actinobacteria	0.0032	0.3998	K02796: PTS-Man-EIID, manZ; PTS system, mannose-specific IID component
palm	Actinobacteria	0.0032	0.3998	K03779: ttdA; L(+)-tartrate dehydratase alpha subunit [EC:4.2.1.32]
palm	Actinobacteria	0.0032	0.3998	K03780: ttdB; L(+)-tartrate dehydratase beta subunit [EC:4.2.1.32]
palm	Actinobacteria	0.0032	0.3998	K10215: ethA; monooxygenase [EC:1.14.13.-]
palm	Actinobacteria	0.0032	0.3998	K19133: csb3; CRISPR-associated protein Csb3
palm	Actinobacteria	0.0042	0.4364	K03310: TC.AGCS; alanine or glycine:cation symporter, AGCS family
palm	Actinobacteria	0.0042	0.4364	K22468: ppk2; polyphosphate kinase [EC:2.7.4.1]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Actinobacteria	0.0046	0.4364	K01223: E3.2.1.86B, bglA; 6-phospho-beta-glucosidase [EC:3.2.1.86]
palm	Actinobacteria	0.0046	0.4364	K02755: PTS-Bgl-EIIA, bglF, bglP; PTS system, beta-glucoside-specific IIA component [EC:2.7.1.-]
palm	Actinobacteria	0.0046	0.4364	K02756: PTS-Bgl-EIIB, bglF, bglP; PTS system, beta-glucoside-specific IIB component [EC:2.7.1.-]
palm	Actinobacteria	0.0046	0.4364	K02757: PTS-Bgl-EIIC, bglF, bglP; PTS system, beta-glucoside-specific IIC component
palm	Actinobacteria	0.0046	0.4364	K03475: PTS-Ula-EIIC, ulaA, sgaT; PTS system, ascorbate- specific IIC component
palm	Actinobacteria	0.0046	0.4364	K21562: flp; CRP/FNR fam- ily transcriptional regulator, anaerobic regulatory protein
palm	Actinobacteria	0.0046	0.4364	K21739: rclA; probable pyri- dine nucleotide-disulfide oxi- doreductase
palm	Actinobacteria	0.0048	0.4364	K01534: zntA; Cd ²⁺ /Zn ²⁺ - exporting ATPase [EC:3.6.3.3 3.6.3.5]
palm	Actinobacteria	0.0048	0.4364	K07112: K07112; uncharacter- ized protein
palm	Actinobacteria	0.0048	0.4364	K11737: cycA; D-serine/D- alanine/glycine transporter

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Bacteroidetes	0	0.0092	K01834: PGAM, gpmA; 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase [EC:5.4.2.11]
palm	Bacteroidetes	0.0001	0.0801	K02230: cobN; cobal-tochelata se CobN [EC:6.6.1.2]
palm	Bacteroidetes	0.0001	0.0801	K01744: aspA; aspartate ammonia-lyase [EC:4.3.1.1]
palm	Bacteroidetes	0.0002	0.0896	K04720: cobD; threonine-phosphate decarboxylase [EC:4.1.1.81]
palm	Bacteroidetes	0.0003	0.0896	K01623: ALDO; fructose-bisphosphate aldolase, class I [EC:4.1.2.13]
palm	Bacteroidetes	0.0003	0.0896	K05791: terZ; tellurium resistance protein TerZ
palm	Bacteroidetes	0.0003	0.0896	K05795: terD; tellurium resistance protein TerD
palm	Bacteroidetes	0.0003	0.0896	K06934: K06934; uncharacterized protein
palm	Bacteroidetes	0.0003	0.0896	K07387: K07387; metalloprotease [EC:3.4.24.-]
palm	Bacteroidetes	0.0003	0.0896	K09155: K09155; uncharacterized protein
palm	Bacteroidetes	0.0004	0.0896	K11537: xapB; MFS transporter, NHS family, xanthosine permease
palm	Bacteroidetes	0.0006	0.1024	K01728: pel; pectate lyase [EC:4.2.2.2]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Bacteroidetes	0.0006	0.1024	K03761: kgtP; MFS transporter, MHS family, alpha-ketoglutarate permease
palm	Bacteroidetes	0.0006	0.1024	K07271: licD; lipopolysaccharide cholinephosphotransferase [EC:2.7.8.-]
palm	Bacteroidetes	0.0006	0.1024	K13771: nsrR; Rrf2 family transcriptional regulator, nitric oxide-sensitive transcriptional repressor
palm	Bacteroidetes	0.0008	0.1069	K03743: pncC; nicotinamide-nucleotide amidase [EC:3.5.1.42]
palm	Bacteroidetes	0.0008	0.1069	K04095: fic; cell filamentation protein
palm	Bacteroidetes	0.0008	0.1069	K20459: nukF, mcdF, sboF; lantibiotic transport system ATP-binding protein
palm	Bacteroidetes	0.0009	0.1069	K00077: panE, apbA; 2-dehydropantoate 2-reductase [EC:1.1.1.169]
palm	Bacteroidetes	0.0009	0.1069	K00661: maa; maltose O-acetyltransferase [EC:2.3.1.79]
palm	Bacteroidetes	0.0009	0.1069	K01759: GLO1, gloA; lactoyl-glutathione lyase [EC:4.4.1.5]
palm	Bacteroidetes	0.0009	0.1069	K20534: gtrB, csbB; polyisoprenyl-phosphate glycosyltransferase [EC:2.4.-.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Bacteroidetes	0.0013	0.1452	K02233: E2.7.8.26, cobS, cobV; adenosylcobinamide-GDP ribazoletransferase [EC:2.7.8.26]
palm	Bacteroidetes	0.0013	0.1452	K12308: bgaB, lacA; beta-galactosidase [EC:3.2.1.23]
palm	Bacteroidetes	0.0014	0.152	K00763: pncB, NAPRT1; nicotinate phosphoribosyltransferase [EC:6.3.4.21]
palm	Bacteroidetes	0.0017	0.1545	K01193: INV, sacA; beta-fructofuranosidase [EC:3.2.1.26]
palm	Bacteroidetes	0.0017	0.1545	K02227: cbiB, cobD; adenosylcobinamide-phosphate synthase [EC:6.3.1.10]
palm	Bacteroidetes	0.0017	0.1545	K02232: cobQ, cbiP; adenosylcobyrinic acid synthase [EC:6.3.5.10]
palm	Bacteroidetes	0.0017	0.1545	K07085: K07085; putative transport protein
palm	Bacteroidetes	0.0022	0.1591	K00768: E2.4.2.21, cobU, cobT; nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase [EC:2.4.2.21]
palm	Bacteroidetes	0.0022	0.1591	K01156: res; type III restriction enzyme [EC:3.1.21.5]
palm	Bacteroidetes	0.0022	0.1591	K02226: cobC, phpB; alpha-ribazole phosphatase [EC:3.1.3.73]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Bacteroidetes	0.0022	0.1591	K07075: K07075; uncharacterized protein
palm	Bacteroidetes	0.0022	0.1591	K07504: K07504; predicted type IV restriction endonuclease
palm	Bacteroidetes	0.0026	0.1591	K05845: opuC; osmoprotectant transport system substrate-binding protein
palm	Bacteroidetes	0.0026	0.1591	K05846: opuBD; osmoprotectant transport system permease protein
palm	Bacteroidetes	0.0026	0.1591	K09474: phoN; acid phosphatase (class A) [EC:3.1.3.2]
palm	Bacteroidetes	0.0026	0.1591	K15924: xynC; glucuronoxylan endo-1,4-beta-xylanase [EC:3.2.1.136]
palm	Bacteroidetes	0.0026	0.1591	K00595: cobL; precorrin-6Y C5,15-methyltransferase (decarboxylating) [EC:2.1.1.132]
palm	Bacteroidetes	0.0026	0.1591	K01155: E3.1.21.4; type II restriction enzyme [EC:3.1.21.4]
palm	Bacteroidetes	0.0026	0.1591	K02858: ribB, RIB3; 3,4-dihydroxy 2-butanone 4-phosphate synthase [EC:4.1.99.12]
palm	Bacteroidetes	0.0026	0.1591	K03303: lctP; lactate permease
palm	Bacteroidetes	0.0026	0.1591	K07481: K07481; transposase, IS5 family

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Bacteroidetes	0.0026	0.1591	K07791: dcuA; anaerobic C4-dicarboxylate transporter DcuA
palm	Bacteroidetes	0.0026	0.1591	K20074: prpC, phpP; PPM family protein phosphatase [EC:3.1.3.16]
palm	Bacteroidetes	0.0031	0.1733	K01091: gph; phosphoglycolate phosphatase [EC:3.1.3.18]
palm	Bacteroidetes	0.0031	0.1733	K01185: E3.2.1.17; lysozyme [EC:3.2.1.17]
palm	Bacteroidetes	0.0031	0.1733	K02231: cobP, cobU; adenosylcobinamide kinase / adenosylcobinamide-phosphate guanylyltransferase [EC:2.7.1.156 2.7.7.62]
palm	Bacteroidetes	0.0031	0.1733	K12267: msrAB; peptide methionine sulfoxide reductase msrA/msrB [EC:1.8.4.11 1.8.4.12]
palm	Bacteroidetes	0.0037	0.1903	K01738: cysK; cysteine synthase [EC:2.5.1.47]
palm	Bacteroidetes	0.0037	0.1903	K01854: glf; UDP-galactopyranose mutase [EC:5.4.99.9]
palm	Bacteroidetes	0.0037	0.1903	K07316: mod; adenine-specific DNA-methyltransferase [EC:2.1.1.72]
palm	Bacteroidetes	0.0042	0.192	K15035: ssrA, sipB; transfer-messenger RNA

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Bacteroidetes	0.0045	0.192	K02000: proV; glycine betaine/proline transport system ATP-binding protein [EC:3.6.3.32]
palm	Bacteroidetes	0.0045	0.192	K02001: proW; glycine betaine/proline transport system permease protein
palm	Bacteroidetes	0.0045	0.192	K02395: flgJ; flagellar protein FlgJ
palm	Bacteroidetes	0.0045	0.192	K04041: fbp3; fructose-1,6-bisphosphatase III [EC:3.1.3.11]
palm	Bacteroidetes	0.0045	0.192	K13001: wbyK; mannosyl-transferase [EC:2.4.1.-]
palm	Bacteroidetes	0.0045	0.192	K16345: xanP; xanthine permease XanP
palm	Bacteroidetes	0.0045	0.192	K16710: wcaK, amsJ; colanic acid/amylovoran biosynthesis protein
palm	Bacteroidetes	0.0045	0.192	K19117: csd1, cas8c; CRISPR-associated protein Csd1
palm	Bacteroidetes	0.0045	0.192	K19118: csd2, cas7; CRISPR-associated protein Csd2
palm	Bacteroidetes	0.0045	0.192	K19119: cas5d; CRISPR-associated protein Cas5d
palm	Bacteroidetes	0.0045	0.192	K21063: K21063; 5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase [EC:3.1.3.104]
palm	Firmicutes	0.0001	0.0276	K05873: cyaB; adenylate cyclase, class 2 [EC:4.6.1.1]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Firmicutes	0.0002	0.0276	K09952: csn1, cas9; CRISPR-associated endonuclease Csn1 [EC:3.1.-.-]
palm	Firmicutes	0.0004	0.0276	K14205: mprF, fmtC; phosphatidylglycerol lysyltransferase [EC:2.3.2.3]
palm	Firmicutes	0.0005	0.035	K13727: pdc; phenolic acid decarboxylase [EC:4.1.1.-]
palm	Firmicutes	0.0008	0.0441	K07482: K07482; transposase, IS30 family
palm	Firmicutes	0.0008	0.0441	K11104: melB; melibiose permease
palm	Firmicutes	0.0012	0.0618	K09765: queH; epoxyqueuosine reductase [EC:1.17.99.6]
palm	Firmicutes	0.0014	0.0713	K02445: glpT; MFS transporter, OPA family, glycerol-3-phosphate transporter
palm	Firmicutes	0.0014	0.0713	K02530: lacR; DeoR family transcriptional regulator, lactose phosphotransferase system repressor
palm	Firmicutes	0.0029	0.1142	K06993: K06993; ribonuclease H-related protein
palm	Firmicutes	0.0029	0.1142	K13677: dgs, bgsA; 1,2-diacylglycerol-3-alpha-glucose alpha-1,2-glucosyltransferase [EC:2.4.1.208]
palm	Firmicutes	0.0035	0.1246	K00799: GST, gst; glutathione S-transferase [EC:2.5.1.18]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Firmicutes	0.0035	0.1246	K07694: vraR; two-component system, NarL family, vancomycin resistance associated response regulator VraR
palm	Firmicutes	0.0035	0.1246	K12268: asp1; accessory secretory protein Asp1
palm	Firmicutes	0.0035	0.1246	K12269: asp2; accessory secretory protein Asp2
palm	Firmicutes	0.0035	0.1246	K12270: asp3; accessory secretory protein Asp3
palm	Firmicutes	0.0035	0.1246	K16235: mmuP; S-methylmethionine transporter
palm	Firmicutes	0.0035	0.1246	K20484: nisC, spaC, epiC; antibiotic biosynthesis protein
palm	Firmicutes	0.004	0.1382	K01819: E5.3.1.26, lacA, lacB; galactose-6-phosphate isomerase [EC:5.3.1.26]
palm	Firmicutes	0.004	0.1382	K10830: psaB, scaC; manganese/zinc transport system ATP-binding protein [EC:3.6.3.35]
palm	Firmicutes	0.0049	0.1657	K01919: gshA; glutamate-cysteine ligase [EC:6.3.2.2]
palm	Proteobacteria	0	0.0307	K06906: K06906; uncharacterized protein
palm	Proteobacteria	0	0.0338	K06908: K06908; uncharacterized protein
palm	Proteobacteria	0	0.0584	K06905: K06905; uncharacterized protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0	0.0747	K04102: pht5; 4,5-dihydroxyphthalate decarboxylase [EC:4.1.1.55]
palm	Proteobacteria	0.0002	0.2134	K07504: K07504; predicted type IV restriction endonuclease
palm	Proteobacteria	0.0002	0.2134	K09935: K09935; uncharacterized protein
palm	Proteobacteria	0.0003	0.2765	K06903: K06903; uncharacterized protein
palm	Proteobacteria	0.0004	0.2765	K07075: K07075; uncharacterized protein
palm	Proteobacteria	0.0004	0.2765	K08195: pcaK; MFS transporter, AAHS family, 4-hydroxybenzoate transporter
palm	Proteobacteria	0.0005	0.2765	K07118: K07118; uncharacterized protein
palm	Proteobacteria	0.0005	0.2765	K06148: ABCC-BAC; ATP-binding cassette, subfamily C, bacterial
palm	Proteobacteria	0.0006	0.2765	K18118: aarC, cat1; succinyl-CoA:acetate CoA-transferase [EC:2.8.3.18]
palm	Proteobacteria	0.0006	0.2765	K03300: TC.CITMHS; citrate-Mg ²⁺ :H ⁺ or citrate-Ca ²⁺ :H ⁺ symporter, CitMHS family
palm	Proteobacteria	0.0006	0.2765	K07031: hddA; D-glycero-alpha-D-mannoheptose-7-phosphate kinase [EC:2.7.1.168]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0007	0.2765	K00116: mco; malate dehydrogenase (quinone) [EC:1.1.5.4]
palm	Proteobacteria	0.0008	0.2765	K12539: prsD; ATP-binding cassette, subfamily C, bacterial PrsD
palm	Proteobacteria	0.0009	0.2765	K01796: E5.1.99.4, AMACR, mcr; alpha-methylacyl-CoA racemase [EC:5.1.99.4]
palm	Proteobacteria	0.0009	0.2765	K02839: prfH; peptide chain release factor
palm	Proteobacteria	0.0009	0.2765	K06909: xtmB; phage terminase large subunit
palm	Proteobacteria	0.0009	0.2765	K10220: ligJ; 4-oxalomesaconate hydratase [EC:4.2.1.83]
palm	Proteobacteria	0.0009	0.2765	K16568: exoZ; exopolysaccharide production protein ExoZ
palm	Proteobacteria	0.0011	0.2765	K01184: E3.2.1.15; polygalacturonase [EC:3.2.1.15]
palm	Proteobacteria	0.0011	0.2765	K14578: nahAb, nagAb, ndoA, nbzAb, dntAb; naphthalene 1,2-dioxygenase ferredoxin component
palm	Proteobacteria	0.0011	0.2765	K15861: fixK; CRP/FNR family transcriptional regulator, nitrogen fixation regulation protein
palm	Proteobacteria	0.0011	0.2765	K09962: K09962; uncharacterized protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0012	0.2765	K06907: K06907; uncharacterized protein
palm	Proteobacteria	0.0013	0.2765	K00983: neuA; N-acetylneuraminate cytidylyl-transferase [EC:2.7.7.43]
palm	Proteobacteria	0.0013	0.2765	K03458: TC.NCS2; nucleobase:cation symporter-2, NCS2 family
palm	Proteobacteria	0.0013	0.2765	K07318: K07318; adenine-specific DNA-methyltransferase [EC:2.1.1.72]
palm	Proteobacteria	0.0013	0.2765	K13633: ftrA; AraC family transcriptional regulator, transcriptional activator FtrA
palm	Proteobacteria	0.0014	0.2765	K00101: lldD; L-lactate dehydrogenase (cytochrome) [EC:1.1.2.3]
palm	Proteobacteria	0.0014	0.2765	K18148: rtcB; release factor H-coupled RctB family protein
palm	Proteobacteria	0.0015	0.2765	K18829: mvpT, vapB; anti-toxin VapB
palm	Proteobacteria	0.0017	0.2765	K18991: mtrA; AraC family transcriptional regulator, activator of mtrCDE
palm	Proteobacteria	0.0018	0.2765	K01051: E3.1.1.11; pectinesterase [EC:3.1.1.11]
palm	Proteobacteria	0.0018	0.2765	K18068: pht3; phthalate 4,5-dioxygenase [EC:1.14.12.7]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0018	0.2765	K19355: MAN; mannan endo-1,4-beta-mannosidase [EC:3.2.1.78]
palm	Proteobacteria	0.002	0.2765	K03383: atzD; cyanuric acid amidohydrolase [EC:3.5.2.15]
palm	Proteobacteria	0.002	0.2765	K09836: crtW, BKT; beta-carotene/zeaxanthin 4-ketolase [EC:1.14.99.63 1.14.99.64]
palm	Proteobacteria	0.002	0.2765	K21681: bcs1; ribitol-5- phosphate 2-dehydrogenase (NADP+) / D-ribitol-5- phosphate cytidylyltransferase [EC:1.1.1.405 2.7.7.40]
palm	Proteobacteria	0.0022	0.2765	K07265: kpsS, lipB; capsular polysaccharide export protein
palm	Proteobacteria	0.0023	0.2765	K13796: cobZ, tcuA; tricar- ballylate dehydrogenase
palm	Proteobacteria	0.0023	0.2765	K17228: sfmG; dimethyl- sulfone monooxygenase [EC:1.14.14.35]
palm	Proteobacteria	0.0024	0.2765	K09952: csn1, cas9; CRISPR- associated endonuclease Csn1 [EC:3.1.-.-]
palm	Proteobacteria	0.0024	0.2765	K12685: ssp; subtilase-type serine protease [EC:3.4.21.-]
palm	Proteobacteria	0.0024	0.2765	K13015: wbpA; UDP-N- acetyl-D-glucosamine dehy- drogenase [EC:1.1.1.136]
palm	Proteobacteria	0.0024	0.2765	K01295: cpg; glutamate car- boxypeptidase [EC:3.4.17.11]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0027	0.2765	K07326: fhaC; hemolysin activation/secretion protein
palm	Proteobacteria	0.0027	0.2765	K20023: talrD-galrD; L-talarate/galactarate dehydratase [EC:4.2.1.156 4.2.1.42]
palm	Proteobacteria	0.0027	0.2765	K07266: kpsC, lipA; capsular polysaccharide export protein
palm	Proteobacteria	0.0028	0.2765	K09022: ridA, tdcF, RIDA; 2-iminobutanoate/2-iminopropanoate deaminase [EC:3.5.99.10]
palm	Proteobacteria	0.0029	0.2765	K01779: racD; aspartate racemase [EC:5.1.1.13]
palm	Proteobacteria	0.003	0.2765	K07098: K07098; uncharacterized protein
palm	Proteobacteria	0.0031	0.2765	K02471: bacA; vitamin B12/bleomycin/antimicrobial peptide transport system ATP-binding/permease protein
palm	Proteobacteria	0.0031	0.2765	K15256: cmoA; tRNA (cmo5U34)-methyltransferase [EC:2.1.1.-]
palm	Proteobacteria	0.0031	0.2765	K01209: abfA; alpha-N-arabinofuranosidase [EC:3.2.1.55]
palm	Proteobacteria	0.0031	0.2765	K01406: prtC; serralysin [EC:3.4.24.40]
palm	Proteobacteria	0.0031	0.2765	K01630: garL; 2-dehydro-3-deoxyglucarate aldolase [EC:4.1.2.20]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0031	0.2765	K09702: K09702; uncharacterized protein
palm	Proteobacteria	0.0031	0.2765	K18009: budC; meso-butanediol dehydrogenase / (S,S)-butanediol dehydrogenase / diacetyl reductase [EC:1.1.1.- 1.1.1.76 1.1.1.304]
palm	Proteobacteria	0.0032	0.2765	K02429: fucP; MFS transporter, FHS family, L-fucose permease
palm	Proteobacteria	0.0032	0.2765	K07666: qseB; two-component system, OmpR family, response regulator QseB
palm	Proteobacteria	0.0032	0.2765	K11747: kefB; glutathione-regulated potassium-efflux system protein KefB
palm	Proteobacteria	0.0035	0.2765	K01654: E2.5.1.56, neuB; N-acetylneuraminate synthase [EC:2.5.1.56]
palm	Proteobacteria	0.0035	0.2765	K20034: dmdB; 3-(methylthio)propionyl—CoA ligase [EC:6.2.1.44]
palm	Proteobacteria	0.0037	0.2765	K07017: K07017; uncharacterized protein
palm	Proteobacteria	0.0038	0.2765	K02168: betT, betS; choline/glycine/proline betaine transport protein
palm	Proteobacteria	0.0038	0.2765	K09688: ABC-2.CPSE.P; capsular polysaccharide transport system permease protein

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0038	0.2765	K11004: hlyB, cyaB; ATP-binding cassette, subfamily B, bacterial HlyB/CyaB
palm	Proteobacteria	0.0039	0.2765	K14266: prnA, rebH, ktzQ; tryptophan 7-halogenase [EC:1.14.19.9]
palm	Proteobacteria	0.0039	0.2765	K15629: CYP152A; fatty-acid peroxygenase [EC:1.11.2.4]
palm	Proteobacteria	0.0039	0.2765	K15765: tmoF, tbuC, touF; toluene monooxygenase electron transfer component [EC:1.18.1.3]
palm	Proteobacteria	0.0039	0.2765	K16873: hmfH; 5-(hydroxymethyl)furfural/furfural oxidase [EC:1.1.3.47 1.1.3.-]
palm	Proteobacteria	0.0039	0.2765	K18067: pht4; phthalate 4,5-cis-dihydrodiol dehydrogenase [EC:1.3.1.64]
palm	Proteobacteria	0.0039	0.2765	K18336: K18336; 2,4-diketo-3-deoxy-L-fuconate hydrolase [EC:3.7.1.-]
palm	Proteobacteria	0.004	0.2765	K00055: E1.1.1.90; aryl-alcohol dehydrogenase [EC:1.1.1.90]
palm	Proteobacteria	0.004	0.2765	K01181: E3.2.1.8, xynA; endo-1,4-beta-xylanase [EC:3.2.1.8]
palm	Proteobacteria	0.004	0.2765	K03535: gudP; MFS transporter, ACS family, glucarate transporter
palm	Proteobacteria	0.004	0.2765	K07217: K07217; Mn-containing catalase

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.004	0.2765	K12941: abgB; aminobenzoyl-glutamate utilization protein B
palm	Proteobacteria	0.004	0.2765	K19707: rsbQ; sigma-B regulation protein RsbQ
palm	Proteobacteria	0.0041	0.2765	K00387: SUOX; sulfite oxidase [EC:1.8.3.1]
palm	Proteobacteria	0.0041	0.2765	K03279: waaJ, rfaJ; UDP-glucose:(galactosyl)LPS alpha-1,2-glucosyltransferase [EC:2.4.1.58]
palm	Proteobacteria	0.0041	0.2765	K05576: ndhE; NAD(P)H-quinone oxidoreductase subunit 4L [EC:7.1.1.2]
palm	Proteobacteria	0.0041	0.2765	K10039: peb1A, glnH; aspartate/glutamate/glutamine transport system substrate-binding protein
palm	Proteobacteria	0.0041	0.2765	K11013: cdtA; cytolethal distending toxin subunit A
palm	Proteobacteria	0.0041	0.2765	K11014: cdtB; cytolethal distending toxin subunit B
palm	Proteobacteria	0.0041	0.2765	K11015: cdtC; cytolethal distending toxin subunit C
palm	Proteobacteria	0.0041	0.2765	K12247: cstII, cstIII; alpha-2,3 sialyltransferase [EC:2.4.99.-]
palm	Proteobacteria	0.0041	0.2765	K14657: nodD; LysR family transcriptional regulator, nod-box dependent transcriptional activator

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0041	0.2765	K14658: nodA; nodulation protein A [EC:2.3.1.-]
palm	Proteobacteria	0.0041	0.2765	K14659: nodB; chitooligosaccharide deacetylase [EC:3.5.1.-]
palm	Proteobacteria	0.0041	0.2765	K14666: nodC; N-acetylglucosaminyltransferase [EC:2.4.1.-]
palm	Proteobacteria	0.0041	0.2765	K14680: E6.5.1.3; RNA ligase [EC:6.5.1.3]
palm	Proteobacteria	0.0041	0.2765	K14733: limB; limonene 1,2-monooxygenase [EC:1.14.13.107]
palm	Proteobacteria	0.0041	0.2765	K18197: yesW; rhamnogalacturonan endolyase [EC:4.2.2.23]
palm	Proteobacteria	0.0041	0.2765	K18704: tarL; CDP-ribitol ribitolphosphotransferase / teichoic acid ribitol-phosphate polymerase [EC:2.7.8.14 2.7.8.47]
palm	Proteobacteria	0.0041	0.2765	K18833: cmr; MFS transporter, DHA3 family, multidrug efflux protein
palm	Proteobacteria	0.0041	0.2765	K19239: sapZ; predicted membrane protein
palm	Proteobacteria	0.0041	0.2765	K03313: nhaA; Na ⁺ :H ⁺ antiporter, NhaA family
palm	Proteobacteria	0.0043	0.2806	K09788: prpF; 2-methylnicotinate isomerase [EC:5.3.3.-]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

palm	Proteobacteria	0.0043	0.2806	K10254: ohyA, sph; oleate hydratase [EC:4.2.1.53]
palm	Proteobacteria	0.0043	0.2806	K03307: TC.SSS; solute:Na+ symporter, SSS family
palm	Proteobacteria	0.0049	0.304	K09936: TC.BAT2; bacterial/archaeal transporter family-2 protein
palm	Proteobacteria	0.0049	0.304	K16088: TC.FEV.OM1, fhuE, fpvA, fptA; outer-membrane receptor for ferric coprogen and ferric-rhodotorulic acid
tongue	Actinobacteria	0.001	1	K01226: treC; trehalose 6-phosphate hydrolase [EC:3.2.1.93]
tongue	Actinobacteria	0.0041	1	K02793: PTS-Man-EIIA, manX; PTS system, mannose-specific IIA component [EC:2.7.1.191]
tongue	Actinobacteria	0.0041	1	K02794: PTS-Man-EIIB, manX; PTS system, mannose-specific IIB component [EC:2.7.1.191]
tongue	Actinobacteria	0.0041	1	K02795: PTS-Man-EIIC, manY; PTS system, mannose-specific IIC component
tongue	Actinobacteria	0.0041	1	K02796: PTS-Man-EIID, manZ; PTS system, mannose-specific IID component
tongue	Proteobacteria	0.0001	0.7022	K09789: bioG; pimeloyl-[acyl-carrier protein] methyl ester esterase [EC:3.1.1.85]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

tongue	Proteobacteria	0.0014	0.7022	K07084: yuiF; putative amino acid transporter
tongue	Proteobacteria	0.0014	0.7022	K21993: fdhC; formate transporter
tongue	Proteobacteria	0.0028	0.7022	K07488: K07488; transposase
tongue	Proteobacteria	0.0028	0.7022	K07741: antB; anti-repressor protein
tongue	Proteobacteria	0.0048	0.7022	K01347: iga; IgA-specific serine endopeptidase [EC:3.4.21.72]
tongue	Proteobacteria	0.0048	0.7022	K06215: pdxS, pdx1; pyridoxal 5-phosphate synthase pdxS subunit [EC:4.3.3.6]
tongue	Proteobacteria	0.0048	0.7022	K07149: K07149; uncharacterized protein
tongue	Proteobacteria	0.0048	0.7022	K07288: tspA; uncharacterized membrane protein
tongue	Proteobacteria	0.0048	0.7022	K12247: cstII, cstIII; alpha-2,3 sialyltransferase [EC:2.4.99.-]
tongue	Proteobacteria	0.0048	0.7022	K19239: sapZ; predicted membrane protein
tongue	Proteobacteria	0.0049	0.7022	K07792: dcuB; anaerobic C4-dicarboxylate transporter DcuB
tongue	Proteobacteria	0.0049	0.7022	K09157: K09157; uncharacterized protein
tongue	Proteobacteria	0.0049	0.7022	K12267: msrAB; peptide methionine sulfoxide reductase msrA/msrB [EC:1.8.4.11 1.8.4.12]

表 3.S2. 各体部位に General が Drifter よりも持つ確率が高い KO リスト

tongue	Proteobacteria	0.0049	0.7022	K21936: chuW; anaerobilin synthase [EC:2.1.1.342]
--------	----------------	--------	--------	--

参考文献

- [1] P Seksik. Alterations of the dominant faecal bacterial groups in patients with crohn's disease of the colon. *Gut*, 52(2):237–242, 2003.
- [2] Ruth E Ley, Peter J Turnbaugh, Samuel Klein, and Jeffrey I Gordon. Microbial ecology: human gut microbes associated with obesity. *Nature*, 444(7122):1022–1023, December 2006.
- [3] Alexander Statnikov, Alexander V Alekseyenko, Zhiguo Li, Mikael Henaff, Guillermo I Perez-Perez, Martin J Blaser, and Constantin F Aliferis. Microbiomic signatures of psoriasis: feasibility and methodology comparison. *Sci. Rep.*, 3:2620, 2013.
- [4] H H Kong, J Oh, C Deming, S Conlan, E A Grice, M A Beatson, E Nomicos, E C Polley, H D Komarow, P R Murray, M L Turner, J A Segre, and NISC Comparative Sequence Program. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res.*, 22(5):850–859, 2012.
- [5] Noah Craft and Huiying Li. Response to the commentaries on the paper: Propionibacterium acnes strain populations in the human skin microbiome associated with acne. *J. Invest. Dermatol.*, 133(9):2295–2297, September 2013.
- [6] Elizabeth K Costello, Christian L Lauber, Micah Hamady, Noah Fierer, Jeffrey I Gordon, and Rob Knight. Bacterial community variation in human body habitats across space and time. *Science*, 326(5960):1694–1697, December 2009.
- [7] Gilberto E Flores, J Gregory Caporaso, Jessica B Henley, Jai Ram Rideout, Daniel Domogala, John Chase, Jonathan W Leff, Yoshiki Vázquez-Baeza, Antonio Gonzalez, Rob Knight, Robert R Dunn, and Noah Fierer. Temporal variability is a personalized feature of the human microbiome, 2014.
- [8] Noah Fierer, Christian L Lauber, Nick Zhou, Daniel McDonald, Elizabeth K Costello, and Rob Knight. Forensic identification using skin bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.*, 107(14):6477–6481, April 2010.
- [9] Thomas H Clarke, Andres Gomez, Harinder Singh, Karen E Nelson, and Lauren M Brinkac. Integrating the microbiome as a resource in the forensics toolkit. *Forensic Sci. Int. Genet.*, 30:141–147, September 2017.
- [10] Sarah E Schmedes, August E Woerner, and Bruce Budowle. Forensic human identification using skin microbiomes. *Appl. Environ. Microbiol.*, September 2017.

- [11] Sarah E Schmedes, August E Woerner, Nicole M M Novroski, Frank R Wendt, Jonathan L King, Kathryn M Stephens, and Bruce Budowle. Targeted sequencing of clade-specific markers from skin microbiomes for forensic human identification. *Forensic Sci. Int. Genet.*, 32:50–61, January 2018.
- [12] Eric A Franzosa, Katherine Huang, James F Meadow, Dirk Gevers, Katherine P Lemon, Brendan J M Bohannon, and Curtis Huttenhower. Identifying personal microbiomes using metagenomic codes. *Proc. Natl. Acad. Sci. U. S. A.*, 112(22):E2930–8, June 2015.
- [13] Elizabeth A Grice, Heidi H Kong, Sean Conlan, Clayton B Deming, Joie Davis, Alice C Young, NISC Comparative Sequencing Program, Gerard G Bouffard, Robert W Blakesley, Patrick R Murray, Eric D Green, Maria L Turner, and Julia A Segre. Topographical and temporal diversity of the human skin microbiome. *Science*, 324(5931):1190–1192, May 2009.
- [14] Hikaru Watanabe, Issei Nakamura, Sayaka Mizutani, Yumiko Kurokawa, Hiroshi Mori, Ken Kurokawa, and Takuji Yamada. Minor taxa in human skin microbiome contribute to the personal identification. *PLoS One*, 13(7):e0199947, 2018.
- [15] Luke R Thompson, Jon G Sanders, Daniel McDonald, Amnon Amir, Joshua Ladau, Kenneth J Locey, Robert J Prill, Anupriya Tripathi, Sean M Gibbons, Gail Ackermann, Jose A Navas-Molina, Stefan Janssen, Evguenia Kopylova, Yoshiki Vázquez-Baeza, Antonio González, James T Morton, Siavash Mirarab, Zhenjiang Zech Xu, Lingjing Jiang, Mohamed F Haroon, Jad Kanbar, Qiyun Zhu, Se Jin Song, Tomasz Kosciolk, Nicholas A Bokulich, Joshua Lefler, Colin J Brislawn, Gregory Humphrey, Sarah M Owens, Jarrad Hampton-Marcell, Donna Berg-Lyons, Valerie McKenzie, Noah Fierer, Jed A Fuhrman, Aaron Clauset, Rick L Stevens, Ashley Shade, Katherine S Pollard, Kelly D Goodwin, Janet K Jansson, Jack A Gilbert, Rob Knight, and Earth Microbiome Project Consortium. A communal catalogue reveals earth’s multiscale microbial diversity. *Nature*, 551(7681):457–463, November 2017.
- [16] Carola Söhngen, Boyke Bunk, Adam Podstawka, Dorothea Gleim, and Jörg Overmann. BacDive—the bacterial diversity metadatabase. *Nucleic Acids Res.*, 42(Database issue):D592–9, January 2014.
- [17] Albert Barberán, Hildamarie Caceres Velazquez, Stuart Jones, and Noah Fierer. Hiding in plain sight: Mining bacterial species records for phenotypic trait information. *mSphere*, 2(4), July 2017.
- [18] I-Min A Chen, Victor M Markowitz, Ken Chu, Krishna Palaniappan, Ernest Szeto, Manoj Pillay, Anna Ratner, Jinghua Huang, Evan Andersen, Marcel Huntemann, Neha Varghese, Michalis Hadjithomas, Kristin Tennesen, Torben Nielsen, Natalia N Ivanova, and Nikos C Kyrpides. IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.*, 45(D1):D507–D516, January 2017.

- [19] Ching-Chia Yang and Wataru Iwasaki. MetaMetaDB: a database and analytic system for investigating microbial habitability. *PLoS One*, 9(1):e87126, January 2014.
- [20] Alex L Mitchell, Maxim Scheremetjew, Hubert Denise, Simon Potter, Aleksandra Tarkowska, Matloob Qureshi, Gustavo A Salazar, Sebastien Pesseat, Miguel A Boland, Fiona M I Hunter, Petra Ten Hoopen, Blaise Alako, Clara Amid, Darren J Wilkinson, Thomas P Curtis, Guy Cochrane, and Robert D Finn. EBI metagenomics in 2017: enriching the analysis of microbial communities, from sequence reads to assemblies. *Nucleic Acids Res.*, 46(D1):D726–D735, January 2018.
- [21] Koichi Higashi, Shinya Suzuki, Shin Kurosawa, Hiroshi Mori, and Ken Kurokawa. Latent environment allocation of microbial community data. *PLoS Comput. Biol.*, 14(6):e1006143, June 2018.
- [22] Dan Knights, Justin Kuczynski, Emily S Charlson, Jesse Zaneveld, Michael C Mozer, Ronald G Collman, Frederic D Bushman, Rob Knight, and Scott T Kelley. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods*, 8(9):761–763, July 2011.
- [23] Simon Lax, Daniel P Smith, Jarrad Hampton-Marcell, Sarah M Owens, Kim M Handley, Nicole M Scott, Sean M Gibbons, Peter Larsen, Benjamin D Shogan, Sophie Weiss, Jessica L Metcalf, Luke K Ursell, Yoshiki Vázquez-Baeza, Will Van Treuren, Nur A Hasan, Molly K Gibson, Rita Colwell, Gautam Dantas, Rob Knight, and Jack A Gilbert. Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*, 345(6200):1048–1052, August 2014.
- [24] Se Jin Song, Christian Lauber, Elizabeth K Costello, Catherine A Lozupone, Gregory Humphrey, Donna Berg-Lyons, J Gregory Caporaso, Dan Knights, Jose C Clemente, Sara Nakielnny, Jeffrey I Gordon, Noah Fierer, and Rob Knight. Cohabiting family members share microbiota with one another and with their dogs. *Elife*, 2:e00458, April 2013.
- [25] Embriette R Hyde, Jose A Navas-Molina, Se Jin Song, Jordan G Kueneman, Gail Ackermann, Cesar Cardona, Gregory Humphrey, Don Boyer, Tom Weaver, Joseph R Mendelson, 3rd, Valerie J McKenzie, Jack A Gilbert, and Rob Knight. The oral and skin microbiomes of captive komodo dragons are significantly shared with their habitat. *mSystems*, 1(4), July 2016.
- [26] Andrew H Loudon, Douglas C Woodhams, Laura Wegener Parfrey, Holly Archer, Rob Knight, Valerie McKenzie, and Reid N Harris. Microbial community dynamics and effect of environmental microbial reservoirs on red-backed salamanders (*plethodon cinereus*). *ISME J.*, 8(4):830–840, April 2014.
- [27] Julia Oh, NISC Comparative Sequencing Program, Allyson L Byrd, Clay Deming, Sean Conlan, Heidi H Kong, and Julia A Segre. Biogeography and individuality shape function in the human skin metagenome. *Nature*, 514(7520):59–64, 2014.

- [28] Julia Oh, Allyson L Byrd, Morgan Park, Heidi H Kong, and Julia A Segre. Temporal stability of the human skin microbiome. *Cell*, 165(4):854–866, 2016.
- [29] Tommi Vatanen, Aleksandar D Kostic, Eva d’Hennezel, Heli Siljander, Eric A Franzosa, Moran Yassour, Raivo Kolde, Hera Vlamakis, Timothy D Arthur, Anu-Maaria Hämäläinen, Aleksandr Peet, Vallo Tillmann, Raivo Uibo, Sergei Mokurov, Natalya Dorshakova, Jorma Ilonen, Suvi M Virtanen, Susanne J Szabo, Jeffrey A Porter, Harri Lähdesmäki, Curtis Huttenhower, Dirk Gevers, Thomas W Cullen, Mikael Knip, DIABIMMUNE Study Group, and Ramnik J Xavier. Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell*, 165(6):1551, June 2016.
- [30] Daniel McDonald, Embriette Hyde, Justine W Debelius, James T Morton, Antonio Gonzalez, Gail Ackermann, Alexander A Aksenov, Bahar Behsaz, Caitriona Brennan, Yingfeng Chen, Lindsay DeRight Goldasich, Pieter C Dorrestein, Robert R Dunn, Ashkaan K Fahimipour, James Gaffney, Jack A Gilbert, Grant Gogul, Jessica L Green, Philip Hugenholtz, Greg Humphrey, Curtis Huttenhower, Matthew A Jackson, Stefan Janssen, Dilip V Jeste, Lingjing Jiang, Scott T Kelley, Dan Knights, Tomasz Kosciolk, Joshua Ladau, Jeff Leach, Clarisse Marotz, Dmitry Meleshko, Alexey V Melnik, Jessica L Metcalf, Hosein Mohimani, Emmanuel Montassier, Jose Navas-Molina, Tanya T Nguyen, Shyamal Peddada, Pavel Pevzner, Katherine S Pollard, Gholamali Rahnavard, Adam Robbins-Pianka, Naseer Sangwan, Joshua Shorenstein, Larry Smarr, Se Jin Song, Timothy Spector, Austin D Swafford, Varykina G Thackray, Luke R Thompson, Anupriya Tripathi, Yoshiki Vázquez-Baeza, Alison Vrbanc, Paul Wischmeyer, Elaine Wolfe, Qiyun Zhu, American Gut Consortium, and Rob Knight. American gut: an open platform for citizen science microbiome research. *mSystems*, 3(3), May 2018.
- [31] Manimozhiyan Arumugam, Jeroen Raes, Eric Pelletier, Denis Le Paslier, Takuji Yamada, Daniel R Mende, Gabriel R Fernandes, Julien Tap, Thomas Bruls, Jean-Michel Batto, Marcelo Bertalan, Natalia Borruel, Francesc Casellas, Leyden Fernandez, Laurent Gautier, Torben Hansen, Masahira Hattori, Tetsuya Hayashi, Michiel Kleerebezem, Ken Kurokawa, Marion Leclerc, Florence Levenez, Chaysavanh Manichanh, H Bjørn Nielsen, Trine Nielsen, Nicolas Pons, Julie Poulain, Junjie Qin, Thomas Sicheritz-Ponten, Sebastian Tims, David Torrents, Edgardo Ugarte, Erwin G Zoetendal, Jun Wang, Francisco Guarner, Oluf Pedersen, Willem M de Vos, Søren Brunak, Joel Doré, MetaHIT Consortium, María Antolín, François Artiguenave, Hervé M Blottiere, Mathieu Almeida, Christian Brechot, Carlos Cara, Christian Chervaux, Antonella Cultrone, Christine Delorme, Gérard Denariáz, Rozenn Dervyn, Konrad U Foerstner, Carsten Friss, Maarten van de Guchte, Eric Guedon, Florence Haimet, Wolfgang Huber, Johan van Hylckama-Vlieg, Alexandre Jamet, Catherine Juste, Ghalia Kaci, Jan Knol, Omar Lakhdari, Severine Layec, Karine Le Roux, Emmanuelle Maguin, Alexandre Mérieux, Raquel Melo Minardi, Christine M’rini, Jean Muller, Raish Oozeer, Julian Parkhill, Pierre Renault, Maria Rescigno, Nicolas Sanchez, Shinichi

- Sunagawa, Antonio Torrejon, Keith Turner, Gaetana Vandemeulebrouck, Encarna Varela, Yohanan Winogradsky, Georg Zeller, Jean Weissenbach, S Dusko Ehrlich, and Peer Bork. Enterotypes of the human gut microbiome. *Nature*, 473(7346):174–180, May 2011.
- [32] Minoru Kanehisa and Susumu Goto. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30, 2000.
- [33] Asaf Levy, Isai Salas Gonzalez, Maximilian Mittelviehhaus, Scott Clingenpeel, Sur Herrera Paredes, Jiamin Miao, Kunru Wang, Giulia Devescovi, Kyra Stillman, Freddy Monteiro, Bryan Rangel Alvarez, Derek S Lundberg, Tse-Yuan Lu, Sarah Lebeis, Zhao Jin, Meredith McDonald, Andrew P Klein, Meghan E Feltcher, Tijana Glavina Rio, Sarah R Grant, Sharon L Doty, Ruth E Ley, Bingyu Zhao, Vittorio Venturi, Dale A Pelletier, Julia A Vorholt, Susannah G Tringe, Tanja Woyke, and Jeffery L Dangl. Genomic features of bacterial adaptation to plants. *Nat. Genet.*, 50(1):138–150, January 2018.
- [34] Anthony M Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30(15):2114–2120, 2014.
- [35] Robert Schmieder, Yan Wei Lim, Forest Rohwer, and Robert Edwards. TagCleaner: Identification and removal of tag sequences from genomic and metagenomic datasets. *BMC Bioinformatics*, 11:341, June 2010.
- [36] Robert C Edgar. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19):2460–2461, October 2010.
- [37] Robert C Edgar, Brian J Haas, Jose C Clemente, Christopher Quince, and Rob Knight. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16):2194–2200, 2011.
- [38] S F Altschul, W Gish, W Miller, E W Myers, and D J Lipman. Basic local alignment search tool. *J. Mol. Biol.*, 215(3):403–410, October 1990.
- [39] Stijn Marinus van Dongen. *Graph Clustering by Flow Simulation*. 2000.
- [40] Pablo Yarza, Michael Richter, Jörg Peplies, Jean Euzéby, Rudolf Amann, Karl-Heinz Schleifer, Wolfgang Ludwig, Frank Oliver Glöckner, and Ramon Rosselló-Móra. The All-Species living tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.*, 31(4):241–250, September 2008.
- [41] Robert C Edgar. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods*, 10(10):996–998, October 2013.
- [42] Ian F Spellerberg and Peter J Fedor. A tribute to claude shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon-Wiener’ index. *Glob. Ecol. Biogeogr.*, 12(3):177–179, 2003.

- [43] Philip Dixon. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.*, 14(6):927, 2003.
- [44] Jun Chen, Kyle Bittinger, Emily S Charlson, Christian Hoffmann, James Lewis, Gary D Wu, Ronald G Collman, Frederic D Bushman, and Hongzhe Li. Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics*, 28(16):2106–2113, 2012.
- [45] Ashley A Ross, Andrew C Doxey, and Josh D Neufeld. The skin microbiome of cohabiting couples. *mSystems*, 2(4), July 2017.
- [46] The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402):207–214, 2012.
- [47] Nakako Shibagaki, Wataru Suda, Cecile Clavaud, Philippe Bastien, Lena Takayasu, Erica Iioka, Rina Kurokawa, Naoko Yamashita, Yasue Hattori, Chie Shindo, Lionel Breton, and Masahira Hattori. Aging-related changes in the diversity of women’s skin microbiomes associated with oral bacteria. *Sci. Rep.*, 7(1), 2017.
- [48] Jacquelyn S Meisel, Geoffrey D Hannigan, Amanda S Tyldsley, Adam J SanMiguel, Brendan P Hodkinson, Qi Zheng, and Elizabeth A Grice. Skin microbiome surveys are strongly influenced by experimental design. *J. Invest. Dermatol.*, 136(5):947–956, May 2016.
- [49] Steven Salzberg. Faculty of 1000 evaluation for reagent and laboratory contamination can critically impact sequence-based microbiome analyses, 2014.
- [50] Geoffrey D Hannigan, Jacquelyn S Meisel, Amanda S Tyldsley, Qi Zheng, Brendan P Hodkinson, Adam J SanMiguel, Samuel Minot, Frederic D Bushman, and Elizabeth A Grice. The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *MBio*, 6(5):e01578–15, October 2015.
- [51] Julia J van Rensburg, Huaiying Lin, Xiang Gao, Evelyn Toh, Kate R Fortney, Sheila Ellinger, Beth Zwickl, Diane M Janowicz, Barry P Katz, David E Nelson, Qunfeng Dong, and Stanley M Spinola. The human skin microbiome associates with the outcome of and is influenced by bacterial infection. *MBio*, 6(5), 2015.
- [52] David Wilkins, Marcus H Y Leung, and Patrick K H Lee. Microbiota fingerprints lose individually identifying features over time. *Microbiome*, 5(1):1, January 2017.
- [53] Souvik Mukherjee, Rupak Mitra, Arindam Maitra, Satyaranjan Gupta, Srikala Kumaran, Amit Chakraborty, and Partha P Majumder. Sebum and hydration levels in specific regions of human face significantly predict the nature and diversity of facial skin microbiome. *Sci. Rep.*, 6(1), 2016.

- [54] Shuta Tomida, Lin Nguyen, Bor-Han Chiu, Jared Liu, Erica Sodergren, George M Weinstein, and Huiying Li. Pan-genome and comparative genome analyses of propionibacterium acnes reveal its genomic diversity in the healthy and diseased human skin microbiome. *MBio*, 4(3):e00003–13, April 2013.
- [55] T Z DeSantis, P Hugenholtz, N Larsen, M Rojas, E L Brodie, K Keller, T Huber, D Dalevi, P Hu, and G L Andersen. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.*, 72(7):5069–5072, July 2006.
- [56] Qiong Wang, George M Garrity, James M Tiedje, and James R Cole. Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.*, 73(16):5261–5267, August 2007.
- [57] Christian Quast, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, Jörg Peplies, and Frank Oliver Glöckner. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.*, 41(Database issue):D590–6, January 2013.
- [58] Hiroshi Mori, Takayuki Maruyama, Masahiro Yano, Takuji Yamada, and Ken Kurokawa. VITCOMIC2: visualization tool for the phylogenetic composition of microbial communities based on 16S rRNA gene amplicons and metagenomic shotgun sequencing. *BMC Syst. Biol.*, 12(Suppl 2):30, March 2018.
- [59] Sira Sriswasdi, Ching-Chia Yang, and Wataru Iwasaki. Generalist species drive microbial dispersion and evolution. *Nat. Commun.*, 8(1):1162, October 2017.
- [60] W Fujibuchi, S Goto, H Migimatsu, I Uchiyama, A Ogiwara, Y Akiyama, and M Kanehisa. DBGET/LinkDB: an integrated database retrieval system. *Pac. Symp. Biocomput.*, pages 683–694, 1998.
- [61] Alex Bateman, Maria Jesus Martin, Claire O’Donovan, Michele Magrane, Emanuele Alpi, Ricardo Antunes, Benoit Bely, Mark Bingley, Carlos Bonilla, Ramona Britto, Borisas Bursteinas, Hema Bye-A-Jee, Andrew Cowley, Alan Da Silva, Maurizio De Giorgi, Tunca Dogan, Francesco Fazzini, Leyla Garcia Castro, Luis Figueira, Penelope Garmiri, George Georghiou, Daniel Gonzalez, Emma Hatton-Ellis, Weizhong Li, Wudong Liu, Rodrigo Lopez, Jie Luo, Yvonne Lussi, Alistair MacDougall, Andrew Nightingale, Barbara Palka, Klemens Pichler, Diego Poggioli, Sangya Pundir, Luis Pureza, Guoying Qi, Alexandre Renaux, Steven Rosanoff, Rabie Saidi, Tony Sawford, Aleksandra Shypitsyna, Elena Speretta, Edward Turner, Nidhi Tyagi, Vladimir Volynkin, Tony Wardell, Kate Warner, Xavier Watkins, Rossana Zaru, Hermann Zellner, Ioannis Xenarios, Lydie Bougueleret, Alan Bridge, Sylvain Poux, Nicole Redaschi, Lucila Aimò, Ghislaine Argoud-Puy, Andrea Auchincloss, Kristian Axelsen, Parit Bansal, Delphine Baratin, Marie-Claude Blatter, Brigitte Boeckmann, Jerven Bolleman, Emmanuel Boutet, Lionel Breuza, Cristina Casal-Casas,

- Edouard de Castro, Elisabeth Coudert, Beatrice Cuche, Mikael Doche, Dolnide Dornevil, Severine Duvaud, Anne Estreicher, Livia Famiglietti, Marc Feuermann, Elisabeth Gasteiger, Sebastien Gehant, Vivienne Gerritsen, Arnaud Gos, Nadine Gruaz-Gumowski, Ursula Hinz, Chantal Hulo, Florence Jungo, Guillaume Keller, Vicente Lara, Philippe Lemercier, Damien Lieberherr, Thierry Lombardot, Xavier Martin, Patrick Masson, Anne Morgat, Teresa Neto, Nevila Nospikel, Salvo Paesano, Ivo Pedruzzi, Sandrine Pilbout, Monica Pozzato, Manuela Pruess, Catherine Rivoire, Bernd Roechert, Michel Schneider, Christian Sigrist, Karin Sonesson, Sylvie Staehli, Andre Stutz, Shyamala Sundaram, Michael Tognolli, Laure Verbregue, Anne-Lise Veuthey, Cathy H Wu, Cecilia N Arighi, Leslie Arminski, Chuming Chen, Yongxing Chen, John S Garavelli, Hongzhan Huang, Kati Laiho, Peter McGarvey, Darren A Natale, Karen Ross, C R Vinayaka, Qinghua Wang, Yuqi Wang, Lai-Su Yeh, and Jian Zhang. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.*, 45(D1):D158–D169, January 2017.
- [62] Kristoffer Forslund, Falk Hildebrand, Trine Nielsen, Gwen Falony, Emmanuelle Le Chatelier, Shinichi Sunagawa, Edi Prifti, Sara Vieira-Silva, Valborg Gudmundsdottir, Helle Krogh Pedersen, Manimozhiyan Arumugam, Karsten Kristiansen, Anita Yvonne Voigt, Henrik Vestergaard, Rajna Hercog, Paul Igor Costea, Jens Roat Kultima, Junhua Li, Torben Jørgensen, Florence Levenez, Joël Dore, MetaHIT consortium, H Bjørn Nielsen, Søren Brunak, Jeroen Raes, Torben Hansen, Jun Wang, S Dusko Ehrlich, Peer Bork, and Oluf Pedersen. Corrigendum: Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature*, 545(7652):116, May 2017.
- [63] Alexandra Zhernakova, Alexander Kurilshikov, Marc Jan Bonder, Ettje F Tigchelaar, Melanie Schirmer, Tommi Vatanen, Zlatan Mijajic, Arnau Vich Vila, Gwen Falony, Sara Vieira-Silva, Jun Wang, Floris Imhann, Eelke Brandsma, Soesma A Jankipersadsing, Marie Joossens, Maria Carmen Cenit, Patrick Deelen, Morris A Swertz, LifeLines cohort study, Rinse K Weersma, Edith J M Feskens, Mihai G Netea, Dirk Gevers, Daisy Jonkers, Lude Franke, Yurii S Aulchenko, Curtis Huttenhower, Jeroen Raes, Marten H Hofker, Ramnik J Xavier, Cisca Wijmenga, and Jingyuan Fu. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science*, 352(6285):565–569, April 2016.
- [64] Noah Fierer, Micah Hamady, Christian L Lauber, and Rob Knight. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc. Natl. Acad. Sci. U. S. A.*, 105(46):17994–17999, November 2008.
- [65] Paul I Costea, Luis Pedro Coelho, Shinichi Sunagawa, Robin Munch, Jaime Huerta - Cepas, Kristoffer Forslund, Falk Hildebrand, Almagul Kushugulova, Georg Zeller, and Peer Bork. Subspecies in the global human gut microbiome. *Mol. Syst. Biol.*, 13(12):960, 2017.

- [66] H Inoue, K Ono, W Masuda, Y Morimoto, T Tanaka, M Yokota, and K Inenaga. Gender difference in unstimulated whole saliva flow rate and salivary gland sizes. *Arch. Oral Biol.*, 51(12):1055–1060, December 2006.
- [67] Simon J Labrie, Julie E Samson, and Sylvain Moineau. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, 8(5):317–327, 2010.
- [68] Sarah A Horst, Timo Jaeger, Luisa A Denkel, Syed Fazle Rouf, Mikael Rhen, and Franz-Christoph Bange. Thiol peroxidase protects salmonella enterica from hydrogen peroxide stress in vitro and facilitates intracellular growth. *J. Bacteriol.*, 192(11):2929–2932, June 2010.
- [69] Hee-Jeon Hong, Matthew I Hutchings, Mark J Buttner, and Biotechnology and Biological Sciences Research Council, UK. Vancomycin resistance VanS/VanR two-component systems. *Adv. Exp. Med. Biol.*, 631:200–213, 2008.
- [70] Marilyn C Roberts. Update on acquired tetracycline resistance genes. *FEMS Microbiol. Lett.*, 245(2):195–203, April 2005.
- [71] MetaCyc factor 420 biosynthesis. <https://biocyc.org/META/NEW-IMAGE?type=PATHWAY&object=PWY-5198>. Accessed: 2019-1-2.
- [72] Patrick Ljm Zeeuwen, Jos Boekhorst, Ellen H van den Bogaard, Heleen D de Koning, Peter M C van de Kerkhof, Delphine M Saulnier, Iris I van Swam, Sacha Aft van Hijum, Michiel Kleerebezem, Joost Schalkwijk, and Harro M Timmerman. Microbiome dynamics of human epidermis following skin barrier disruption. *Genome Biol.*, 13(11):R101, November 2012.
- [73] T Murata, M Kawano, K Igarashi, I Yamato, and Y Kakinuma. Catalytic properties of na(+)-translocating V-ATPase in enterococcus hirae. *Biochim. Biophys. Acta*, 1505(1):75–81, May 2001.
- [74] Juke S Lolkema, Yuriy Chaban, and Egbert J Boekema. Subunit composition, structure, and distribution of bacterial v-type ATPases. *J. Bioenerg. Biomembr.*, 35(4):323–335, August 2003.