Construction of a complete set of *Neisseria meningitidis* defined mutants – the NeMeSys collection – and its use for the phenotypic profiling of the genome of an important human pathogen

Alastair Muir[1], Ishwori Gurung[1], Ana Cehovin[1], Adelme Bazin[2], David Vallenet[2], Vladimir Pelicic[1,*]

[1]MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London, United Kingdom

[2]LABGeM, Génomique Métabolique, CEA, Genoscope, Institut François Jacob, Université d'Evry, Université Paris-Saclay, CNRS, Evry, France

*Corresponding author

E-mail: v.pelicic@imperial.ac.uk

Abstract

One of the most important challenges in biology is to determine the function of millions of genes of unknown function. Even in model bacterial species, there is a sizeable proportion of such genes, which has important theoretical and practical consequences. Here, we constructed a complete collection of defined mutants – named NeMeSys – in the important human pathogen *Neisseria meningitidis*, consisting of individual mutants in 1,584 non-essential genes. This effort identified 391 essential meningococcal genes – highly conserved in other bacteria – leading to a full panorama of the minimal genome in this species, associated with just four underlying basic biological functions: 1) expression of genome information, 2) preservation of genome information, 3) cell membrane structure/function, and 4) cytosolic metabolism. Subsequently, we illustrated the utility of the NeMeSys collection for determining gene function by identifying 1) a novel and conserved family of histidinol-phosphatase, 2) all genes, including three new ones, involved in the biology of type IV pili, a widespread virulence factor, and 3) several conditionally essential genes found in regions of genome plasticity, likely to encode antitoxins and/or immunity proteins. These findings have widespread implications in bacteria. The NeMeSys collection is an invaluable resource paving the way for a global phenotypic landscape in a major human bacterial pathogen.

Introduction

Low-cost massively parallel sequencing methods have transformed biology, with availability of complete genome sequences increasing at breathtaking pace. In bacteria, genome sequences can be determined for hundreds of isolates in parallel in just a matter of days. For bacterial pathogens, genome sequences are often available for thousands of clinical isolates, a bonanza for epidemiology and comparative genomics[1]. This explosion has also led to the identification of millions of new genes, with the important insight that many are genes of unknown function. Even in a model species such as *Escherichia coli* K-12, perhaps the most thoroughly studied biological entity, only 57 % of the 4,182 protein-coding genes have experimentally verified functions[2]. The remaining 43 % have either no predicted function, or a predicted function that is yet to be verified[2]. In less well studied species, this gene partition is dramatically skewed towards the latter class, with an overwhelming majority of genes having predicted but not experimentally verified functions, or no predicted function at all. This indicates that there are still major gaps in our understanding of basic bacterial physiology, including for well-studied metabolic pathways[3]. This has important theoretical consequences for systems and synthetic biology, by hindering holistic understanding of the life-supporting functions and processes allowing bacteria to grow and replicate. It also has practical consequences by slowing down the identification of new means for controlling bacterial pathogens, which are needed more than ever in this era of antimicrobial resistance. Therefore, developing methods for systematically elucidating gene function is a critical biological endeavour.

Molecular genetics, which involves the creation and phenotypic analysis of large collections of mutants, is the most direct path to determining gene function on a genome scale. Such efforts provide a crucial link between phenotype and genotype. In many bacteria, genome-wide mutant collections have been constructed using either targeted mutagenesis or random transposon (Tn) mutagenesis. Because of its

time and cost-effectiveness, Tn mutagenesis is way more widely used[4-6], especially since it has been coupled with massively parallel sequencing of Tn insertion sites (Tn-Seq)[7], which allows simultaneous monitoring of the fitness of pools of hundreds of thousands of mutants[8,9]. However, Tn-Seq is not without drawbacks. 1) Not all intragenic Tn insertions lead to a loss of function. 2) Specific mutants cannot be recovered from pools for further individual study. 3) Libraries of Tn mutant are never complete, no matter how large. In contrast, complete collections of mutants constructed by targeted mutagenesis do not suffer from these limitations, but their construction is far more laborious and expensive. For this reason, such collections are available for only a handful of bacterial species: the model Gram-negative *E. coli* K-12 (Gammaproteobacteria)[10], the model Gram-positive *Bacillus subtilis* (Firmicute)[11], *Acinetobacter baylyi* (Gammaproteobacteria) studied for its metabolic properties and possible biotechnological applications[12], and *Streptococcus sanguinis* (Firmicute) an opportunistic pathogen causing infective endocarditis[13]. The utility of such resources is best exemplified by a large-scale study using the Keio collection of approx. 4,000 mutants in *E. coli* K-12, which identified mutants presenting altered colony size upon growth in more than 300 different conditions[14]. More than 10,000 phenotypes were identified for approx. 2,000 mutants, including many in genes of unknown function[14]. Another major advantage of mutant collections constructed by targeted mutagenesis is that they also reveal complete lists of essential genes, those that cannot be mutated because their loss is lethal, which together compose the minimal genome. These genes that encode proteins essential for cellular life, many of which are expected to be common to all living organisms[15], have attracted particular interest because they are expected to shed light on the origin of life. Additionally, essential genes specific to bacteria are also interesting because they are prime targets for the design of new drugs to tackle antimicrobial resistance.

Since the few species in which complete collections of mutants have been constructed represent only a tiny fraction of bacterial diversity, additional complete

collections of mutants in diverse species are needed to tackle more effectively the challenge of genes of unknown function. Human bacterial pathogens are particularly attractive candidates for this venture because of the potential practical implications for human health. Among these, *Neisseria meningitidis* – a Gram-negative Betaproteobacteria causing life-threatening meningitis and sepsis – stands out as an ideal candidate for several reasons. 1) It has a small genome with approx. 2,000 protein-coding genes, and yet displays a robust metabolism allowing rapid growth (doubling time of 40 min), even on minimal medium. 2) It has been the subject of intensive investigation for decades and there are several thousand publicly available meningococcal genome sequences. 3) It is naturally competent, which makes it a workhorse for genetics. Previously, we have used these advantages to design a modular toolbox for *Neisseria meningitidis* systematic analysis of gene function, named NeMeSys[16]. The central module of NeMeSys was an ordered library of approx. 4,500 Tn mutants in which the Tn insertion sites were sequence-defined[16,17], which contained insertions in more than 900 genes[16].

In the present study, we have 1) used the original library of Tn mutants as a starting point to generate a complete collection of defined mutants in *N. meningitidis*, comprising one mutant in each non-essential gene, 2) defined and analysed in depth the minimal meningococcal genome, and 3) illustrated the potential of this collection for the phenotypic profiling of the meningococcal genome by identifying the function of multiple genes of unknown function.

Results and Discussion

Construction of the complete NeMeSys collection of mutants

The current study gave new impetus to an update of the genome annotation of *N. meningitidis* 8013 published in 2009[16]. The re-annotation took into account new information in the databases, new publications, an RNA-Seq analysis performed in 8013[18], and findings in the present study. In brief, we dropped one previously annotated gene, updated the gene product annotations of 521 genes, added 69 non-coding RNAs, and changed 176 gene start sites. The genome of *N. meningitidis* 8013 contains 2,060 protein-coding genes (labelled NMV_). In parallel, we re-sequenced the genome of the 8013 wild-type (WT) strain by Illumina whole genome sequencing (WGS), which confirmed the accuracy of the original genome sequence determined by Sanger sequencing more than 12 years ago[16], with an error rate of approx. 1 per 134,000 bases. We identified 17 differences with the published sequence, which are likely to be genuine since they were also found in the few mutants that were verified by WGS (see below). Of these differences, 11 correspond to single-nucleotide polymorphisms, five are single-nucleotide indels, and one is the deletion of a GT dinucleotide within a repeated GT tract in NMV_0677, which therefore corresponds to a phase variation event[19].

In keeping with previous studies describing the construction of complete collections of mutants[10-13], 85 genes (4.1 %) were not considered to be significant targets for mutagenesis (Fig. 1) because they encode transposases for highly repeated insertion sequences, or they correspond to short remnants of truncated genes and/or non-expressed cassettes (Supplemental spreadsheet 1). We therefore set out to systematically mutagenise the remaining 1,975 genes (95.9 %) following a two-step approach described below. First, since we previously constructed an arrayed library of Tn mutants in strain 8013 in which Tn insertion sites were sequence-defined[16,17], we selected a subset of suitable mutants from this library.

Specifically, mutations with a Tn insertion closer to the centre of the genes (excluding the first and last 15 % that are often not gene-inactivating) were first re-transformed into strain 8013, which is naturally competent. After extraction of the genomic DNA of the transformants, we PCR-verified that they contained a Tn inserted in the expected target gene using suitable pairs of flanking primers (Supplemental spreadsheet 2). We thus selected a subset of 801 Tn mutants (40.6 % of target genes), harbouring Tn insertions expected to be gene-inactivating (Fig. 1) (Supplemental spreadsheet 3). As a control, Illumina WGS of one mutant, in *nlaB* gene, confirmed that it contained the expected mutation, a Tn inserted after a TA dinucleotide at position 2,157,750 in the genome.

Next, we set out to systematically mutagenise the remaining 1,174 target genes (59.4 %) by allelic exchange. For this, we used a previously validated no-cloning mutagenesis method[20], based on splicing PCR (sPCR), to construct mutants in which a portion of the target gene would be deleted and replaced by the same kanamycin (Km) resistance cassette present in the above Tn. In brief, three PCR products were first amplified, corresponding to the Km cassette and regions upstream and downstream target genes. For failed reactions, we undertook as many rounds of primer design as needed, until all PCRs were successful. The three PCR products were then combined and spliced together. The final sPCR product was transformed directly into *N. meningitidis* 8013. Typically, for non-essential genes, we obtained hundreds of colonies resistant to kanamycin (Km^R). For each successful transformation, two Km^R colonies were isolated and PCR-verified to contain the expected mutation. To minimise false positive identification of essential genes, each transformation that yielded no transformants was repeated at least three times with different sPCR products. Only when all transformations failed to yield transformants, was the target gene deemed to be essential. In summary (Fig. 1), out of the 1,174 genes that were targeted, we could disrupt 783 (Supplemental spreadsheet 3), while 391 could not be mutated and are thus essential (Supplemental spreadsheet 4). As

above, WGS of a few mutants constructed by sPCR, in *lnt*, *tatB* and *secB* genes, confirmed that they each contained the expected mutation.

Taken together, using the above two-step approach, we constructed a comprehensive set of 1,584 defined mutants in strain 8013, named the NeMeSys collection. In addition, we also identified 391 candidate essential genes (19 % of all protein-coding genes), which are required for growth of the meningococcus on rich medium. For each gene that was successfully mutated, one mutant was stored at -80ºC in glycerol, while the corresponding genomic DNA was stored at -20ºC. This allows for easy distribution of mutants to the community and/or re-transformation of the corresponding mutations in 8013, other meningococcal strains, or even genetically closely related species like the gonococcus, which all share most of their genome with the meningococcus.

In depth analysis of the meningococcal minimal genome

Considering their theoretical and practical importance, we first focused on the 391 genes essential for meningococcal life. Since corresponding proteins are expected to be highly conserved, we determined the partition of the 391 essential genes between conserved and variable genomes. To do this, we first used the recently described PPanGGOLiN software[21] – an expectation-maximisation algorithm based on multivariate Bernoulli mixture model coupled with a Markov random field – to compute the pangenome of *N. meningitidis*, based on complete genome sequences of 108 meningococcal isolates available in RefSeq. We thereby classified genes in three categories[21], persistent (gene families present in almost all genomes), shell (present at intermediate frequencies), or cloud (present at low frequency). We then determined how the 2,060 protein-coding genes of *N. meningitidis* 8013 partitioned between these three classes (Supplemental spreadsheet 5). Only 1,664 genes (80.8 %) belong to the persistent genome, while 396 genes (19.2 %) correspond to gene families present at intermediate/low frequency in meningococci since they are part of

the shell (241 genes) and cloud genomes (155 genes) (Fig. 2A). Importantly, besides highlighting the well-known genomic plasticity of this naturally competent species, this analysis confirmed our original prediction by revealing that the essential genes that we have identified are overwhelmingly conserved in meningococcal isolates. Indeed, 382 essential genes (97.4 %) are part of the persistent meningococcal genome (Fig. 2B) (Supplemental spreadsheet 5). Of the remaining nine essential genes, six are part of the shell genome, while three belong to the cloud genome (Supplemental spreadsheet 5).

Furthermore, many meningococcal essential genes (30.2 %) are conserved and essential in other bacteria where systematic mutagenesis efforts have been performed (Fig. 3A), including in phylogenetically distant species such as *S. sanguinis* (Supplemental spreadsheet 6)[10,12,13]. When conservation was assessed in Proteobacteria only (*E. coli* and *A. baylyi*), which are more closely related to *N. meningitidis*, half of the 391 essential meningococcal genes (195 in total) are conserved and essential in these two Gram-negative species (Fig. 3A). Interestingly, a similar proportion of meningococcal essential genes (176 in total) was also conserved in JCVI-syn3.0 (Fig. 3B) (Supplemental spreadsheet 7), a synthetic *Mycoplasma mycoides* bacterium engineered with a minimal genome[22] smaller than that of any naturally occurring autonomously replicating cell. Taken together, these observations are consistent with the notion that a sizeable portion of the minimal genome essential for cellular life is widely conserved in bacteria.

Next, to have a more global understanding of the biological functions that are essential in the meningococcus, we analysed the distribution of the 391 essential genes in functional categories (Supplemental spreadsheet 8). Genes were classified in specific functional categories using the bioinformatic tools embedded in the MicroScope platform[23], which hosts NeMeSys. In particular, we used predictions from MultiFun[24], MetaCyc[25], eggNOG[26], COG[27], FIGfam[28], and/or InterProScan[29]. As can be seen in Table 1, the essential meningococcal genes could be distributed in a

surprisingly limited number of pathways (Table 1). In keeping with expectations, many essential genes are involved in key cellular processes such as 1) transcription (11 genes encode subunits of the RNA polymerase, a series of factors modulating transcription, and two transcriptional regulators), 2) RNA modification/degradation (19 genes), and 3) protein biosynthesis (120 genes) (Table 1). The large class of genes involved in protein biosynthesis encode 1) all 24 proteins involved in tRNA charging, 2) virtually all ribosomal proteins (51/56), 3) five proteins involved in ribosome biogenesis/maturation, 4) 10 factors modulating translation, 5) two enzymes introducing post-translational protein modifications, 6) seven factors facilitating protein folding, and 7) 21 factors involved in protein export (signal peptidase, Lol system involved in lipoprotein export, general secretion Sec system, Tat translocation system, and TAM translocation and assembly module specific for autotransporters) (Table 1). Furthermore, many other essential genes encode proteins involved in additional key processes such as 1) genome replication and maintenance (20 proteins including subunits of the DNA polymerase, topoisomerases, gyrase, ligase *etc.*), 2) cell division (13 proteins), 3) peptidoglycan cell wall biogenesis (17 proteins), and 4) membrane biogenesis (37 proteins) (Table 1). The latter class includes all 12 proteins required for fatty acid biosynthesis, eight proteins required for the biosynthesis of phospholipids (phosphatidylethanolamine (PE) and phosphatidylglycerol (PG)), and 17 proteins involved in the biosynthesis and export of the lipo-oligosaccharide (LOS), which constitutes the external leaflet of the outer membrane. Finally, most of the remaining genes of known function (Table 1) are involved in energy generation, production of important metabolic intermediates (such as dihydroxyacetone phosphate (DHAP) or 5-phosphoribosyl diphosphate (PRPP)) and a variety of metabolic pathways leading to the biosynthesis of 1) vitamins (B1, B2, B6, and B9), 2) nucleotides, 3) amino acids (notably *meso*-diaminopimelate (M-DAP) a component of the peptidoglycan), 4) isoprenoid, 5) heme, 6) CoA/ac-CoA, 7) NAD/NADP, 8) iron-sulfur (Fe-S) clusters, 9) ubiquinol, 10)

lipoate, and 11) S-adenosyl-methionine (SAM). Many of these compounds are cofactors and/or coenzymes known to be essential for the activity of many bacterial enzymes.

Remarkably, when essential genes were further integrated into networks and metabolic pathways – using primarily MetaCyc[25] – the above picture became dramatically simpler. As many as 91.3 % of the meningococcal essential genes partitioned into just four major functional categories (Fig. 4). Namely, 1) gene/protein expression, 2) genome/cell replication, 3) membrane/cell wall biogenesis, and 4) cytosolic metabolism. Strikingly, for most multi-step enzymatic pathways, most, and very often all, the corresponding genes were identified as essential (Fig. 5), although they are most often scattered throughout the genome. We view this as an important quality control of our library. Only 34 essential genes (8.7 %) could not be clearly assigned to these four categories. Critically, this functional partition is highly coherent with the one previously determined for the minimal synthetic bacterium JCVI-syn3.0[22]. As a result, we were able to generate a cellular panorama of the meningococcal essential genome (Fig. 5), in which many of the above pathways are detailed and often linked by critical metabolic intermediates. This panorama provides a useful blueprint for systems and synthetic biology and a global understanding of meningococcal biology.

Essential genes, which are part of the variable genome, are conditionally essential

A puzzling finding was that some essential genes, which could not be assigned to the above four functional groups, are part of shell/cloud genomes present at intermediate/low frequency in the meningococcus (Supplemental spreadsheet 5). Interestingly, many of the corresponding genes are located in regions of genome plasticity (RGP) – often thought to have been acquired by horizontal gene transfer – as confirmed by the novel predictive method panRGP[30] that identifies RGP using

pangenome data generated by PPanGGOLiN[21]. A panRGP analysis identified 32 RGP in the genome of strain 8013, encompassing a total of 348,885 bp (15.3 % of the genome) (Supplemental spreadsheet 9). Interestingly, one of these genes offered a plausible explanation to the apparent paradox of essential genes that are not part of the persistent meningococcal genome. NMV_2289 is predicted to encode a DNA-methyltransferase, based on the presence of a DNA methylase N-4/N-6 domain (InterPro IPR002941). Since the neighbouring NMV_2288 is predicted to encode a restriction enzyme, the probable role of the NMV_2289 methyltransferase is to protect host DNA against degradation by this restriction enzyme. This would explain the lethal phenotype of the ΔNMV_2289 mutant. A closer examination of the other essential genes that are part of shell/cloud genomes and often found in RGP suggests that a similar scenario might be often applicable. We noticed that NMV_1478, which encodes a protein of unknown function in RGP_2, is likely to be co-transcribed with NMV_1479 that is predicted to encode the toxin of a toxin-antitoxin (TA) system[31] (Fig. 6A). Therefore, since a toxin and its cognate neutralising antitoxin are often encoded by closely linked genes, we hypothesized that NMV_1478 might be the antitoxin for the neighbouring NMV_1479 toxin. This would explain why NMV_1478 is essential, *i.e.* in its absence the NMV_1479 toxin would kill the cell. If this is true, we reasoned that it should be possible to delete NMV_1478 together with NMV_1479, which was attempted. As predicted, in contrast to NMV_1478 which could not be deleted individually despite repeated attempts, a double deletion mutant ΔNMV_1478/1479 could be readily obtained (Fig. 6A). This confirms that the NMV_1478 gene is conditionally essential, and only required for cellular life if the toxin product of NMV_1479 is present. Such a scenario was also tested for NMV_0559, which is found in a large RGP of 86.8 kbp (RGP_0), encompassing genes NMV_0527 to NMV_0618/0619 (Fig. 6B). We named this RGP *tps* because it contains multiple *tpsA* genes predicted to encode large haemagglutinin/haemolysin-related proteins of two partner secretion systems[32].

Since NMV_0559 is the only essential gene in the *tps* RGP, we tested whether this gene might be conditionally essential by attempting to delete most of RGP_0. As predicted, while NMV_0559 could not be deleted on its own despite repeated attempts, an 80 kbp Δ*tps* deletion in RGP_0 encompassing NMV_0559 could be readily obtained (Fig. 6B). This suggests that NMV_0559 is conditionally essential, only required for cellular life in the presence of another undefined gene in RGP_0. We predict that several other essential genes of unknown function that are not part of the persistent meningococcal genome may be similarly conditionally essential, including 1) NMV_1305 and NMV_1310 part of a putative prophage (RGP_1), 2) NMV_1541 the putative antitoxin for neighbouring NMV_1539 toxin, 3) NMV_1544 putative antitoxin for neighbouring NMV_1543 toxin, 4) NMV_1761 putative immunity protein against neighbouring MafB toxins (RGP_4), which represent a newly identified family of secreted toxins in pathogenic *Neisseria* species[33], 5) NMV_1918 part of a second *tps* RGP (RGP_7), 6) NMV_2010 the putative immunity protein against neighbouring bacteriocins (RGP_3), and 7) NMV_2333 and NMV_2335 putative immunity proteins in a second *maf* RGP (RGP_11).

Filling holes in metabolic pathways: identification of a novel and widespread histidinol-phosphatase

Another major utility of the NeMeSys collection of mutants is to help determining the function of genes of unknown function by reverse genetics. We therefore wished to illustrate this aspect. We noticed that the 10-step metabolic pathway leading to histidine biosynthesis presents an apparent "hole" (Fig. 7A). The "missing" enzyme in 8013 corresponds to histidinol-phosphatase (EC 3.1.3.15), which catalyses the dephosphorylation of histidinol-P into histidinol (Fig. 7A). The meningococcal genome does not have homologs of known histidinol-phosphatases[34], a feature shared by many other bacterial species according to MetaCyc[25]. First, we excluded the possibility that the meningococcus might be auxotrophic for histidine by showing that

8013 can grow on M9 minimal medium without added histidine (Fig. 7B). Next, we sought to identify the unknown histidinol-phosphatase using NeMeSys, by specifically testing mutants in genes encoding putative phosphatases of unknown function, in search of a mutant that would not grow on M9 without added histidine. Using this approach, we identified ΔNMV_1317 as a histidine auxotroph, growing on M9 plates only when histidine was added (Fig. 7B). The product of this gene was annotated as a putative hydrolase of unknown function, belonging to the IB subfamily of the haloacid dehalogenase superfamily of aspartate-nucleophile hydrolases (IPR006385). The corresponding sequences include a variety of phosphatases none of which is annotated as a histidinol-phosphatase. Complementation of ΔNMV_1317 mutant with NMV_1317 restored growth on M9 without added histidine (Fig. 7B), confirming that the auxotrophic phenotype was due to the mutation in NMV_1317. To confirm that NMV_1317 encodes the missing enzyme in histidine biosynthesis, we performed a cross-species complementation assay with the *hisB* gene from *E. coli* DH5α ($hisB_{EC}$), which encodes a histidinol-phosphatase unrelated to NMV_1317. Complementation of ΔNMV_1317 with $hisB_{EC}$ restored growth on M9 without added histidine (Fig. 7B). Complementation was gene specific, since $hisB_{EC}$ could not complement the growth deficiency of a different histidine auxotrophic mutant in NMV_1778 (*hisH*) (Fig. 7B). Together, these results show that NMV_1317 defines a previously unrecognised class of histidinol-phosphatases, allowing us to fill a hole in the histidine biosynthesis pathway in the meningococcus. This finding, which illustrates the utility of NeMeSys for annotating genes of unknown function by reverse genetics, has implications for other species lacking known histidinol-phosphatases. Indeed, NMV_1317 homologs are widespread in Betaproteobacteria and Gammaproteobacteria, in many of which no histidinol-phosphatase has been identified according to MetaCyc[25]. We expect that in a variety of Burkholderiales and Pseudomonadales, including many important pathogenic species of *Bordetella*,

363 *Burkholderia* and *Pseudomonas*, NMV_1317 homologs will encode the elusive
364 histidinol-phosphatase.
365
366 **Identification of the complete set of genes involved in type IV pilus biology in**
367 ***N. meningitidis*, including three new ones**
368 Another utility of NeMeSys we wished to illustrate is the possibility to perform whole-
369 genome phenotypic screens to identify all the genes responsible for a phenotype of
370 interest. We chose to focus on type IV pili (T4P), which are pivotal virulence factors in
371 the meningococcus[35], for two reasons. First, T4P, are ubiquitous in prokaryotes,
372 which has made them a hot topic for research for the past 30 years[36]. Second,
373 although many aspects of T4P biology remain incompletely understood, most genes
374 composing the multi-protein machinery involved in the assembly of these filaments
375 and their multiple functions have been identified, including in the meningococcus that
376 is one of the mainstream T4P models[35]. Hence, we could readily benchmark the
377 results of our screen against previous mutational analyses, especially for the
378 recovery rate of expected mutants. The two T4P-linked phenotypes we decided to
379 study are the formation of bacterial aggregates and twitching motility. Both
380 phenotypes can be simultaneously assessed – allowing thus a dual screen – by
381 observing the mutants growing in liquid medium by phase-contrast microscopy[37]. We
382 thus scored the 1,589 mutants for the presence of round aggregates, and for the
383 continuous and vigorous jerky movements of cells within these aggregates, which
384 corresponds to twitching motility[37] (Supplemental spreadsheet 10). This analysis
385 revealed that 20 mutants (1.2 %) present phenotypic defects in these two T4P-linked
386 phenotypes (Table 2). Importantly, we identified 100 % of the expected mutants in 17
387 *pil* genes known to affect these phenotypes in the meningococcus[37-39]. This is an
388 important quality control of the NeMeSys collection of mutants, which confirms
389 excellent correlation between phenotype and genotype, and shows that when a
390 robust screening method is used all the genes involved in a given phenotype can be

identified. Furthermore, we could readily identify mutants such as Δ*pilT* able to form (irregular) aggregates but showing no twitching motility[37]. PilT is known to encode the motor powering pilus retraction, which is directly responsible for twitching motility[40]. No other mutant in the library exhibited a similar phenotype. Strikingly, we also identified mutants in three genes not previously associated with T4P biology in *N. meningitidis* (Table 2), which was unexpected considering that T4P have been studied for 30 years in this species. TsaP has been shown in the closely related pathogen *Neisseria gonorrhoeae* to interact with the secretin PilQ[41], which forms a pore in the outer membrane through which T4P translocate onto the cell surface. TsaP plays a poorly understood role in T4P biology in *N. gonorrhoeae* and *Myxococcus xanthus*[41], but has apparently no role in *Pseudomonas aeruginosa*[42]. In addition, we identified NMV_1205 and NMV_2228, which to the best of our knowledge have never been previously linked to T4P biology. NMV_1205 is predicted to encode a periplasmic protein of unknown function and is found mainly in Neisseriales, where it is widespread. NMV_2228 is predicted to encode a carbonic anhydrase, an enzyme catalysing the reversible hydration of carbon dioxide[43].

Next, we analysed the three new mutants in detail using an approach previously validated in the meningococcus with many *pil* genes[37,44]. We first tested whether piliation was affected in the corresponding mutants. We purified T4P using a procedure in which filaments sheared by vortexing are precipitated using ammonium sulfate[37]. Pilus preparations, obtained from equivalent numbers of cells, were separated by SDS-PAGE and either stained using Coomassie blue, or tested by immunoblotting using an anti-PilE antibody. This revealed the major pilin PilE as a 17 kDa species (Fig. 8A) and confirmed that all three mutants were piliated. However, when compared to the WT, pilus yields differed between the mutants. While piliation in ΔNMV_2228 was apparently normal, it was decreased in Δ*tsaP* and ΔNMV_1205. The decrease was dramatic in ΔNMV_1205, where filaments could be detected only by immunoblotting (Fig. 8A). Quantification of piliation using a whole-cell ELISA

procedure[37] confirmed these findings (Fig. 8B) and showed that piliation levels were 40 %, 31 % and 154 % of WT, in Δ*tsaP*, ΔNMV_1205 and ΔNMV_2228, respectively. Therefore, none of the corresponding proteins is absolutely required for T4P biogenesis.

We then further characterised the new mutants for T4P-linked functions. Since it is known that in other meningococcal mutants piliation and aggregation defects can be restored when filament retraction is abolished by a concurrent mutation in *pilT*[37,44], we tested if that might be the case for the new mutants. First, we showed that aggregation, which is abolished in Δ*tsaP* and ΔNMV_1205 and dramatically affected in ΔNMV_2228, could be restored when mutants were complemented with the corresponding WT alleles (Fig. 9A). This confirmed that the phenotypic defects in these mutants were indeed due to the mutations in the above three genes. Similarly, aggregation was restored in Δ*tsaP*/Δ*pilT* and ΔNMV_1205/Δ*pilT* (ΔNMV_2228/Δ*pilT* could not be constructed), which harboured a concurrent mutation in *pilT* that encodes the retraction motor PilT. These findings suggest that *tsaP* and NMV_1205 participate in the formation of aggregates by counterbalancing PilT-mediated pilus retraction. Lastly, we checked whether the mutants were affected for another important T4P-linked phenotype, competence for DNA uptake which makes the meningococcus naturally transformable. We quantified competence in the three mutants as previously described[45] by transformation to rifampicin resistance (Fig. 9B). We found that Δ*tsaP* was almost as transformable as the WT, ΔNMV_1205 showed a 26-fold decrease in transformation, while competence was completely abolished in ΔNMV_2228 (explaining why the ΔNMV_2228/Δ*pilT* double mutant could not be constructed). These findings show that *tsaP*, NMV_1205 and NMV_2228 contribute to the fine-tuning of multiple functions mediated by T4P in *N. meningitidis*.

Taken together, these findings are of significance in two ways. The identification of three new genes playing a role in T4P biology (how exactly remains

to be understood), will contribute to a better understanding of these filaments. This has broad implications because T4P and T4P-related filamentous nanomachines are ubiquitous in Bacteria and Archaea[46]. In addition, the identification of new genes contributing to phenotypes that have been extensively studied in multiple species for the past 30 years is unambiguous evidence of the potential of NeMeSys to lead to a global phenotypic profiling of the meningococcal genome.

Concluding remarks

Here, we described the construction of a complete collection of defined mutants in *N. meningitidis*, which has been fully integrated in our modular NeMeSys toolbox[16], accessible online in MicroScope (https://mage.genoscope.cns.fr/microscope/mage/nemesys). Furthermore, we illustrated NeMeSys utility for tackling the important challenge of genes of unknown function, following a variety of approaches. Although they were limited at this stage to a couple of different biological properties, these experiments have provided new information on meningococcal biology with possible implications for many other species and have clearly demonstrated the potential of NeMeSys to make a significant contribution to ongoing efforts directed towards a comprehensive understanding of a bacterial cell. Such potential is further amplified by several useful properties of *N. meningitidis* such as 1) its small genome with limited functional redundancy, which in species with larger genomes can obscure the link between phenotype and genotype, 2) its hardy nature with a robust metabolism allowing it to grow on minimal medium, 3) its taxonomy making it the first Betaproteobacteria in which a complete library of mutants is available, and 4) the fact that it is a major human pathogen, which allows addressing important virulence properties absent in non-pathogenic model species. This latter point makes the meningococcus a prime model for the identification of new means for controlling bacterial pathogens, which would have crucial practical implications in human health.

The other major achievement in the present study is the identification of the minimal meningococcal genome, comprising 391 genes that could not be disrupted. The finding that more than 90 % of these genes are associated with just four basic biological functions, helped us generate a remarkably coherent cellular panorama of the meningococcal essential metabolism. We surmise that most, if not all, of the remaining essential genes of unknown function, which are highly conserved in other bacterial species, will be involved in either expression of genome information,

preservation of genome information, cell membrane structure/function, or cytosolic metabolism (in particular the production of key metabolic intermediates, vitamins and crucial cofactors/coenzymes). Our cellular panorama provides a useful starting point for systems biology, which together with previous genome-scale metabolic network models[47,48], could help us progress towards a global deciphering of meningococcal biology. This, in turn, would have widespread theoretical implications for our understanding of the bases of cellular life, because many of the meningococcal genes, especially the essential ones, are widely conserved in bacterial.

Materials and methods

Strains and growth conditions

All the *N. meningitidis* strains that were generated and used in this study were derived from a highly adhesive variant – sometimes called clone 12 or 2C43 – of the clinical isolate 8013[49]. This serogroup C strain, which belongs to the sequence type 177 and clonal complex ST-18, has been previously sequenced[16]. Meningococci were routinely grown on plates with GC Agar Base (GCB) (Difco) containing 5 g/l agar (all chemicals were from Sigma unless stated otherwise) and Kellogg's supplements (4 g/l glucose, 0.59 µM thiamine hydrochloride, 12.37 µM $Fe(NO_3)_3 \cdot 9H_2O$, 68.4 µM L-glutamine). Plates were incubated overnight (O/N) at 37°C in a moist atmosphere containing 5 % CO_2. Alternatively, we used agar plates with M9 minimal medium (4 g/l glucose, 4.78 mM Na_2HPO_4, 2.2 mM KH_2PO_4, 1.87 mM NH_4Cl, 8.56 mM NaCl, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$). When required, plates contained 100 µg/ml kanamycin, 3 µg/ml erythromycin, 5 µg/ml rifampicin, 25 µg/ml L-histidine, and/or 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Merck). Strains were stored at -80ºC in 10 % glycerol in liquid GC (15 g/l protease peptone No. 3, 23 mM K_2HPO_4, 7.34 mM KH_2PO_4, 85.6 mM NaCl). *E. coli* TOP10 and DH5α were grown at 37°C in liquid or solid lysogeny broth (LB), which contained 100 µg/ml spectinomycin or 50 µg/ml kanamycin, when appropriate.

Construction of strains

Genomic DNA from *N. meningitidis* and *E. coli* strains was prepared using the Wizard Genomic DNA Purification kit (Promega) following the manufacturer's instructions. Plasmid DNA from *E. coli* strains was purified using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions. Bacteria were transformed as follows. *N. meningitidis* is naturally competent for transformation. A loopful of bacteria grown on GCB plates was resuspended in liquid GC containing 5 mM $MgCl_2$ (GC

transfo), 200 µl was aliquoted in the well of a 24-well plate, and DNA was added. After incubation for 30 min at 37°C on an orbital shaker, 0.8 ml GC transfo was added to the wells and the plates were further incubated for 3 h at 37°C, without shaking. Transformants were selected on GCB plates containing suitable antibiotics. For transformation of *E. coli*, ultra-competent cells were prepared as described elsewhere[50] and transformed by a standard heat shock procedure[51]. Transformants were selected on LB plates containing the suitable antibiotic.

The construction of the NeMeSys library of meningococcal mutants followed a two-step procedure explained in the Results section. First, we selected a subset of potentially suitable mutants from an archived library of Tn mutants with sequence-defined Tn insertion sites, which was described previously[16,17]. The corresponding genomic DNAs were PCR-verified using primers flanking the Tn insertion sites (Supplemental spreadsheet 2). When the mutants were confirmed to be correct, the corresponding mutations were re-transformed in 8013, the mutants were PCR-verified once again and stored at -80ºC. The remaining target genes, for which no Tn mutants were available, were then systematically submitted to targeted mutagenesis using a validated no-cloning method[20]. In brief, using high-fidelity PfuUltra II Fusion HS DNA Polymerase (Agilent) and two sets of specific primers (F1/R1 and F2/R2), we amplified 500-750 bp PCR products upstream and downstream from each target gene, respectively. The R1 and F2 primers were consecutive, non-overlapping and chosen within the target gene (excluding the first and last 30 %). These primers contained 20-mer overhangs complementary to the F3/R3 primers used to amplify the 1,518 bp kanamycin resistance cassette present in the Tn (Supplemental spreadsheet 2). In the first step, three PCR products were amplified separately using F1/R1, F2/R2 and F3/R3 pairs of primers. The first two PCRs contained a 4-fold excess of outer primers F1 and R2[52]. The three products were then combined (3.3 µl of each) and spliced together by PCR using the excess F1/R2 primers added in the first reaction[52]. The sPCR products were directly transformed into *N. meningitidis*. For

each successful transformation, two colonies were isolated and verified by PCR using F1/R2. When transformations yielded no transformants, they were repeated at least three times with different sPCR products. This method was also used to construct several polymutants in which we deleted the 80 kbp RGP_0, the TA system NMV_1478/1479, and three repeated prophages (NMV_1286/1294, NMV_1387/1398, and NMV_1412/1418). For each gene that was successfully mutated, one mutant was individually stored at -80ºC in glycerol, and the tubes were ordered according to the gene NMV_ label. Likewise, corresponding genomic DNAs that allow easy re-transformation of these mutations (even in other *Neisseria* strains), were individually ordered in Eppendorf tubes and stored at -20ºC. Several mutants, as well as the WT strain, were verified by WGS. Sequencing was performed by MicrobesNG on an Illumina sequencer using standard Nextera protocols. An average of 122 Mb of read sequences (between 80 and 179 Mb) were obtained for each sequencing project, representing an average 53-fold genome coverage. The corresponding Illumina reads have been deposited in the European Nucleotide Archive under accession number PRJEB39197.

The Δ*tsaP*/Δ*pilT* and ΔNMV_1205/Δ*pilT* double mutants were constructed by transforming Δ*tsaP* and ΔNMV_1205 with genomic DNA extracted from a *pilT* mutant disrupted by a cloned *ermAM* cassette conferring resistance to erythromycin[53]. ΔNMV_2228/Δ*pilT* could not be constructed because both mutations abolish competence. For complementation assays, complementing genes were amplified using specific indF/indR primers (Supplemental spreadsheet 2), with overhangs introducing flanking *Pac*I sites for cloning, and a ribosome binding site (RBS) in front of the gene. The corresponding PCR products were cloned into pCR8/GW/TOPO (Invitrogen), verified by Sanger sequencing and subcloned into pGCC4[54]. This placed the genes under the transcriptional control of an IPTG-inducible promoter, within a DNA an intragenic region of the gonococcal chromosome conserved in *N. meningitidis*[54]. The resulting plasmids were transformed in the desired

meningococcal mutant, leading to ectopical insertion of the complementing gene. We thus constructed ΔNMV_1317::NMV_1317, ΔNMV_1317::*hisB*$_{EC}$, ΔNMV_1718::*hisB*$_{EC}$, Δ*tsaP*::*tsaP*, ΔNMV_1205::NMV_1205 and ΔNMV_2228::NMV_2228. Since ΔNMV_2228 is not competent, the last strain was constructed by first transforming the corresponding pGCC4 derivative into the WT strain, before transforming the ΔNMV_2228 mutation. Expression of the complementing genes in *N. meningitidis* was induced by growing the strains on GCB plates containing 0.5 mM IPTG.

Dual screen for mutants affected for aggregation and twitching motility

N. meningitidis mutants grown O/N on GCB plates were resuspended individually with a 200 µl pipette tip in the wells of 24-well plates containing 500 µl pre-warmed RPMI 1640 with L-glutamine (PAA Laboratories), supplemented with 10 % heat-inactivated fetal bovine serum Gold (PAA Laboratories). Plates were incubated for approx. 2 h at 37ºC. Aggregates forming on the bottom of the wells were visualised by phase-contrast microscopy using a Nikon TS100F microscope[37]. Digital images of aggregates were recorded using a Sony HDR-CX11 camcorder mounted onto the microscope. In parallel, we scored twitching motility by observing whether bacteria in aggregates exhibited continuous and vigorous jerky movement[37]. This movement is abolished in a Δ*pilT* mutant, in which the T4P retraction motor that powers twitching motility is not produced anymore.

Detection and quantification of T4P

Pilus purification by ammonium sulfate precipitation was carried out as follows. Bacteria grown O/N on GCB plates were first resuspended in 1.5 ml ethanolamine buffer (150 mM ethanolamine, 65 mM NaCl) at pH 10.5. Filaments were sheared by vortexing at maximum speed for 1 min, before OD_{600} was adjusted to 9-12 using ethanolamine buffer (in 1.5 ml final volume). Bacteria were then pelleted by

centrifugation at 17,000 *g* at 4°C during 10 min. The supernatant (1.35 ml) was recovered, topped to 1.5 ml with ethanolamine buffer. This step was repeated once, and filaments were precipitated for approx. 1 h at room temperature by adding 150 µl ethanolamine buffer saturated with ammonium sulphate. Filaments were then pelleted by centrifugation at 17,000 *g* at 4°C during 15 min. Pellets were rinsed once with 100 µl Tris-buffered saline at pH 8, and finally resuspended in 100 µl of Laemmli buffer (BioRad) containing β-mercaptoethanol.

Piliation was quantified by performing whole-cell ELISA[37] using the 20D9 mouse monoclonal antibody that is specific for the T4P in strain 8013[55]. Bacteria grown O/N on GCB plates were resuspended in PBS, adjusted to OD_{600} 0.1 and heat-killed during 1 h at 56ºC. Serial 2-fold dilutions were then aliquoted (100 µl) in the wells of 96-well plates. Each well was made in triplicate. The plates were dried O/N at room temperature in a running safety cabinet. The next day, wells were washed seven times with washing solution (0.1 % Tween 80 in PBS), before adding 100 µl/well of 20D9 antibody (diluted 1/1,000 in washing solution containing 5 % skimmed milk). Plates were incubated 1 h at room temperature, and then washed seven times with washing solution. Next, we added 100 µl/well of peroxidase-linked anti-mouse IgG antibody (Amersham) diluted 1/10,000. Plates were incubated 1 h at room temperature, and then washed seven times with washing solution. We then added 100 µl/well of TMB solution (Thermo Scientific) and incubated the plates during 20 min at room temperature in the dark. Finally, we stopped the reaction by adding 100 µl/well of 0.18 M sulfuric acid and we read the plates at 450 nm using a plate reader. Statistical analyses were performed with Prism (GraphPad Software) using appropriate tests as described in the legends to figures.

SDS-PAGE, Coomassie staining and immunoblotting

Purified T4P were separated by SDS-PAGE using 15 % polyacrylamide gels. Gels were stained using Bio-Safe Coomassie stain (BioRad) or blotted to Amersham

629 Hybond ECL membranes (GE Healthcare) using standard molecular biology
630 techniques[51]. Blocking, incubation with primary/secondary antibodies and detection
631 using Amersham ECL Plus reagents (GE Healthcare) were done following the
632 manufacturer's instructions. The primary antibody was a previously described rabbit
633 anti-PilE antiserum (1/2,500)[56], while the secondary antibody (1/10,000) was a
634 commercial Amersham ECL-HRP linked rabbit IgG (GE Healthcare).
635
636 **Quantifying transformation in the meningococcus**
637 Since the tested mutants are Km^R, we tested their competence by transforming them
638 to Rif^R using DNA from a mutant of 8013 spontaneously resistant to rifampicin[57]. We
639 first amplified by PCR using rpoF/rpoR primers a 1,172 bp internal portion of *rpoB*,
640 which usually contains point mutations[58] leading to Rif^R. This PCR product was
641 cloned in pCR8/GW/TOPO and sequenced, which revealed a single point mutation,
642 leading to a His → Tyr substitution at position 553 in RpoB. This pCR8/GW/TOPO
643 derivative was used to amplify the PCR product used for quantifying competence,
644 which was done as follows. Bacteria grown O/N on GCB plates were resuspended in
645 pre-warmed liquid GC transfo at an OD_{600} of 0.1. The number of bacterial cells in this
646 suspension was quantified by performing colony-forming unit (CFU) counts. We
647 mixed 200 µl of bacterial suspension with 100 ng of transforming DNA in the wells of
648 a 24-well plate. After incubating for 30 min at 37ºC on an orbital shaker, 0.8 ml GC
649 transfo was added to the wells and the plates were further incubated during 3 h at
650 37ºC without shaking. Transformants were selected by plating appropriate dilutions
651 on plates containing rifampicin and by counting, the next day, the number of Rif^R
652 CFU. Transformation frequencies are expressed as % of transformed recipient cells.
653
654 **Bioinformatics**
655 The Qiagen CLC Genomics Workbench software was used for WGS analysis. In
656 brief, Illumina reads from each strain were mapped onto the published sequence of

8013[16], and good quality and high frequency (>70 %) base changes and small deletions/insertions were identified. Large deletions/insertions were identified using annotated assemblies (N50 between 38,959 and 45,097 bp), which were visualised in Artemis[59], by performing pairwise sequence alignments using DNA Strider[60].

The genome of strain 8013 and a total of 108 complete genomes from NCBI RefSeq (last accessed June 8th, 2020) were used to compute the *N. meningitidis* pangenome using the PPanGGOLiN software[21] (version 1.1.85). The original annotations of the genomes have been kept in order to compute gene families. The options "--use_pseudo" and "--defrag" were used to consider pseudogenes in the gene family computation, and to associate families made of fragmented genes with their original gene family, respectively. The other parameters have been used with default values to compute the persistent, shell and cloud partitions. PPanGGOLiN results were used to predict regions of genomic plasticity using the panRGP method with default options[30].

All the datasets generated during this study have been stored within MicroScope[23], which is publicly accessible. This web interface can be used to visualise genomes (simultaneously with synteny maps in other microbial genomes), perform queries (by BLAST or keyword searches) and download all the above datasets in a variety of formats (including EMBL and GenBank).

Acknowledgments

This work was supported by funding from the Wellcome Trust (092290/Z/10/Z) to VP. The France Génomique and French Bioinformatics Institute national infrastructures – funded as part of Investissement d'Avenir program managed by the Agence Nationale pour la Recherche (contracts ANR-10-INBS-09 and ANR-11-INBS-0013) – are acknowledged for support of the MicroScope annotation platform. We thank Angelika Gründling (Imperial College London) and Christoph Tang (University of Oxford) for critical reading of the manuscript.

References

1. Jolley, K.A. & Maiden, M.C. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11**, 595 (2010).
2. Chang, Y.C. *et al.* COMBREX-DB: an experiment centered database of protein function: knowledge, predictions and knowledge gaps. *Nucleic Acids Research* **44**, D330-5 (2016).
3. Galperin, M.Y. & Koonin, E.V. From complete genome sequence to 'complete' understanding? *Trends in Biotechnology* **28**, 398-406 (2010).
4. Akerley, B.J. *et al.* A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 966-971 (2002).
5. Liberati, N.T. *et al.* An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2833-8 (2006).
6. Deutschbauer, A. *et al.* Evidence-based annotation of gene function in *Shewanella oneidensis* MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genetics* **7**, e1002385 (2011).
7. van Opijnen, T., Lazinski, D.W. & Camilli, A. Genome-wide fitness and genetic interactions determined by Tn-seq, a high-throughput massively parallel sequencing method for microorganisms. *Current Protocols in Molecular Biology* **106**, 7.16.1-7.16.24 (2014).
8. Langridge, G.C. *et al.* Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Research* **19**, 2308-16 (2009).
9. Christen, B. *et al.* The essential genome of a bacterium. *Molecular Systems Biology* **7**, 528 (2011).
10. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**, 2006.0008 (2006).

11. Kobayashi, K. *et al.* Essential *Bacillus subtilis* genes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 4678-83 (2003).

12. de Berardinis, V. *et al.* A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Molecular Systems Biology* **4**, 174 (2008).

13. Xu, P. *et al.* Genome-wide essential gene identification in *Streptococcus sanguinis*. *Scientific Reports* **1**, 125 (2011).

14. Nichols, R.J. *et al.* Phenotypic landscape of a bacterial cell. *Cell* **144**, 143-56 (2011).

15. Mushegian, A.R. & Koonin, E.V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 10268-73 (1996).

16. Rusniok, C. *et al.* NeMeSys: a resource for narrowing the gap between sequence and function in the human pathogen *Neisseria meningitidis*. *Genome Biology* **10**, R110 (2009).

17. Geoffroy, M., Floquet, S., Métais, A., Nassif, X. & Pelicic, V. Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. *Genome Research* **13**, 391-8 (2003).

18. Heidrich, N. *et al.* The primary transcriptome of *Neisseria meningitidis* and its interaction with the RNA chaperone Hfq. *Nucleic Acids Research* **45**, 6147-6167 (2017).

19. Martin, P. *et al.* Experimentally revised repertoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation. *Molecular Microbiology* **50**, 245-57 (2003).

20. Georgiadou, M., Castagnini, M., Karimova, G., Ladant, D. & Pelicic, V. Large-scale study of the interactions between proteins involved in type IV pilus biology in *Neisseria meningitidis*: characterization of a subcomplex involved in pilus assembly. *Molecular Microbiology* **84**, 857-873 (2012).

21. Gautreau, G. *et al.* PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph. *PLoS Computational Biology* **16**, e1007732 (2020).

22. Hutchison, C.A., 3rd *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253 (2016).

23. Vallenet, D. *et al.* MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Resesarch* **48**, D579-D589 (2020).

24. Serres, M.H. & Riley, M. MultiFun, a multifunctional classification scheme for *Escherichia coli* K-12 gene products. *Microbial and Comparative Genomics* **5**, 205-22 (2000).

25. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Research* **48**, D445-D453 (2020).

26. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research* **47**, D309-D314 (2019).

27. Galperin, M.Y., Makarova, K.S., Wolf, Y.I. & Koonin, E.V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Research* **43**, D261-9 (2015).

28. Meyer, F., Overbeek, R. & Rodriguez, A. FIGfams: yet another set of protein families. *Nucleic Acids Research* **37**, 6643-54 (2009).

29. Jones, P. *et al.* InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236-40 (2014).

30. Bazin, A., Gautreau, G., Médigue, C., Vallenet, D. & Calteau, A. panRGP: a pangenome-based method to predict genomic islands and explore their diversity. *BioRxiv* **doi**, 10.1101/2020.03.26.007484 (2020).

31. Yamaguchi, Y., Park, J.H. & Inouye, M. Toxin-antitoxin systems in bacteria and archaea. *Annual Review of Genetics* **45**, 61-79 (2011).

767 32. van Ulsen, P., Rutten, L., Feller, M., Tommassen, J. & van der Ende, A. Two-
768 partner secretion systems of *Neisseria meningitidis* associated with invasive clonal
769 complexes. *Infection and Immunity* **76**, 4649-58 (2008).
770 33. Jamet, A. *et al.* A new family of secreted toxins in pathogenic *Neisseria*
771 species. *PLoS Pathogens* **11**, e1004592 (2015).
772 34. Kulis-Horn, R.K., Ruckert, C., Kalinowski, J. & Persicke, M. Sequence-based
773 identification of inositol monophosphatase-like histidinol-phosphate phosphatases
774 (HisN) in *Corynebacterium glutamicum*, Actinobacteria, and beyond. *BMC*
775 *Microbiology* **17**, 161 (2017).
776 35. Pelicic, V. Type IV pili: *e pluribus unum*? *Molecular Microbiology* **68**, 827-37
777 (2008).
778 36. Berry, J.L. & Pelicic, V. Exceptionally widespread nano-machines composed
779 of type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiology Reviews*
780 **39**, 134-154 (2015).
781 37. Brown, D., Helaine, S., Carbonnelle, E. & Pelicic, V. Systematic functional
782 analysis reveals that a set of 7 genes is involved in fine tuning of the multiple
783 functions mediated by type IV pili in *Neisseria meningitidis*. *Infection and Immunity*
784 **78**, 3053-3063 (2010).
785 38. Carbonnelle, E., Helaine, S., Prouvensier, L., Nassif, X. & Pelicic, V. Type IV
786 pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occuring after
787 pilus assembly, essential for fiber stability and function. *Molecular Microbiology*
788 **55**, 54-64 (2005).
789 39. Helaine, S. *et al.* PilX, a pilus-associated protein essential for bacterial
790 aggregation, is a key to pilus-facilitated attachment of *Neisseria meningitidis* to
791 human cells. *Molecular Microbiology* **55**, 65-77 (2005).
792 40. Merz, A.J., So, M. & Sheetz, M.P. Pilus retraction powers bacterial twitching
793 motility. *Nature* **407**, 98-102 (2000).

41. Siewering, K. *et al.* Peptidoglycan-binding protein TsaP functions in surface assembly of type IV pili. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E953-61 (2014).

42. Koo, J., Lamers, R.P., Rubinstein, J.L., Burrows, L.L. & Howell, P.L. Structure of the *Pseudomonas aeruginosa* type IVa pilus secretin at 7.4 A. *Structure* **24**, 1778-1787 (2016).

43. Smith, K.S. & Ferry, J.G. Prokaryotic carbonic anhydrases. *FEMS Microbiology Reviews* **24**, 335-66 (2000).

44. Carbonnelle, E., Helaine, S., Nassif, X. & Pelicic, V. A systematic genetic analysis in *Neisseria meningitidis* defines the Pil proteins required for assembly, functionality, stabilization and export of type IV pili. *Molecular Microbiology* **61**, 1510-22 (2006).

45. Berry, J.L., Cehovin, A., McDowell, M.A., Lea, S.M. & Pelicic, V. Functional analysis of the interdependence between DNA uptake sequence and its cognate ComP receptor during natural transformation in *Neisseria* species. *PLoS Genetics* **9**, e1004014 (2013).

46. Denise, R., Abby, S.S. & Rocha, E.P.C. Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biology* **17**, e3000390 (2019).

47. Baart, G.J. *et al.* Modeling *Neisseria meningitidis* metabolism: from genome to metabolic fluxes. *Genome Biology* **8**, R136 (2007).

48. Mendum, T.A., Newcombe, J., Mannan, A.A., Kierzek, A.M. & McFadden, J. Interrogation of global mutagenesis data with a genome scale model of *Neisseria meningitidis* to assess gene fitness *in vitro* and in sera. *Genome Biology* **12**, R127 (2011).

49. Nassif, X. *et al.* Antigenic variation of pilin regulates adhesion of *Neisseria meningitidis* to human epithelial cells. *Molecular Microbiology* **8**, 719-725 (1993).

821 50. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of
822 *Escherichia coli* with plasmids. *Gene* **96**, 23-8 (1990).
823 51. Sambrook, J. & Russell, D.W. *Molecular cloning. A laboratory manual*, (Cold
824 Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001).
825 52. Metzgar, D. *et al. Acinetobacter* sp. ADP1: an ideal model organism for
826 genetic analysis and genome engineering. *Nucleic Acids Research* **32**, 5780-90
827 (2004).
828 53. Pujol, C., Eugène, E., Marceau, M. & Nassif, X. The meningococcal PilT
829 protein is required for induction of intimate attachment to epithelial cells following
830 pilus-mediated adhesion. *Proceedings of the National Academy of Sciences of the*
831 *United States of America* **96**, 4017-4022 (1999).
832 54. Mehr, I.J., Long, C.D., Serkin, C.D. & Seifert, H.S. A homologue of the
833 recombination-dependent growth gene, *rdgC*, is involved in gonococcal pilin
834 antigenic variation. *Genetics* **154**, 523-32 (2000).
835 55. Pujol, C., Eugène, E., de Saint Martin, L. & Nassif, X. Interaction of *Neisseria*
836 *meningitidis* with a polarized monolayer of epithelial cells. *Infection and Immunity*
837 **65**, 4836-4842 (1997).
838 56. Morand, P.C. *et al.* Type IV pilus retraction in pathogenic *Neisseria* is
839 regulated by the PilC proteins. *The EMBO Journal* **23**, 2009-17 (2004).
840 57. Cehovin, A. *et al.* Specific DNA recognition mediated by a type IV pilin.
841 *Proceedings of the National Academy of Sciences of the United States of America*
842 **110**, 3065-3070 (2013).
843 58. Carter, P.E., Abadi, F.J., Yakubu, D.E. & Pennington, T.H. Molecular
844 characterization of rifampin-resistant *Neisseria meningitidis*. *Antimicrobial Agents*
845 *and Chemotherapy* **38**, 1256-61 (1994).
846 59. Carver, T., Harris, S.R., Berriman, M., Parkhill, J. & McQuillan, J.A. Artemis:
847 an integrated platform for visualization and analysis of high-throughput sequence-
848 based experimental data. *Bioinformatics* **28**, 464-9 (2012).

849 60. Marck, C. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein
850 sequences on the Apple Macintosh family of computers. *Nucleic Acids Research*
851 **16**, 1829-36 (1988).
852 61. Luo, H., Lin, Y., Gao, F., Zhang, C.T. & Zhang, R. DEG 10, an update of the
853 database of essential genes that includes both protein-coding genes and
854 noncoding genomic elements. *Nucleic Acids Research* **42**, D574-80 (2014).

855 **Table 1**. Essential meningococcal genes listed by functional category/sub-category.

856

Functional category/sub-category	Number of genes	% of total
gene/protein expression	**150**	**38.4**
ribosomal protein	51	13
aminoacyl-tRNA synthesis	24	6.1
protein export	21	5.4
RNA modification/degradation	19	4.9
DNA transcription	11	2.8
translation factors	10	2.6
protein folding	7	1.8
ribosome biogenesis/maturation	5	1.3
protein post-translational modification	2	0.5
genome/cell replication	**33**	**8.4**
chromosome replication/maintenance	20	5.1
cell division	13	3.3
cell membrane/wall biogenesis	**54**	**13.8**
peptidoglycan biosynthesis/recycling	17	4.3
LOS biosynthesis	14	3.6
fatty acid biosynthesis	12	3.1
phospholipids biosynthesis	8	2
UDP-GlcNAc biosynthesis	3	0.8
cytosolic metabolism	**120**	**30.7**
vitamins biosynthesis	24	6.1
nucleotide biosynthesis	15	3.8
aerobic respiration/cytochrome c	12	3.1
amino acid metabolism	9	2.3
isoprenoid biosynthesis	9	2.3
electron transport	9	2.3
heme biosynthesis	9	2.3
sugar metabolism	9	2.3
Fe-S cluster biosynthesis	6	1.5
CoA/ac-CoA biosynthesis	6	1.5
ubiquinol biosynthesis	4	1
NAD/NADP biosynthesis	3	0.8
lipoate biosynthesis	2	0.5
SAM biosynthesis	1	0.3
phosphate metabolism	1	0.3
sulfur metabolism	1	0.3
unassigned	**34**	**8.7**
Total	**391**	**100**

857

858 **Table 2**. Mutants in the complete NeMeSys library of meningococcal mutants
859 affected for two functions mediated by T4P: aggregation and twitching motility. The
860 other 1,569 mutants are capable of forming aggregates and exhibit twitching motility.
861 N/A, not assayable.

862

Mutated gene	Product	Aggregation	Twitching
Genes previously known to be involved in T4P biology in *N. meningitidis*			
pilE	major pilin PilE	-	N/A
pilT	type IV pilus retraction ATPase PilT	+ (irregular)	-
pilF	type IV pilus extension ATPase PilF	-	N/A
pilD	leader peptidase / N-methyltransferase PilD	-	N/A
pilG	type IV pilus biogenesis protein PilG	-	N/A
pilW	type IV pilus biogenesis lipoprotein PilW	-	N/A
pilX	minor pilin PilX	-	N/A
pilK	type IV pilus biogenesis protein PilK	-	N/A
pilJ	type IV pilus biogenesis protein PilJ	-	N/A
pilI	type IV pilus biogenesis protein PilI	-	N/A
pilH	type IV pilus biogenesis protein PilH	-	N/A
pilZ	PilZ protein	-	N/A
pilM	type IV pilus biogenesis protein PilM	-	N/A
pilN	type IV pilus biogenesis protein PilN	-	N/A
pilO	type IV pilus biogenesis protein PilO	-	N/A
pilP	type IV pilus biogenesis lipoprotein PilP	-	N/A
pilQ	type IV pilus secretin PilQ	-	N/A
Genes not previously known to be involved in T4P biology in *N. meningitidis*			
tsaP	secretin-associated protein TsaP	-	N/A
NMV_1205	conserved hypothetical periplasmic protein	-	N/A
NMV_2228	putative carbonic anhydrase	+/-	N/A

863

Legends to figures

Fig. 1. Flowchart of the construction of the NeMeSys complete collection of mutants in *N. meningitidis* 8013. As for similar efforts in other bacteria[10,12,13], we first selected protein-coding genes to be targeted by systematic mutagenesis, excluding 85 genes (4.1 %) because they encode transposases of highly repeated insertion sequences, or correspond to short remnants of truncated genes or cassettes. We then followed a two-step mutagenesis approach explained in the text. In brief, we first selected a subset of sequence-verified Tn mutants from a previously constructed arrayed library[16,17]. Mutations were re-transformed in strain 8013 and PCR-verified. We thus selected 801 Tn mutants with a disrupting transposon in the corresponding target genes. Next, we systematically mutagenised the remaining 1,174 target genes using a validated no-cloning mutagenesis method relying on sPCR[20]. For each successful transformation, two colonies were isolated and PCR-verified. To minimise false positive identification of essential genes, each transformation that yielded no transformants was repeated at least three times. In total, we could construct an additional 783 mutants, generating an ordered library of defined mutants in 1,584 meningococcal genes. This effort also identified 391 candidate essential genes, which could not be disrupted, encoding proteins required for *N. meningitidis* growth on rich medium.

Fig. 2. Partition of essential meningococcal genes into persistent, shell, and cloud genomes. To perform this analysis, we used the PPanGGOLiN[21] method as explained in the text. **A**) Partition of the 2,060 genes in the genome of *N. meningitidis* 8013. **B**) Partition of the subset of 391 essential genes identified in this study.

Fig. 3. Comparison of *N. meningitidis* minimal genome (black line) to those of other bacteria. **A**) Three bacteria in which complete libraries of mutants have been

892 constructed: *E. coli* (green line)[10], *A. baylyi* (red line)[12] and *S. sanguinis* (blue line)[13].
893 To perform this analysis, we queried the DEG database[61] of essential genes using
894 our set of essential genes. **B**) JCVI-syn3.0 (red line), a synthetic *M. mycoides*
895 designed with a minimal genome[22].
896
897 **Fig. 4.** Partition of the meningococcal essential genes into four major functional
898 groups. Gene/protein expression (38.4 %, orange), genome/cell replication (8.4 %,
899 green), cell membrane/wall biogenesis (13.8 %, blue), and cytosolic metabolism
900 (30.7 %, yellow). Only 34 essential genes (8.3 %) could not be clearly assigned to
901 one of these four categories.
902
903 **Fig. 5.** Concise cellular overview of the essential meningococcal genome. The four
904 basic functional groups are highlighted using the same colour as in Fig. 4, *i.e.* orange
905 (gene/protein expression), green (genome/cell replication), blue (cell membrane/wall
906 biogenesis), and yellow (cytosolic metabolism). The 34 essential genes that could not
907 be clearly assigned to one of these four categories are not represented on the figure.
908 Genes involved in the different reactions are indicated by their name or NMV_ label,
909 in red when essential, in black when dispensable. *Genes involved in more than one
910 pathway. Key compounds/proteins are abbreviated as follows. ACP, acyl carrier
911 protein; AIR, aminoimidazole ribotide; CDP-DAG, CDP-diacylglycerol; CMP-Kdo,
912 CMP-ketodeoxyoctonate; DHAP, dihydroxyacetone phosphate; DPP, dimethylallyl
913 diphosphate; DXP, 1-deoxyxylulose-5P; Fe-S, iron-sulfur; FMN, flavin
914 mononucleotide; G3P, glyceraldehyde-3P Glu-6P, glucose-6P; IMP, inosine
915 monophosphate; IPP, isopentenyl diphosphate; LOS, lipo-oligosaccharide; M-DAP,
916 *meso*-diaminopimelate; NaMN, nicotinate D-ribonucleotide; OPP, all-*trans*-octaprenyl
917 diphosphate; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PHBA, *p*-
918 hydroxybenzoate; PRPP, 5-phosphoribosyl diphosphate; Rib-5P, ribulose-5P; SAM,

S-adenosyl-methionine; UDP-GlcNAc, UDP-*N*-acetyl-glucosamine; UP, di-*trans*-poly-*cis*-undecaprenyl monophosphate; UPP, di-*trans*-poly-*cis*-undecaprenyl diphosphate.

Fig. 6. Essential genes in RGP – NMV_1479 and NMV-0559 (highlighted in red) – are conditionally essential. **A**) Gene organisation of the putative antitoxin NMV_1478 with its neighbouring NMV_1479 toxin. Results of the mutagenesis (✂) are shown. , viable mutant; , lethal phenotype. **B**) Gene organisation of the *tps* RGP (RGP_0) to which NMV_0559 belongs. Genes on the + strand are in white, genes on the - strand are in black. Results of the mutagenesis (✂) are shown. , viable mutant; , lethal phenotype. In contrast to NMV_0559, each of the other target genes in the *tps* RGP could be mutated individually (not shown for readability).

Fig. 7. NMV_1317 encodes a novel histidinol-phosphatase. **A**) Histidine biosynthesis pathway in the meningococcus. PRPP, 5-phosphoribosyl diphosphate; PRFAR, phosphoribulosylformimino-AICAR-P; IGP, erythro-imidazole-glycerol-P. **B**) Growth on M9 minimal medium, with or without added histidine (His). The plates also contained 0.5 mM IPTG for inducing expression of the complementing genes. WT, strain 8013; Δ1317, ΔNMV_1317 mutant; Δ1317::1317, ΔNMV_1317 complemented with NMV_1317; Δ1317::$hisB_{EC}$, ΔNMV_1317 cross-complemented with $hisB_{EC}$ from *E. coli*, which encodes the unrelated histidinol-phosphatase present in this species; Δ1718, ΔNMV_1718 mutant; Δ1718::$hisB_{EC}$, control showing that ΔNMV_1718 (*hisH*) cannot be cross-complemented by $hisB_{EC}$.

Fig. 8. Assaying piliation in the mutants in genes not previously associated with T4P biology in *N. meningitidis*. The WT strain and a non-piliated *ΔpilD* mutant were included as positive and negative controls, respectively. *ΔtsaP*, *ΔtsaP* mutant; Δ1205, ΔNMV_1205 mutant; Δ2228, ΔNMV_2228 mutant. **A**) T4P purified using a

946 shearing/precipitation method were separated by SDS-PAGE and either stained with
947 Coomassie blue (upper panel) or analysed by immunoblotting using an antibody
948 against the major pilin PilE (lower panel)[56]. Samples were prepared from equivalent
949 numbers of cells and identical volumes were loaded in each lane. MW, molecular
950 weight marker lane, with values in kDa. **B**) T4P were quantified by whole-cell ELISA
951 using a monoclonal antibody specific for the filaments of strain 8013[55]. Equivalent
952 numbers of cells, based on OD_{600} readings, were applied to the wells of microtiter
953 plates. Results are expressed in % piliation (ratio to WT) and are the average ±
954 standard deviations from five independent experiments. For statistical analysis, one-
955 way ANOVAs followed by Dunnett's multiple comparison tests were performed (*** p
956 < 0.0001).
957
958 **Fig 9.** Functional analysis of the mutants in genes not previously associated with T4P
959 biology in *N. meningitidis*. The WT strain and a non-piliated *pilD* mutant were
960 included as positive and negative controls, respectively. **A**) Aggregation in liquid
961 culture as assessed by phase-contrast microscopy[37]. Δ*tsaP*::*tsaP*, Δ*tsaP*
962 complemented with *tsaP*; Δ*tsaP*/Δ*pilT*, double mutant in *tsaP* and *pilT*; Δ1025::1205,
963 ΔNMV_1205 complemented with NMV_1205; Δ1205/Δ*pilT*, double mutant in
964 NMV_1205 and *pilT*; Δ2228::2228, ΔNMV_2228 complemented with NMV_2228. **B**)
965 Quantification of the competence for DNA transformation. Equivalent numbers of
966 recipient cells were transformed using a *rpoB* PCR product, with a point mutation
967 leading to rifampicin resistance. Results are expressed as transformation frequencies
968 and are the average ± standard deviations from four independent experiments. For
969 statistical analysis, one-way ANOVAs followed by Dunnett's multiple comparison
970 tests were performed (*** p < 0.0001).

Legends to supplemental files

Supplemental spreadsheet 1. List of the 1,975 genes in *N. meningitidis* 8013, which were targeted during our systematic mutagenesis, with their essential features. The 85 genes that were not valid targets are also listed separately.

Supplemental spreadsheet 2. List of the primers used in this study. **A**) List of 6,320 primers that were used for the creation of the complete NeMeSys library of mutants. Primers were either used for verifying available Tn mutants (not highlighted), or for sPCR mutagenesis of remaining target genes (highlighted in orange). The 20-mer overhangs in R1 and F2, which are complementary to the primers used to amplify the kanamycin resistance cassette, are in lower case. **B**) List of other primers. Overhangs are in lower case, with restriction sites underlined.

Supplemental spreadsheet 3. List of the 1,589 meningococcal mutants constituting the complete NeMeSys library of mutants. The 801 Tn mutants are not highlighted, while the 788 mutants constructed by sPCR mutagenesis are highlighted in orange.

Supplemental spreadsheet 4. List of the 391 essential meningococcal genes, which could not be mutagenised.

Supplemental spreadsheet 5. Partition of meningococcal genes into persistent (present in almost all meningococci, not highlighted), shell (present at intermediate frequencies, highlighted in gold), or cloud (present at low frequency, highlighted in orange). Partitioning was done using the PPanGGOLiN[21] software on 108 complete *N. meningitidis* genomes, publicly available in RefSeq. The partitioning of the 391 essential genes is also listed separately.

Supplemental spreadsheet 6. Essentiality conservation of meningococcal essential genes in three other bacteria in which complete libraries of mutants have been constructed. Two are Gram-negative Proteobacteria (*E. coli* K-12 and *A. baylyi* ADP1), while one is a Gram-positive Firmicute (*S. sanguinis* SK36).

Supplemental spreadsheet 7. Conservation of meningococcal essential genes in JCVI-syn3.0, a synthetic *M. mycoides* bacterium engineered with a minimal genome[22].

Supplemental spreadsheet 8. Partition of meningococcal essential genes into four major functional groups: 1) gene/protein expression (150 genes), 2) genome/cell replication (33 genes), 3) cell membrane/wall biogenesis (54 genes), and 4) cytosolic metabolism (120 genes). Only 34 essential genes could not be assigned to one of these functional categories.

Supplemental spreadsheet 9. List of the 32 RGP, with the genes they encompass, identified in strain 8013 using the panRGP sotware[30]. Essential genes in these RGP are highlighted in orange.

Supplemental spreadsheet 10. Results of a dual phenotypic screen, using the NeMeSys complete library of mutants, for mutants affected for aggregation and/or twitching motility in liquid culture. The 20 mutants impaired for the formation of aggregates and/or twitching motility are highlighted in orange. The corresponding genes with their essential features are also listed separately.

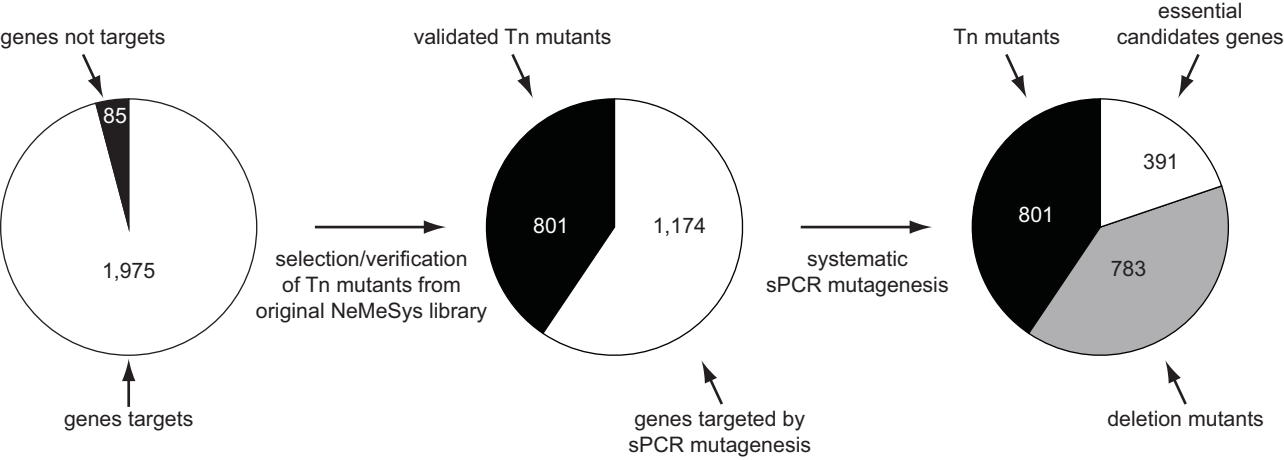

genes not targets
85
1,975
genes targets
validated Tn mutants
801
1,174
genes targeted by
sPCR mutagenesis
selection/verification
of Tn mutants from
original NeMeSys library
Tn mutants
801
systematic
sPCR mutagenesis
essential
candidates genes
391
783
deletion mutants

Figure 1

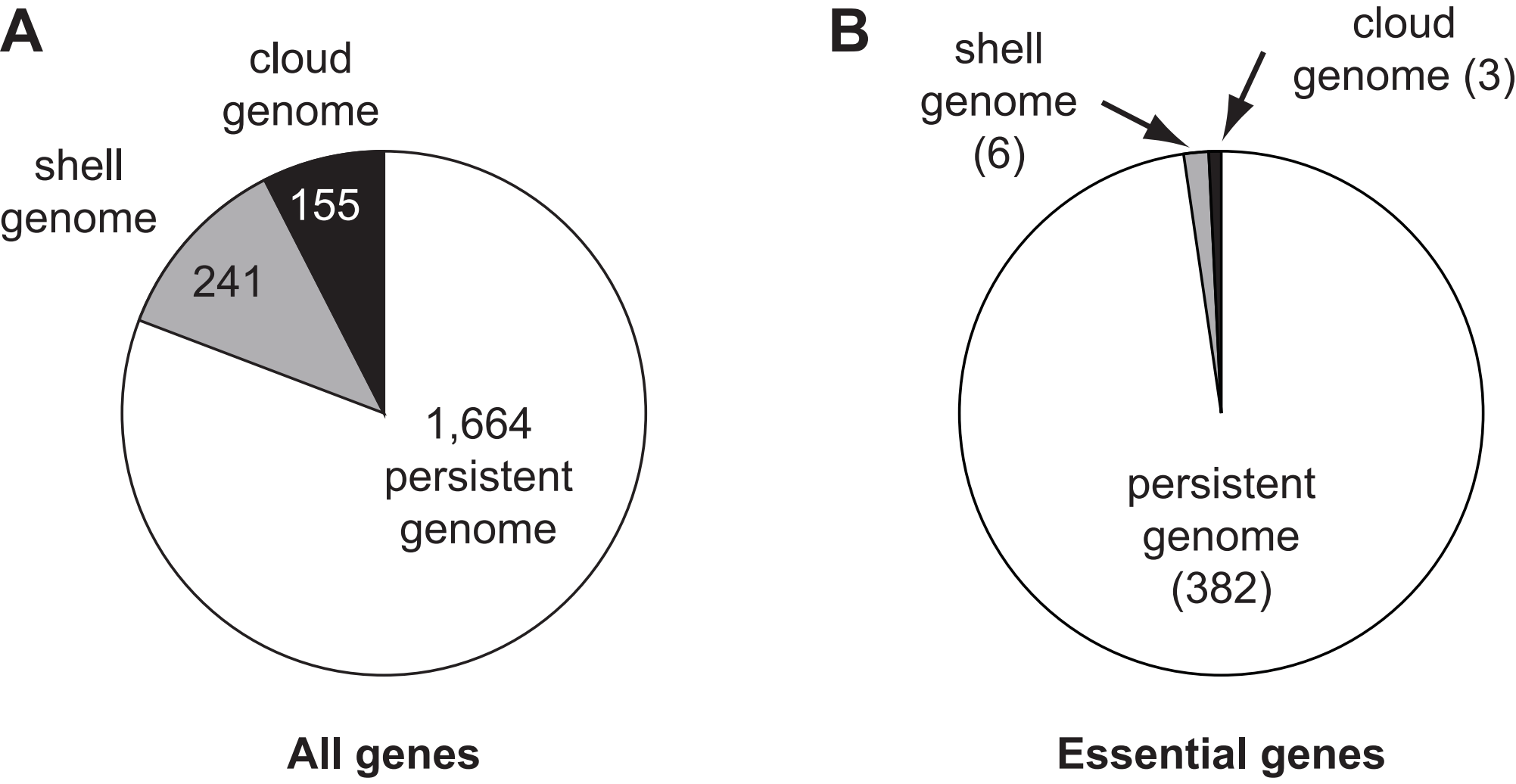
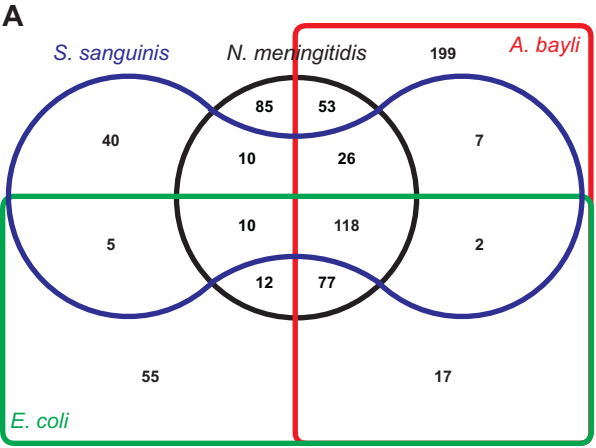

A
cloud genome
shell genome
155
241
1,664 persistent genome
All genes
B
shell genome (6)
cloud genome (3)
persistent genome (382)
Essential genes

Figure 2


A
S. sanguinis
N. meningitidis
A. bayli
E. coli
40
85
53
199
7
10
26
5
10
118
2
12
77
55
17

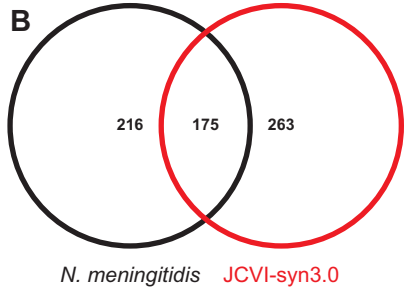

B
216
175
263
N. meningitidis
JCVI-syn3.0

Figure 3

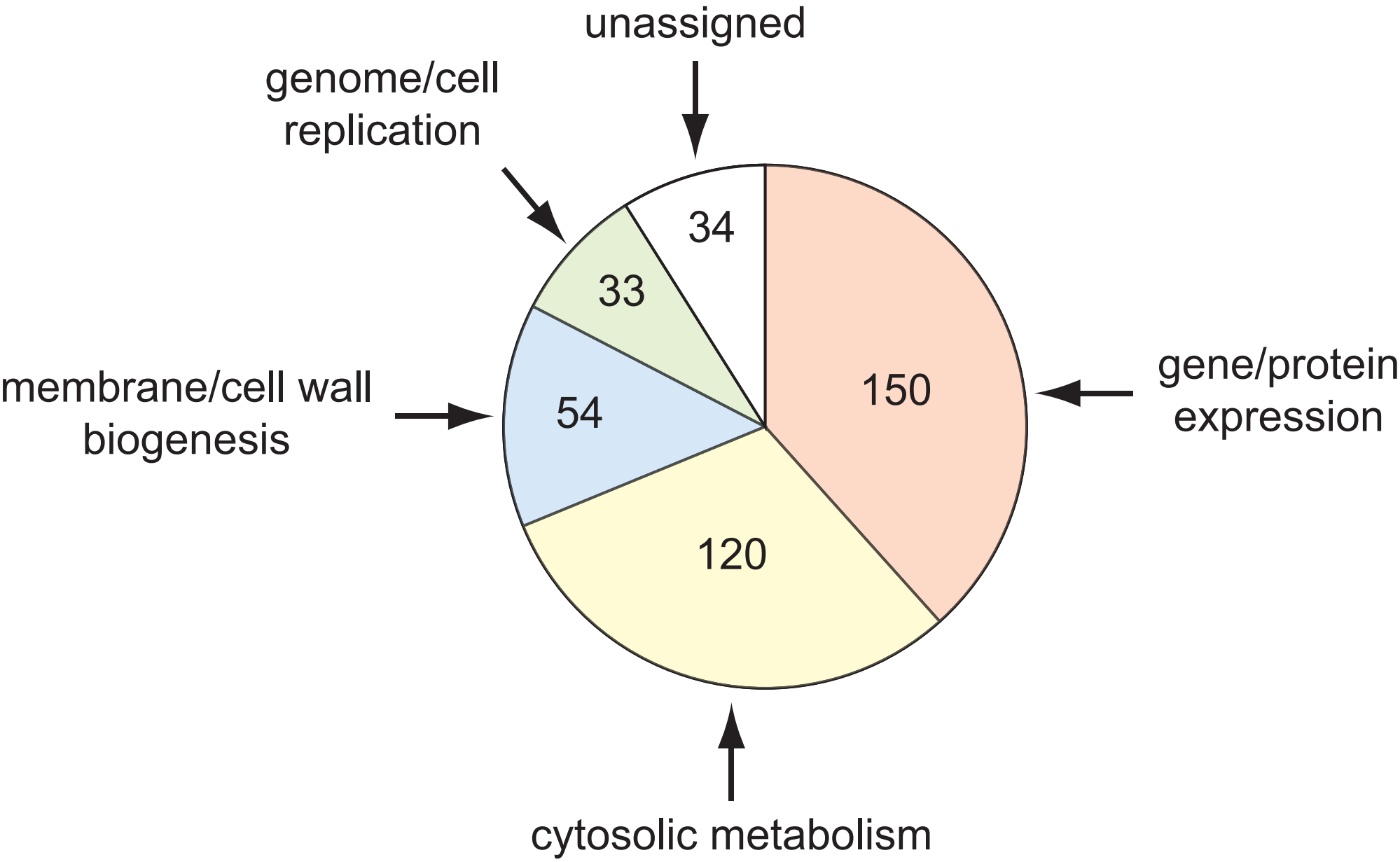

unassigned
genome/cell replication
membrane/cell wall biogenesis
gene/protein expression
cytosolic metabolism
34
33
54
150
120

Figure 4

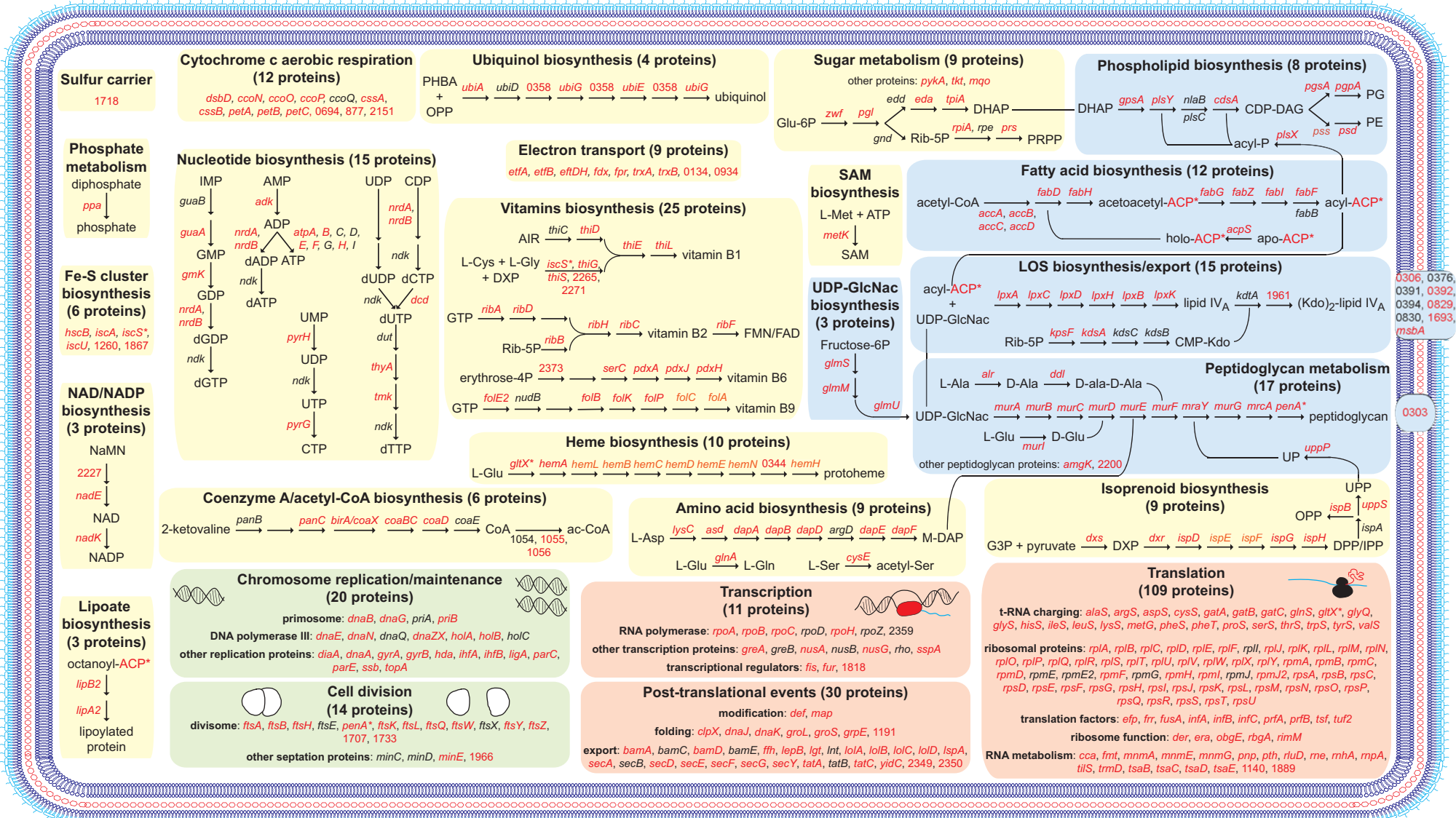

Sulfur carrier
1718
Cytochrome c aerobic respiration (12 proteins)
dsbD, ccoN, ccoO, ccoP, ccoQ, cssA, cssB, petA, petB, petC, 0694, 877, 2151
Ubiquinol biosynthesis (4 proteins)
PHBA + OPP
ubiA
ubiD
0358
ubiG
0358
ubiE
0358
ubiG
ubiquinol
Sugar metabolism (9 proteins)
other proteins: pykA, tkt, mqo
Glu-6P
zwf
pgl
edd
eda
tpiA
DHAP
gnd
Rib-5P
rpiA, rpe
prs
PRPP
Phospholipid biosynthesis (8 proteins)
DHAP
gpsA
plsY
nlaB
plsC
cdsA
CDP-DAG
pgsA
pgpA
PG
pss
psd
PE
acyl-P
plsX
Phosphate metabolism
diphosphate
ppa
phosphate
Nucleotide biosynthesis (15 proteins)
IMP
guaB
guaA
GMP
gmK
GDP
nrdA, nrdB
dGDP
ndk
dGTP
AMP
adk
ADP
nrdA, nrdB
atpA, B, C, D, E, F, G, H, I
dADP ATP
ndk
dATP
UMP
pyrH
UDP
ndk
UTP
pyrG
CTP
UDP
CDP
nrdA, nrdB
ndk
dUDP
dCTP
ndk
dcd
dUTP
dut
thyA
tmk
ndk
dTTP
Electron transport (9 proteins)
etfA, etfB, eftDH, fdx, fpr, trxA, trxB, 0134, 0934
SAM biosynthesis
L-Met + ATP
metK
SAM
Fatty acid biosynthesis (12 proteins)
acetyl-CoA
accA, accB, accC, accD
fabD
fabH
acetoacetyl-ACP*
fabG
fabZ
fabI
fabF
fabB
acyl-ACP*
holo-ACP*
acpS
apo-ACP*
Vitamins biosynthesis (25 proteins)
AIR
thiC
thiD
thiE
thiL
vitamin B1
L-Cys + L-Gly + DXP
iscS*, thiG, thiS, 2265, 2271
GTP
ribA
ribD
ribH
ribC
vitamin B2
ribF
FMN/FAD
Rib-5P
ribB
erythrose-4P
2373
serC
pdxA
pdxJ
pdxH
vitamin B6
GTP
folE2
nudB
folB
folK
folP
folC
folA
vitamin B9
UDP-GlcNac biosynthesis (3 proteins)
Fructose-6P
glmS
glmM
glmU
LOS biosynthesis/export (15 proteins)
acyl-ACP*
+
UDP-GlcNac
lpxA
lpxC
lpxD
lpxH
lpxB
lpxK
lipid IVA
kdtA
1961
(Kdo)2-lipid IVA
Rib-5P
kpsF
kdsA
kdsC
kdsB
CMP-Kdo
0306, 0376, 0391, 0392, 0394, 0829, 0830, 1693, msbA
NAD/NADP biosynthesis (3 proteins)
NaMN
2227
nadE
NAD
nadK
NADP
Peptidoglycan metabolism (17 proteins)
L-Ala
alr
D-Ala
ddl
D-ala-D-Ala
UDP-GlcNac
murA
murB
murC
murD
murE
murF
mraY
murG
mrcA
penA*
peptidoglycan
L-Glu
murI
D-Glu
other peptidoglycan proteins: amgK, 2200
0303
UP
uppP
Heme biosynthesis (10 proteins)
L-Glu
gltX*
hemA
hemL
hemB
hemC
hemD
hemE
hemN
0344
hemH
protoheme
Isoprenoid biosynthesis (9 proteins)
UPP
ispB
uppS
OPP
ispA
G3P + pyruvate
dxs
DXP
dxr
ispD
ispE
ispF
ispG
ispH
DPP/IPP
Coenzyme A/acetyl-CoA biosynthesis (6 proteins)
2-ketovaline
panB
panC
birA/coaX
coaBC
coaD
coaE
CoA
1054, 1055, 1056
ac-CoA
Amino acid biosynthesis (9 proteins)
L-Asp
lysC
asd
dapA
dapB
dapD
argD
dapE
dapF
M-DAP
L-Glu
glnA
L-Gln
L-Ser
cysE
acetyl-Ser
Translation (109 proteins)
t-RNA charging: alaS, argS, aspS, cysS, gatA, gatB, gatC, glnS, gltX*, glyQ, glyS, hisS, ileS, leuS, lysS, metG, pheS, pheT, proS, serS, thrS, trpS, tyrS, valS
ribosomal proteins: rplA, rplB, rplC, rplD, rplE, rplF, rplI, rplJ, rplK, rplL, rplM, rplN, rplO, rplP, rplQ, rplR, rplS, rplT, rplU, rplV, rplW, rplX, rplY, rpmA, rpmB, rpmC, rpmD, rpmE, rpmE2, rpmF, rpmG, rpmH, rpmI, rpmJ, rpmJ2, rpsA, rpsB, rpsC, rpsD, rpsE, rpsF, rpsG, rpsH, rpsI, rpsJ, rpsK, rpsL, rpsM, rpsN, rpsO, rpsP, rpsQ, rpsR, rpsS, rpsT, rpsU
translation factors: efp, frr, fusA, infA, infB, infC, prfA, prfB, tsf, tuf2
ribosome function: der, era, obgE, rbgA, rimM
RNA metabolism: cca, fmt, mnmA, mnmE, mnmG, pnp, pth, rluD, rne, rnhA, rnpA, tilS, trmD, tsaB, tsaC, tsaD, tsaE, 1140, 1889
Fe-S cluster biosynthesis (6 proteins)
hscB, iscA, iscS*, iscU, 1260, 1867
Chromosome replication/maintenance (20 proteins)
primosome: dnaB, dnaG, priA, priB
DNA polymerase III: dnaE, dnaN, dnaQ, dnaZX, holA, holB, holC
other replication proteins: diaA, dnaA, gyrA, gyrB, hda, ihfA, ihfB, ligA, parC, parE, ssb, topA
Transcription (11 proteins)
RNA polymerase: rpoA, rpoB, rpoC, rpoD, rpoH, rpoZ, 2359
other transcription proteins: greA, greB, nusA, nusB, nusG, rho, sspA
transcriptional regulators: fis, fur, 1818
Lipoate biosynthesis (3 proteins)
octanoyl-ACP*
lipB2
lipA2
lipoylated protein
Cell division (14 proteins)
divisome: ftsA, ftsB, ftsH, ftsE, penA*, ftsK, ftsL, ftsQ, ftsW, ftsX, ftsY, ftsZ, 1707, 1733
other septation proteins: minC, minD, minE, 1966
Post-translational events (30 proteins)
modification: def, map
folding: clpX, dnaJ, dnaK, groL, groS, grpE, 1191
export: bamA, bamC, bamD, bamE, ffh, lepB, lgt, lnt, lolA, lolB, lolC, lolD, lspA, secA, secB, secD, secE, secF, secG, secY, tatA, tatB, tatC, yidC, 2349, 2350

Figure 5

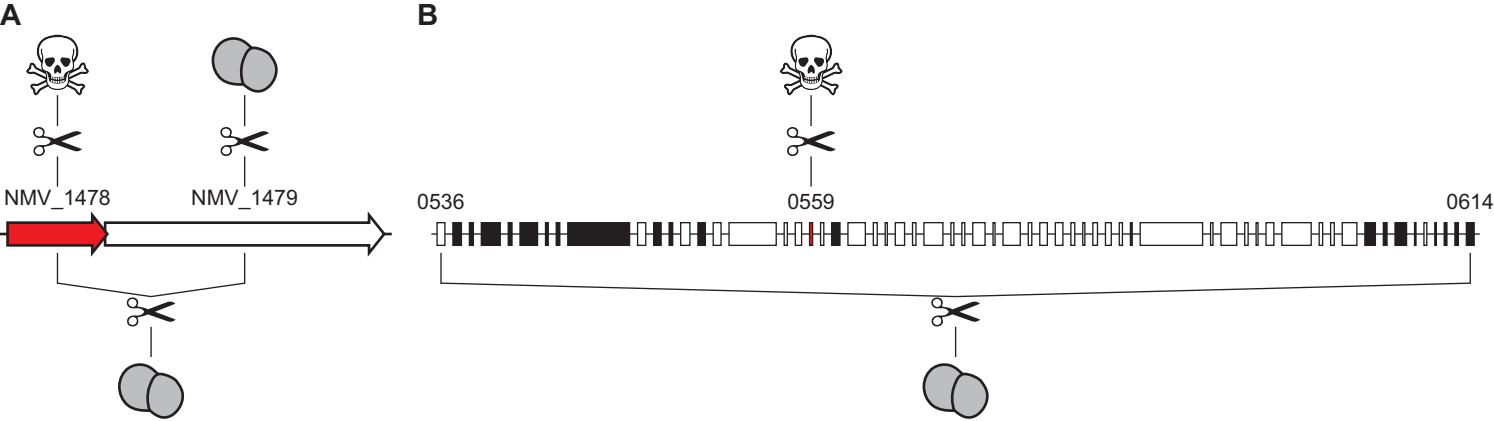


Figure 6

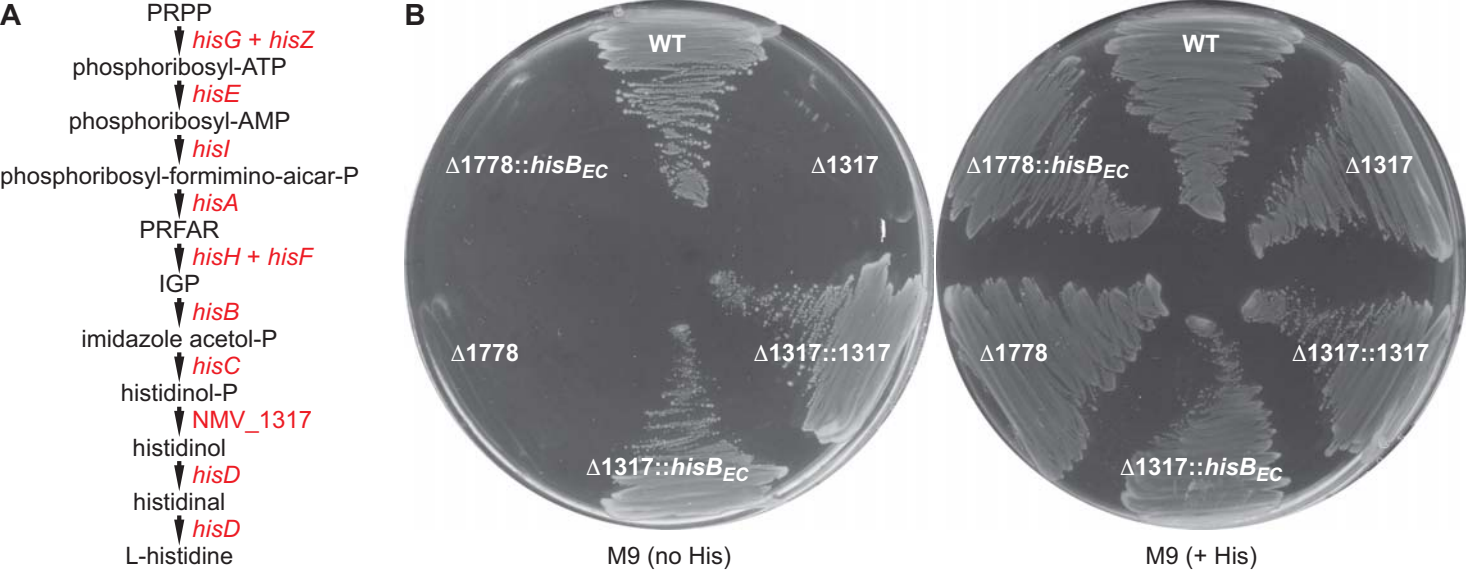

A
PRPP
hisG + hisZ
phosphoribosyl-ATP
hisE
phosphoribosyl-AMP
hisI
phosphoribosyl-formimino-aicar-P
hisA
PRFAR
hisH + hisF
IGP
hisB
imidazole acetol-P
hisC
histidinol-P
NMV_1317
histidinol
hisD
histidinal
hisD
L-histidine
B
WT
Δ1778::hisB_EC
Δ1317
Δ1778
Δ1317::1317
Δ1317::hisB_EC
M9 (no His)
WT
Δ1778::hisB_EC
Δ1317
Δ1778
Δ1317::1317
Δ1317::hisB_EC
M9 (+ His)

Figure 7

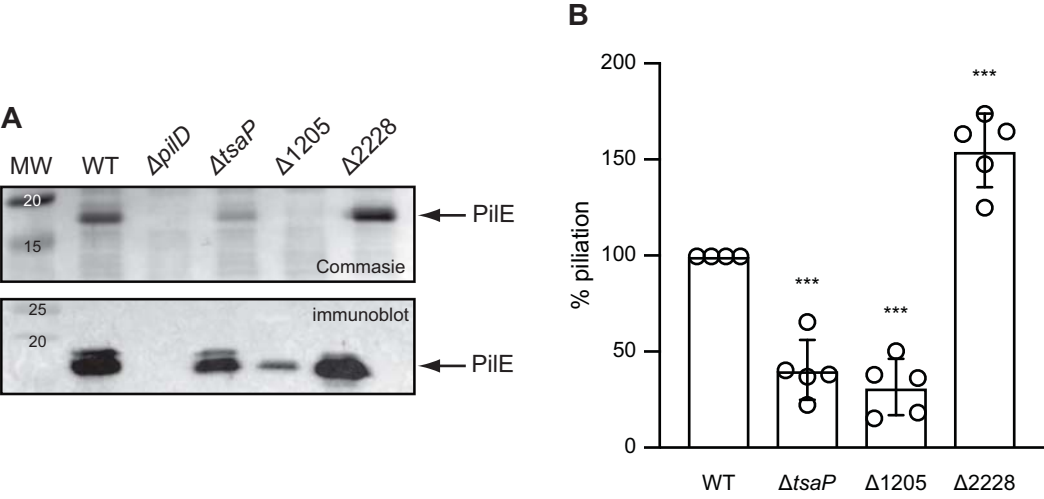
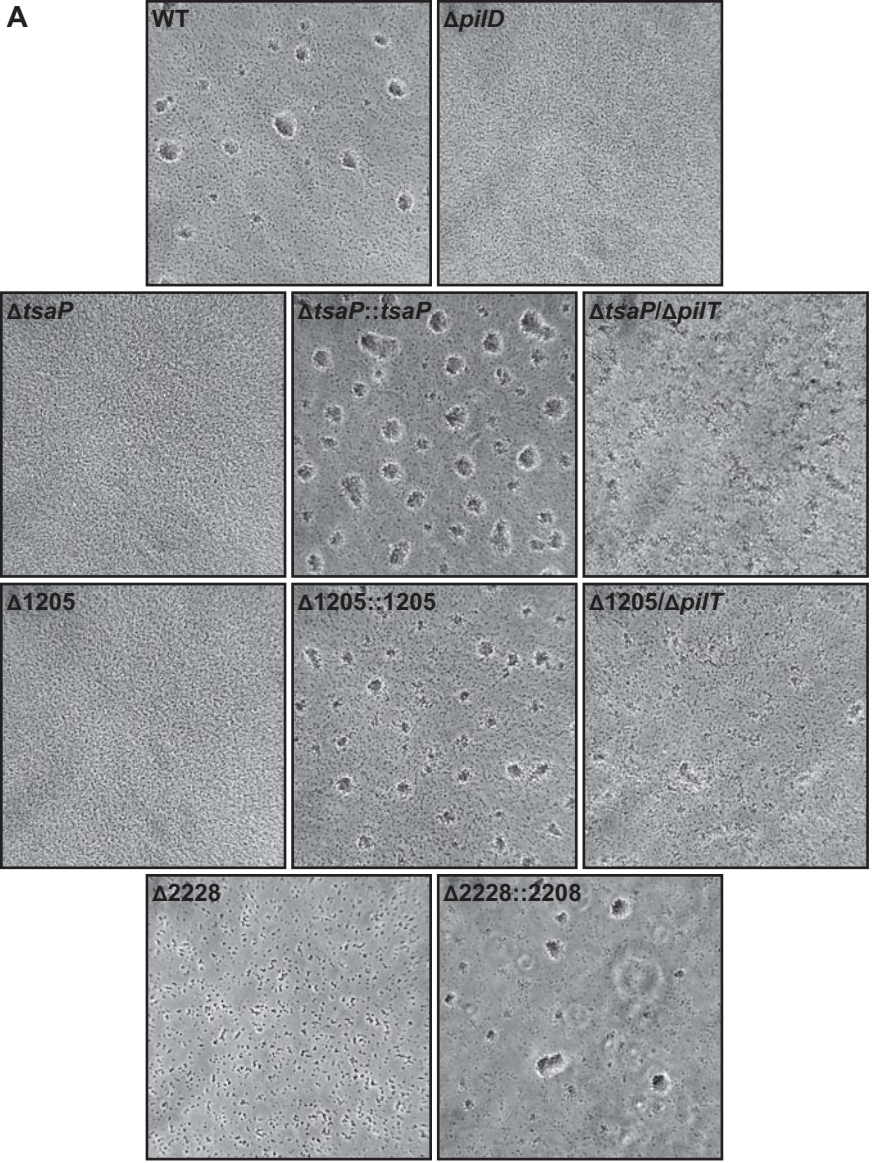

A
MW
WT
ΔpilD
ΔtsaP
Δ1205
Δ2228
20
15
PilE
Commasie
25
20
immunoblot
PilE
B
200
150
100
50
0
% piliation
***
***
***
WT
ΔtsaP
Δ1205
Δ2228

Figure 8


A
WT
ΔpilD
ΔtsaP
ΔtsaP::tsaP
ΔtsaP/ΔpilT
Δ1205
Δ1205::1205
Δ1205/ΔpilT
Δ2228
Δ2228::2208

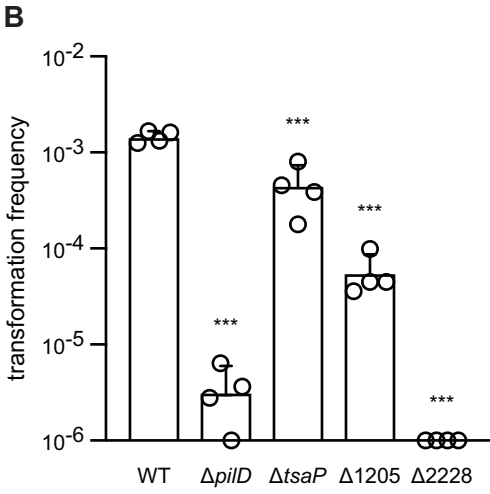

B
transformation frequency
10-2
10-3
10-4
10-5
10-6
***
***
***
***
WT
ΔpilD
ΔtsaP
Δ1205
Δ2228

Figure 9