Autotransporters drive biofilm formation and auto-aggregation in the
diderm Firmicute Veillonella parvula.

Nathalie Béchon1,2,#, Alicia Jiménez-Fernández1,#, Jerzy Witwinowski3, Emilie Bierque1,4,7, Najwa Taib3,5, Thomas Cokelaer5,6, Laurence Ma6, Jean-Marc Ghigo1, Simonetta Gribaldo3 and Christophe Beloin1,*

1 Genetics of Biofilm Laboratory, Institut Pasteur, UMR CNRS2001, Paris, 75015, France
2 Université de Paris, Sorbonne Paris Cité, Paris, France
3 Unit Evolutionary Biology of the Microbial Cell, Institut Pasteur, UMR CNRS2001, Paris, France
4 Sorbonne Université, Collège doctoral, F-75005 Paris, France
5 Hub de Bioinformatique et Biostatistique – Département Biologie Computationnelle, Institut Pasteur, USR 3756 CNRS, Paris, France
6 Plate-forme Technologique Biomics – Centre de Ressources et Recherches Technologiques (C2RT), Institut Pasteur, Paris, France
7 Present address: Leptospirosis Research and Expertise Unit, Institut Pasteur in New Caledonia, Institut Pasteur International Network, Noumea, New Caledonia.

#The authors equally contributed to the work. Author order was determined alphabetically.

Running Title: Autotransporters drive V. parvula biofilm.

*Corresponding author: Christophe Beloin (christophe.beloin@pasteur.fr)
ABSTRACT

The Negativicutes are a clade of Firmicutes that have retained the ancestral diderm character and possess an outer membrane. One of the best studied Negativicute, *Veillonella parvula*, is an anaerobic commensal and opportunistic pathogen inhabiting complex human microbial communities, including the gut and the dental plaque microbiota. Whereas adhesion and biofilm capacity of *V. parvula* is expected to be crucial for its maintenance and development in these environments, studies of *V. parvula* adhesion have been hindered by the lack of efficient genetic tools to perform functional analyses in this bacterium. Here, we took advantage of a recently described naturally transformable *V. parvula* isolate, SKV38, and adapted tools developed for the closely related *Clostridia* spp. to perform random transposon and targeted mutagenesis to identify *V. parvula* genes involved in biofilm formation. We show that type V secreted autotransporters -typically found in diderm bacteria- are the main determinants of *V. parvula* auto-aggregation and biofilm formation, and compete with each other for binding either to cells or to surfaces, with strong consequences on *V. parvula* biofilm formation capacity. The identified trimeric autotransporters have an original structure compared to classical autotransporters identified in Proteobacteria with an additional C-terminal domain. We also show that inactivation of the gene coding for a poorly characterized metal-dependent phosphohydrolase HD domain protein conserved in the Firmicutes and their closely related diderm phyla inhibits autotransporter-mediated biofilm formation. This study paves the way for further molecular characterization of *V. parvula* interactions with other bacteria and the host within complex microbiota environments.

IMPORTANCE

*Veillonella parvula* is an anaerobic commensal and opportunistic pathogen whose ability to adhere to surfaces or other bacteria and form biofilms is critical to inhabit complex human microbial communities such as the gut and oral microbiota. Although the adhesive capacity of *V. parvula* has been previously described, very little is known about the underlying molecular mechanisms due to a lack of genetically amenable *Veillonella* strains. In this study, we took advantage of a naturally transformable *V. parvula* isolate and newly adapted genetic tools to identify surface exposed adhesin called autotransporters as the main molecular determinants of adhesion in this bacterium. This work therefore provides new insights on an important aspect of *V. parvula* lifestyle, opening new possibilities for mechanistic studies of the contribution of biofilm formation to the biology of this major commensal of the oral-digestive tract.
INTRODUCTION

Negativicutes are atypical and poorly studied Firmicute lineages displaying an outer envelope with lipopolysaccharide (1). Among the Negativicutes, Veillonella spp. are anaerobic diderm cocci that commonly inhabit the human and animal microbiota. One of their best studied species, Veillonella parvula (2), is a natural inhabitant of multiple different microbiota, including the human gut (3, 4). V. parvula is considered a commensal organism, and proposed to play a role in the development of immunity through its capacity to colonize the infant gut (5, 6). It is a key early colonizer of the dental plaque during the establishment of sessile microbial communities called biofilms (7), promoting multi-species growth and playing a central role in the metabolism of community members through lactic acid consumption (8). However, V. parvula is also described as an opportunistic pathogen and has been associated with diverse infections, including osteomyelitis, endocarditis, spondylodiscitis, abscesses as well as systemic infections (9–13).

The importance of V. parvula in the development of microbial community spurred our interest in identifying the determinants of its adhesion and biofilm formation capacities. Moreover, considering the presence of an outer membrane in this atypical Firmicute, we wondered whether V. parvula uses known diderm or monoderm biofilm determinants, or currently undescribed adhesion factors. We recently studied V. parvula DSM2008 as a model diderm Firmicute strain (14) to investigate its outer membrane (OM) protein composition and detected 78 OM proteins, thirteen of which being potential adhesins belonging to the type V family of secreted autotransporter proteins (T5SS) (15). Autotransporter proteins are specifically found in diderms and all share common structural and functional features: a Sec-dependent signal peptide, a passenger domain providing the protein function, and an outer-membrane β-barrel domain that allows secretion of the passenger domain (16). However, the challenge of genetic manipulation in V. parvula DSM2008 severely limited the study of these adhesins in this strain.

Here, we have sequenced and annotated the genome of V. parvula SKV38, a recently isolated, naturally transformable and genetically amenable strain (17). We adapted and developed genetic tools for this organism, permitting random and site directed mutagenesis, plasmid complementation and controlled expression using an inducible promoter. This enabled us to identify and characterize factors involved in V. parvula biofilm formation. We find that the main V. parvula biofilm modulating determinants are T5SS adhesins, i.e. typical diderm...
determinants. Interestingly, the identified adhesins possess an additional C-terminal domain compared to the known domain architecture of classical autotransporters. We also show that a locus encoding a metal-dependent phosphohydrolase HD domain protein is involved in biofilm formation, similarly to what was shown in the prototypical monoderm *Bacillus subtilis* (18). Therefore, our results demonstrate that diderm Firmicutes use a mixture of diderm and monoderm factors to modulate their ability to engage into biofilm lifestyle, supporting the idea that monoderm and diderm molecular systems could have co-evolved in these atypical Firmicutes.
RESULTS

Random transposon mutagenesis reveals two *V. parvula* SKV38 genes involved in biofilm formation

In order to obtain a framework for genetic work in the recently described naturally competent *V. parvula* SKV38 isolate, we sequenced it using PacBio technology. We obtained a completely assembled genome of 2.146Mbp, encoding 1,912 predicted protein-encoding open reading frame (ORF), 12 rRNA, 49 tRNA and one tmRNA (see Material and Methods). We performed random transposon mutagenesis in *V. parvula* SKV38 using the pRPF215 plasmid carrying an inducible transposase and a mariner-based transposon previously used to mutagenize *Clostridioides difficile* (19), a close relative of the Negativicutes. We screened 940 individual transposon mutants for biofilm formation using crystal violet staining (CV) static biofilm assay in 96-well microtiter plates and identified eight independent mutants with significant reduction in biofilm formation (Figure 1A). Whole genome sequencing localized the transposons in two loci putatively implicated in biofilm formation (Figure 1B). The most affected mutants correspond to insertions in *FNLLGLLA_00516* (seven mutants), encoding a T5SS type Vc trimeric autotransporter. One transposon mutant was inserted in *FNLLGLLA_01127*, encoding a putative HD phosphatase (Figure 1B).

*FNLLGLLA_00516* encodes a trimeric autotransporter involved in auto-aggregation

*FNLLGLLA_00516* encodes a protein containing several domains usually identified in the T5SS type Vc trimeric autotransporters. Trimeric autotransporters are OM proteins specific of diderm bacteria that have been widely studied for their ability to bind to different surfaces or to other bacteria (20). *FNLLGLLA_00516* is a homolog of *V. parvula* DSM2008 vpar_0464, which encodes a protein that was detected in the OM (15). *FNLLGLLA_00516* was annotated by PROKKA as BtaF, a trimeric autotransporter identified in *Brucella suis* involved in adhesion to extracellular matrix and abiotic surfaces (21). Here, we renamed it *Veillonella* trimeric autotransporter A (VtaA), as the first trimeric autotransporter involved in biofilm formation identified in *V. parvula* SKV38. We deleted the vtaA coding sequence and showed that ΔvtaA had no growth defect (Figure S1A) but displayed a marked reduction of biofilm formation in 96-well polystyrene microtiter plate (Figure 2A). Moreover, while *V. parvula* SKV38 cultures strongly aggregated, ΔvtaA did not (Figure 2B and S2). We constructed the strain p*Tet-vtaA*, where the chromosomal vtaA gene is placed under the control of a functional tetracycline/anhydrotetracycline (aTc) inducible promoter (Figure S3) and showed that its...
aggregation capacity and biofilm formation in 96-well polystyrene microtiter plate directly correlated with aTc concentration (Figure 2C-D), demonstrating that VtaA-mediated cell-to-cell interactions are critical for biofilm formation in these conditions. Whereas the microtiter plate assay corresponds to a static biofilm assay, we also used continuous flow glass microfermentors to investigate the contribution of VtaA to biofilm formation in dynamic conditions. Surprisingly, ΔvtaA formed almost six times more biofilm than the WT strain in these conditions (Figure 2E). Accordingly, scanning electronic microscopy (SEM) images of mature biofilms on microscopic plastic slides in microfermentor showed that ΔvtaA formed a much thicker biofilm when compared to WT (Figure S4). Altogether, these results suggest that auto-aggregation differentially contributes to biofilm formation in static conditions on hydrophobic surfaces versus continuous flow conditions on hydrophilic surfaces.

*V. parvula* SKV38 encodes sixteen putative autotransporters in addition to VtaA

The strong biofilm phenotype displayed by the ΔvtaA mutant in microfermentor led us to suspect that additional adhesins could modulate *V. parvula* biofilm formation capacity. Indeed, searching the *V. parvula* SKV38 genome revealed multiple genes encoding autotransporters (Table 1): three Va classical monomeric autotransporters with a characteristic PFAM_PF03797 autotransporter β-domain (renamed *Veillonella* monomeric autotransporter A to C: VmaA to C), and eight other putative Vc trimeric autotransporters with a characteristic PFAM_PF03895 YadA _anchor_domain (renamed *Veillonella* trimeric autotransporter B to I: VtaB to I). We also identified several partial autotransporters: FNLLGLLA_00035, that only contains a PFAM_PF11924 Ve inverse autotransporter β-domain but no putative α-domain that normally carries the function of the protein, and FNLLGLLA_00036-37 and FNLLGLLA_00040-41, which are homologs of *V. parvula* DSM2008 Vpar_0041 and Vpar_0048, respectively, and that appear to be split in SKV38 (Table 1). Interestingly, domain analysis of all trimeric ATs of *V. parvula* SKV38 showed that they possess an extra C-terminal-domain (SLH or coiled-coil domain) after the YadA anchor domain that is not found in classical trimeric ATs. Among those, six autotransporters plus FNLLGLLA_00035, FNLLGLLA_00036-37 and FNLLGLLA_00040-41 form a potential genomic cluster coding for adhesins (Figure 3A), whereas the six others are located in different areas of the genome (Figure 3B).

We selected eight *Veillonella* strains to study more precisely the evolution of the adhesin cluster, including SKV38 and DSM2008. The trimeric autotransporter adhesins seem to evolve dynamically with numerous domain swaps, duplications and reductions of gene copies, likely...
through homologous recombination suggesting rapid evolutionary changes in the repertoire of
Veillonella adhesins (Figure 4). Duplications and deletions could be eased by the presence of
short ORFs annotated as hypothetical proteins presenting a high degree of sequence identity.
The most basal strain in the Veillonella phylogeny has a minimal cluster of only three adhesins
genes. Throughout the Veillonella genus, the size of the cluster is very variable with a minimal
form in V. atypica, with only two adhesins. This specific adhesin locus, immediately upstream
of rRNA coding genes, is to our knowledge a peculiar genomic character of the Veillonella
genus and is not found in other genera of the Veillonellaceae, suggesting it originated in the
common ancestor of all Veillonella species.

The cluster of trimeric autotransporters is involved in surface binding and not
aggregation.

To assess the function of the potential adhesins identified in the V. parvula SKV38
genome, we constructed -within the cluster of adhesin genes- independent deletion mutants for
the two first autotransporters (vmaA and vtaB) and a large deletion for the eight adjacent genes
encoding trimeric autotransporters or partial trimeric autotransporters, hereafter called Δ8
(ΔFNLLGLLA_00036 to vtaF). We also generated independent individual mutants for each of
the six additional autotransporters located outside of the cluster. These mutants were all tested
for biofilm formation in 96-well polystyrene plate and aggregation capacities. None of the
mutants, with the exception of the previously mentioned ΔvtaA strain, was affected for
aggregation capacities (Figure 5A). The Δ8 mutant was the sole mutant, in addition to the ΔvtaA
mutant, to display lower biofilm formation in 96-well polystyrene microtiter plate (Figure 5B
and C), suggesting that the adhesins of this cluster could be involved in biofilm formation
independently of cell-to-cell interactions. When tested in microfermentor, Δ8 displayed a
slightly reduced ability to form mature biofilm, however, not statistically different from WT
(Figure 5D). This reduced ability to form mature biofilms was actually more visible when
observing SEM images, since the Δ8 mutant only poorly covered the coverslip with sporadic
aggregates of cells producing extracellular matrix (Figure S4). Initial adhesion assay to glass
spatula showed that both ΔvtaA and Δ8 displayed a lower percentage of initial adhesion than
WT, suggesting that VtaA-mediated auto-aggregation contributed to initial adhesion of the WT
strain while the adhesin cluster is probably directly involved in surface binding (Figure 5E).
This also indicates that ΔvtaA does not adhere to glass better than WT, and so the increased
biofilm formation of Δvta in microfermenter arises during the continuous flow culture step. The
effect of deleting vtaA and the 8 adhesin genes on initial adhesion was additive since a ΔvtaAΔ8
double mutant showed a reduced initial adhesion on microfermentor spatula compared to either
WT, ΔvtaA or Δ8 (Figure 5E). In addition, ΔvtaAΔ8 formed 17 times less biofilm than ΔvtaA
in microfermentor, indicating that in the absence of VtaA, the adhesins encoded by some of
these eight genes strongly promote mature biofilm formation in microfermentor (Figure 5D).

Taken together, these results demonstrate the differential contribution of VtaA and part
of the cluster of adhesin to V. parvula SKV38 adhesion and highlight the existence of potential
interference mechanisms between them.

**FNLLGLLA_01127 encodes an HD phosphatase that inhibits biofilm formation**

In addition to genes encoding potential T5SS proteins, we also identified a transposon
mutant in FNLLGLLA_01127, encoding a protein of the HD phosphatase superfamily (Figure
1B). The FNLLGLLA_01127 gene is homologous to YqeK, a putative phosphatase required for
pellicle formation and the development of biofilm in B. subtilis (18). FNLLGLLA_01127/yqeK
is found in a cluster of genes (obb, yhbY, proB, proA, nadD, yqeK, lytR, and rsfS), whose
synteny is very well conserved among Negativicutes. This cluster, or part of it, is also well
conserved in almost all Firmicutes genomes we analyzed, both monoderm and diderm, as well
as in members of other diderm phyla phylogenetically close to the Firmicutes, notably
Deinococcus-Thermus (Figure 6 and S5, DataSet 2). A FNLLGLLA_01127 deletion mutant
(Δ1127) had a lower carrying capacity compared to WT, maybe due to higher mortality during
the stationary phase (Figure S1) and a moderate 1.5-fold decrease in biofilm formation in
microtiter plate after correcting for the growth defect (Figure 7A). This mutant also displayed
a slightly faster aggregation rate than the WT during early time points (Figure 7B). The
strongest phenotype of this mutant was detected in microfermentor with a 9-fold increase in
biofilm formation compared to WT (Figure 7C). Expression of FNLLGLLA_01127 gene in
trans (plasmid p1127) did not complement the observed growth defect (Figure S1B) but it did
complement the increased biofilm formation in microfermentor (Figure 7D), showing that
deletion of FNLLGLLA_01127 might have had polar effects on downstream genes of the operon
causing a growth defect, but that FNLLGLLA_01127 alone was responsible for the observed
inhibition of biofilm formation. Scanning electronic microscopy showed that Δ1127, similarly
to ΔvtaA, formed a thick layered biofilm, although with fewer filaments and protein deposits
compared to WT (Figure 7E). However, contrary to ΔvtaA or Δ8 mutants, Δ1127 showed no
defect in initial adhesion to a glass spatula (Figure 7F). Interestingly, a Δ1127Δ8 double mutant
formed almost 20 times less biofilm than Δ1127 in microfermentor (Figure 7C), suggesting that at least some of the autotransporters of the cluster were necessary for Δ1127 observed strong biofilm formation in microfermentor.
DISCUSSION

Originally described as a social organism mostly living in biofilm communities (8), *Veillonella* is a known bacterial member of multiple human microbiota. Biofilm formation and adhesion are important in these niches, but their study in *Veillonella* has been hindered by the lack of efficient genetic tools. Here, we used genetics tools adapted from *Clostridia* to characterize factors promoting biofilm formation in a naturally competent *Veillonella parvula* isolate.

We identified a T5SS type Vc trimeric autotransporter, FNLLGLLA_0516 (VtaA), as an important biofilm factor promoting *V. parvula* SKV38 auto-aggregation. In addition to Hag1, a YadA-like autotransporter identified from the related species *V. atypica* involved in interspecies interactions (22), VtaA represents the second *Veillonella* protein described which is involved in adhesion, and the first involved in abiotic surface adhesion and auto-aggregation in diderm Firmicutes. Beyond the potential impact on *Veillonella* niche colonization, aggregation capacity is known to contribute to bacterial protection from environmental stresses or host responses (23), promotion of host colonization (24), or pathogenesis (25) in various bacterial species. VtaA is homologous to *Brucella suis* trimeric autotransporter BtaF. However, while *B. suis* BtaF promotes biofilm formation in vitro, it was not shown to promote aggregation (21), suggesting that these two proteins have different functions.

In diderm bacteria such as *E. coli*, self-associating autotransporters (SAATs) from the type Va family and type Vc trimeric autotransporters were shown to contribute to biofilm formation through their self-recognition properties (26–32). However, in *V. parvula* VtaA-mediated auto-aggregation either promoted (plastic surface and static conditions) or strongly impaired (glass surface and continuous-flow conditions) biofilm formation depending on the model used. ∆vtaA initially adhered less to glass spatula compared to WT, even though later on it formed much more biofilm, thus we suspect that the material (glass vs. plastic) is not responsible for the observed difference between our two systems. We hypothesize instead that in the WT, VtaA-mediated aggregates are more sensitive to flow than individual cells, and are thus washed out faster of the microfermentor, and that adhesion to surfaces or to the biofilm extracellular matrix is more important than cell-to-cell interactions when the culture is performed under continuous flow.

Interference between cell surface structures is a well-described mechanism by which bacteria modulate their adhesion properties. In *E. coli*, multiple structures, such as chaperone-usher...
fimbriae, LPS O-antigen or capsules, interfere with the self-recognizing autotransporter Ag43 though unknown mechanisms (33–36). Therefore, it is possible that in *V. parvula*, VtaA could compete with other adhesins through steric hindrance or competition for membrane export and thus limit biofilm formation under continuous-flow conditions. Consistently, ΔvtaA enhanced biofilm formation in microfermentor was dependent on the presence of eight genes of the cluster of trimeric autotransporters, suggesting a competition between VtaA and adhesin(s) of this cluster. Moreover, we noticed that both VtaA and the 8-gene cluster are necessary for full initial adhesion to glass spatula in an independent manner, suggesting that any competition between them only arises later on, during continuous-flow cultures. The exact contribution of these different trimeric autotransporters to biofilm formation and their interplay with VtaA will require further characterization.

Analysis of *V. parvula* SKV38 genome revealed the presence of seven other potential full-length autotransporters, but no other types of classical dierm adhesins. None of them appeared to be involved in cell-to-cell interactions or biofilm formation on abiotic surfaces, and their function remains to be fully elucidated. As *V. parvula* is present in different microbiota, it is expected that a large arsenal of adhesion factors is necessary to adhere under different mechanical constraints and on different surfaces, such as tooth enamel or various epithelia. Moreover, *Veillonella* is known to co-aggregate with *Streptococci* (37–39), that produce *Veillonella* favored carbon source, lactate (40), and they were shown to specifically co-aggregate with *Streptococci* and *Actinomyces* strains from the same microbiota, showing that co-aggregation could have strong implication in niche colonization of these bacteria (41). *V. parvula* and other *Veillonella* are also associated to different opportunistic infections and the contribution of their adhesins to pathogenicity remains to be addressed. Finally, some autotransporters have been shown to carry non-adhesive functions, including protease activity (42), but we detected no classical protease domain in the *Veillonella* autotransporters.

Trimeric autotransporters possess a characteristic YadA_anchor domain (PF03895) that is found mainly in Proteobacteria, but also in Cyanobacteria, Verrucomicrobia, Planctomycetes, Kiritimatiellaeota, Chlorobi, Synergistetes, Fusobacteria and Negativicutes (https://pfam.xfam.org/family/PF03895, Dec 2019 (43)). Interestingly, the YadA_anchor of *V. parvula* SKV38 and all *Veillonella* trimeric autotransporters is not at the very end of the C-terminus, where it is usually found in Proteobacteria, but is pre-C-terminal, followed by either a coiled domain or a S-layer homology (SLH) domain (Figure 3, DataSet 1). While the function
of the coiled domain is unknown, in some bacteria the periplasmic SLH domain binds to peptidoglycan (44), suggesting that *Veillonella* trimeric autotransporters could be non-covalently attached to the peptidoglycan. These extra-domains after the YadA_anchor are also found in other Negativicutes (notably the extra SLH domain) and in some other dierm phyla phylogenetically related to the Firmicutes such as Synergistetes and Fusobacteria (DataSet 1).

In addition to possessing trimeric autotransporters with an extra coiled C-terminus domain, the Fusobacterium *Streptobacillus moniliformis* ATCC14647 carries eight genes encoding unique trimeric autotransporters with an extra OmpA_family domain (PF00691) at their extreme C-terminus, a domain known to display affinity to peptidoglycan (45) (DataSet 1). These data suggest that a subset of phylogenetically close dierm bacteria have evolved trimeric autotransporters integrating different peptidoglycan binding domains. Whether these domains have an impact on trimeric autotransporters function or exposure to the surface, or more generally on outer membrane stabilization is presently unknown.

Our screening also led to the identification of FNLLGLLA_01127, the homolog of *B. subtilis* YqeK, a putative phosphatase required for pellicle formation and the development of biofilm (18). *Staphylococcus aureus* YqeK was recently shown to be a nucleosidase hydrolyzing diadenosine-tetraphosphate (Ap4A) into ADP (46). In *Pseudomonas fluorescens*, an increased level of Ap4A increases cyclic-di-GMP concentration and enhances cell-surface exposure of a large adhesin LapA, thus inducing biofilm formation (47). c-di-GMP regulates biofilm formation by modulating production of a variety cell-surface appendages or exopolysaccharides in both monoderm and dierm bacteria (48–52). Interestingly, *B. subtilis* YqeK induces the epsA-O operon, involved in the production of biofilm matrix-forming polysaccharides (53).

Deletion of *V. parvula* FNLLGLLA_01127 only led to a minor decrease in biofilm formation in 96-well plate, but to a strong increase in continuous-flow biofilm formation that was dependent on the presence of the cluster of trimeric autotransporters. Further work is needed to determine whether FNLLGLLA_01127 directly impacts production of the adhesins of the cluster or participates to the production/regulation of an unknown exopolysaccharide, which, contrary to *B. subtilis*, would interfere with the function or exposure of the adhesins of the cluster rather than favor community development.

In this study we have shown that classical dierm trimeric autotransporters and a potential nucleotidase, conserved both in monoderm and dierm are crucial for adhesion both between cells and/or to surfaces in the dierm Firmicute *V. parvula*. Our work also underscores the rapid
evolution of a diverse arsenal of trimeric autotransporters in the *Veillonella* genus, both in numbers and size, probably by efficient recombination favored by gene clustering, allowing rapid adaptation to changing environments. Taken together, our results suggest a complex interplay at the surface of *V. parvula* between different cell surface structures that may have co-evolved for a long time in these atypical Firmicutes. Much remains to be discovered on the regulatory circuits controlling these adhesion factors and their role in diderm Firmicutes biology.
MATERIALS AND METHODS

Genome preparation and sequencing

*V. parvula* SKV38 genomic DNA was extracted using Qiagen genomic tip 20G kit. It was sequenced to 1,500X coverage using PacBio sequencing of one single molecule real time (SMRT) cell with no multiplexing using the V2.1 chemistry. Only one SMRT cell was used but with no multiplexing, leading to an unusually large amount of subreads: 3 Gbp, meaning about 1,500X coverage assuming a 2.1 Mbp genome. This yielded 338,310 reads with a mean subread length of 9,080 bp and N50 read length of 13,500 bp. The longest subread length is above 70 kbp. We randomly subsampled the data to avoid misassemblies keeping only 100,000 subreads, which resulted in a 430X coverage. The genome was then assembled using Canu version 1.8 (54) keeping the default parameters. In particular, subread below 1,000 bp were dropped. The error correction steps of the Canu algorithm were not tuned, keeping the parameters that control alignment seed length, read length, overlap length and error rates to their default values. We obtained one contig of 2.146 Mbp and an additional contig of only 1,972 bp that was abandoned due to lack of supporting data and was removed by the circularization process. The obtained assembled genome closely matched the genome size (2.1422 Mbp) and GC content (38.7%, expected 38.6%) of the reference *V. parvula* DSM2008 strain. The resulting assembled genome was polished using Pilon (55) but no correction was required. No gaps or drops of coverage was detected based on sequana_coverage output (56, 57). The completeness of the candidate assembly was assessed to be 98% using the bacteria mode and the bacteria_db9 lineage-specific profile library of BUSCO software (58), while the number of complete duplicated or fragmented BUSCOs remains at 0, indicative of complete assembly. Alignment of all reads show that only 4% (13,028) remained unmapped and 80% of their length were below 2 kbp. The remaining reads (2000 reads) map on various species and could not be further assembled. Overall, these analyses indicate that the final genome assembly is complete and of good quality.

Bioinformatic analyses

The *V. parvula* SKV38 genome was annotated using PROKKA (59). The SKV38 annotated genome sequence was deposited in the ENA (European Nucleotide Archive) under the accession number ERZ1303164.
For protein domain visualization, PFAM domains (pfam.xfam.org, Pfam 32.0. (43)) were detected using HMMER (60). Domains with an e-value lower than $10^{-3}$ were kept and, in case of overlapping domains, the domain having the best e-value was kept. Presence of C-terminal coils structure was determined using the COILS program (https://embnet.vital-it.ch/software/COILS_form.html) (61).

The search for HD phosphatase (YqeK) cluster homologs was conducted as follows: a local databank containing 390 genomes representative of bacterial diversity was mined for the presence of a phosphatase containing HD domain (PF01966) using HMMSEARCH and the --cut_ga option. Protein sequences were then filtered using alignment, functional annotation, protein domains presence and phylogeny. Synteny was investigated in the locus around yqeK by looking, using MacSyFinder (62), for the presence of at least one of the 7 genes surrounding yqeK in V. parvula SKV38, namely obg (containing GTP1_OBG domain, PF01018), yhbY (containing CRS1_YhbY domain, PF01985), proB (containing AA_kinase domain, PF00696), proA (containing Aldedh domain, PF00171), nadD (containing CTP_transf_like domain, PF01467), lytR (containing LytR_cpsA_psr domain, PF03816) and rsfS (containing RsfS domain, PF02410), with no more than eight other genes separating them. All HMM profiles were downloaded from the PFAM site (pfam.xfam.org). As YqeK homologs are widespread in the Firmicutes, another local databank containing 230 representative Firmicutes genomes was queried by the MacSyFinder approach as described above. All trees were visualized with ITOL (63). Details of the results are presented in Dataset S2.

**Strains and growth conditions**

Bacterial strains and plasmids are listed in Table 2. V. parvula was grown in either Brain Heart infusion medium (Bacto Brain Heart infusion, Difco) supplemented with 0.1 % L-cysteine and 0.6 % sodium DL-lactate (BHILC) or SK medium (10 g L$^{-1}$ tryptone (Difco), 10 g L$^{-1}$ yeast extract (Difco), 0.4 g L$^{-1}$ disodium phosphate, 2 g L$^{-1}$ sodium chloride, and 10 mL L$^{-1}$ 60 % w/v sodium DL-lactate, described in (17)) and incubated at 37°C in anaerobic conditions in anaerobic bags (GENbag anaero, Biomerieux, ref. 45534) or in a C400M Ruskinn anaerobic-microaerophilic station. *Escherichia coli* was grown in Lysogeny Broth (LB) (Corning) medium under aerobic conditions at 37°C. 20 mg L$^{-1}$ chloramphenicol (Cm), 200 mg L$^{-1}$ erythromycin (Ery) or 2.5 mg L$^{-1}$ tetracycline (Tc) were added to V. parvula cultures, 100 mg L$^{-1}$ carbenicillin (Cb) or 5 mg L$^{-1}$ tetracycline (Tc) were added to *E. coli* cultures when needed.
100 µg L⁻¹ anhydrotetracycline (aTc) was added to induce the pTet promoter unless stated otherwise. All chemicals were purchased from Sigma-Aldrich unless stated otherwise.

**Natural transformation**

Cells were resuspended in 1 mL SK media adjusted to 0.4-0.8 OD₆₀₀ and 10 µL were dotted on SK-agar petri dishes. On each drop, 0.5-1 µg plasmid or 75-200 ng µL⁻¹ linear dsDNA PCR product was added, or water for the negative control. The plates were incubated 48 hours. The biomass was resuspended in 500 µL SK medium and plated on SK-agar supplemented with the corresponding antibiotic and incubated for another 48 hours. Colonies were streaked on fresh selective plates and the correct integration of the construct was confirmed by PCR and sequencing.

**Random mariner transposon mutagenesis**

Plasmid pRPF215, described for *Clostridium* mutagenesis (Addgene 106377) (19) was transformed into *V. parvula* SKV38 by natural transformation and selected on Cm supplemented SK-agar medium. An overnight culture of *V. parvula* SKV38-pRPF215 in BHILC was then diluted to 0.1 OD₆₀₀ in the same media, supplemented with aTc and grown 5 hours to induce the transposase. After induction, the culture was diluted and plated on BHILC supplemented with Ery and aTc for selection and incubated for 48 hours. From the resulting colonies, 940 were inoculated in Greiner Bio-one polystyrene flat-bottom 96-well plates (655101) and grown in BHILC supplemented with either Ery and aTc, or Cm, to confirm both the presence of the transposon and the loss of pRPF215 and then kept in 15 % glycerol at -80°C. Selected transposon mutants were grown overnight and the genomic DNA was harvested using DNeasy blood and tissue kit (Qiagen). The genomic DNA was then sent for whole genome sequencing at the Mutualized platform for Microbiology of Institut Pasteur.

**Cloning-independent allelic exchange mutagenesis**

Site-directed mutagenesis of *V. parvula* SK38 strain was performed as described by Knapp and colleagues (17). Briefly, 1-Kb regions upstream and downstream the target sequence and *V. atypica* tetracycline resistance cassette (*tetM* in pBSJL2) or catP resistance cassette from *C. difficile* (catP in pRPF185, Addgene 106367, from (64)) were PCR-amplified with overlapping
primers using Phusion Flash High-Fidelity PCR Master-Mix (Thermo Scientific, F548). PCR products were used as templates in a second PCR round using only the external primers that generated a linear dsDNA with the tetracycline resistance cassette flanked by the upstream and downstream sequences. This construct was transformed into *V. parvula* by natural transformation and its integration into the genome was selected by plating on Tc or Cm supplemented medium. Positive candidates were further confirmed by a set of PCRs and sequencing around the site. Primers used in this study are listed in Table S1.

**Complementation**

We replaced the tetracycline resistance gene and its *gyrA* promoter of the shuttle vector pBSJL2 by a chloramphenicol resistance gene, *P*cat:cat from pRPF185 by Gibson assembly. Briefly, the inserts and the plasmids were PCR amplified and then mixed with Gibson master mix 2x (100µL 5X ISO Buffer, 0.2 µL 10,000 U/mL T5 exonuclease (NEB #M0363S), 6.25 µL 2,000 U/mL Phusion HF polymerase (NEB #M0530S), 50 µL 40,000 U/mL Taq DNA ligase (NEB #M0208S), 87 µL dH2O for 24 reactions) and incubated at 50°C for 30-60 min.

The resulting plasmid pBSJL2-cat was digested by Fastdigest *BamHI* (Thermo scientific) and the band obtained was purified from agarose gel using QIAquick gel extraction kit (Qiagen) to be used as linear plasmid in a second Gibson assembly. The genes and the *P*mdh promoter of *V. parvula* SKV38 were amplified by PCR using PhusionFlash Master-mix and cloned in pBSJL2-cat using Gibson assembly. The mix was then transformed in *E. coli* DH5α and plated on LB with carbenicillin. The plasmid was harvested by miniprep using QIAprep spin miniprep kit (Qiagen) and transformed in *V. parvula* as described above.

Alternatively, the anhydrotetracycline inducible expression cassette of pRPF185, hereafter referred to as *pTet*, (Addgene 106367, (64)) was inserted along with a chloramphenicol marker right before the ATG of the target gene, following the procedure described above for cloning-independent allelic exchange mutagenesis. The functionality of *pTet* in *V. parvula* was previously verified using measurement of the aTc dependent β-glucuronidase activity generated by the presence of pRPF185 transformed in *V. parvula* SKV38 (Figure S3).

**Biofilm formation in 96-well microtiter plates**

Overnight cultures in BHILC medium were diluted to 0.05 OD600 and transferred to three Greiner Bio-one polystyrene flat bottom 96-well plates adding 150 µL per well. After 24 hours of static incubation, one of the three plates was resuspended by pipetting to measure OD600
using a TECAN Infinite-M200-Pro spectrophotometer. The two other plates were used for
coloration: cultures were removed by pipetting carefully the supernatant out and biofilms fixed
with 150 µL Bouin solution (HT10132, Sigma-Aldrich) for 15 minutes. Bouin solution was
removed by inversion and the biofilms were washed once in water. The biofilms were stained
with 150 µL crystal violet 1 % (V5265, Sigma-Aldrich) for 15 minutes without shaking, then
washed in water twice and left to dry. All washes were made by flicking of the plate. After
drying the plate, crystal violet was dissolved with 200 µL absolute ethanol and transferred to a

clean 96-well plate for OD₆₂₀ measurement (TECAN Infinite-M200-Pro spectrophotometer).

### Biofilm formation in microfermentor

Continuous flow non-bubbled microfermentor containing a removable spatula were used as
described in (65, 66) (see https://research.pasteur.fr/en/tool/biofilm-microfermenters/). Briefly,
a glass spatula was dipped in an overnight culture diluted to 0.5 OD₆₀₀ in 15 mL BHILC for 15
minutes and returned to the fermentor. Biofilm was grown on the spatula for 48 hours at 37°C.
BHILC was constantly supplied through a peristaltic pump at 4 rpm. During the last hour, the
speed was increased to 10 rpm to remove planktonic bacteria. A mix of filtered 90% nitrogen/5%
hydrogen/5% carbon dioxide was also constantly supplied to maintain anaerobic condition.
After 48 hours of growth, the spatula was removed, and the biofilm was resuspended by
vortexing in 15 mL BHILC. We measured OD₆₀₀ of the resuspended biofilms with Smart Spec
Plus spectrophotometer (BioRad).

### Aggregation curve

Overnight cultures were diluted to 0.8 OD₆₀₀ in Brain-heart infusion (BHI) media in semi-micro
spectrophotometry cuvette (Fisherbrand) and left to sediment on the bench in presence of
oxygen, so no growth should occur. OD₆₀₀ was measured every hour in a single point of the
cuvette using SmartSpec spectrophotometer (BioRad).

### Initial adhesion on glass

Glass spatula from microfermentor (described above) were dipped in overnight cultures diluted
to 0.5 OD₆₀₀ in 15 mL Brain-Heart Infusion (BHI) media for 30 minutes to let bacteria adhere.
The spatulas were washed once in 15 mL BHI by submersion and the adhering bacteria were
resuspended in 15 mL clean BHI by vortexing. The culture used for inoculation, as well as the
resuspended bacteria were serially diluted and plated on SK-agar plate for colony forming unit
(CFU) counting.
Statistical analysis was performed using either R and Rstudio software or GraphPad Prism8 (GraphPad software, Inc.). We used only non-parametric test, and when applicable corrected for multiple testing. For microfermentor experiments, 4 replicates of each condition were used. For all the other experiments, at least 6 biological replicates in at least 2 independent experiment were used. A cut-off of p-value of 5% was used for all tests. * p<0.05; ** p<0.05; *** p<0.005. For growth curve analyses, we computed the growth rate and carrying capacity of each biological replicates using Growthcurver 0.3.0 package in R and we performed a Mann-Whitney test comparing both parameters for each mutant to the corresponding WT.
COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

We thank Justin Merritt for providing *V. parvula* SKV38 strain, Bruno Dupuy and Robert P. Fagan for providing the different *Clostridium* plasmids, Pierre Simon Garcia for help with the Figure 3 preparation, Daniela Megrian Nuñez and Panagiotis Adam for the genome databank preparation and the platforms France Génomique and IBISA. We wish to acknowledge funding from the French National Research Agency (ANR) (Fir-OM ANR-16-CE12-0010), from the Institut Pasteur “Programmes Transversaux de Recherche” (PTR 39-16), from the French government's Investissement d'Avenir Program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant n°ANR-10-LABX-62-IBEID) and from the Fondation pour la Recherche Médicale (grant DEQ20180339185). N.B. was supported by a MENESR fellowship. A.J.F. was supported by a PRESTIGE program from Campus France.

AUTHORS CONTRIBUTIONS

C.B., N.B. and A.J.F. designed the experiments. N.B., A.J.F., E.B. and L.M. performed the experiments. J.W., N.T., and T.C. carried out all genomics and phylogeny analyses under the supervision of SG. N.B., C.B. and A.J.F. wrote the paper, with contribution from S.G., J.W., T.C. and JM.G. All authors have read and approved the manuscript.
REFERENCES

1. Megrian D, Taib N, Witwinowski J, Beloin C, Gribaldo S. 2020. One or two membranes? Diderm Firmicutes challenge the Gram-positive/Gram-negative divide. Mol Microbiol mmi.14469.

2. Veillon A, Zuber A. 1898. Recherches sur quelques microbes strictement anaérobies et leur rôle en pathologie humaine. Arch Med Exp Anat Pathol 10.

3. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu W-H, Lakshmanan A, Wade WG. 2010. The human oral microbiome. J Bacteriol 192:5002–17.

4. van den Bogert B, Erkus O, Boekhorst J, de Goffau M, Smid EJ, Zoetendal EG, Kleerebezem M. 2013. Diversity of human small intestinal Streptococcus and Veillonella populations. FEMS Microbiol Ecol 85:376–388.

5. Aujoulat F, Roudière L, Picaud J-C, Jacquot A, Filleron A, Neveu D, Baum T-P, Marchandin H, Jumas-Bilak E. 2014. Temporal dynamics of the very premature infant gut dominant microbiota. BMC Microbiol 14:325.

6. Arrieta M-C, Stiensma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutch S, Kuzeljevic B, Gold MJ, Britton HM, Lefebvre DL, Subbarao P, Mandhane P, Becker A, McNagny KM, Sears MR, Kollmann T, Investigators the CS, Mohn WW, Turvey SE, Finlay BB. 2015. Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci Transl Med 7:307ra152-307ra152.

7. Kolenbrander PE. 2011. Multispecies communities: Interspecies interactions influence growth on saliva as sole nutritional source, p. 49–54. In International Journal of Oral Science.

8. Periasamy S, Kolenbrander PE. 2010. Central role of the early colonizer Veillonella sp. in establishing multispecies biofilm communities with initial, middle, and late colonizers of enamel. J Bacteriol 192:2965–72.

9. Li J, Chen P, Li J, Gao X, Chen X, Chen J. 2017. A new treatment of sepsis caused by veillonella parvula : A case report and literature review. J Clin Pharm Ther 42:649–652.

10. Hirai J, Yamagishi Y, Kinjo T, Hagihara M, Sakanashi D, Suematsu H, Fujita J, Mikamo H. 2016. Osteomyelitis caused by Veillonella species: Case report and review of the literature. J Infect Chemother 22:417–420.
11. Gouze H, Noussair L, Padovano I, Salomon E, de Laroche M, Duran C, Felter A, Carlier R, Breban M, Dinh A. 2019. Veillonella parvula spondylodiscitis. Médecine Mal Infect 49:54–58.

12. Hyo Y, Fukushima H, Harada T, Hara H. 2019. Nasal septal abscess caused by anaerobic bacteria of oral flora. Auris Nasus Larynx 46:147–150.

13. Wellens L, Casteels I, Huygens M. 2019. Veillonella parvula periorbital cellulitis: an unusual pathogen causing a common clinical sign. GMS Ophthalmol cases 9:Doc17.

14. Gronow S, Welnitz S, Lapidus A, Nolan M, Ivanova N, del Rio TG, Copeland A, Chen F, Tice H, Pitluck S, Cheng JF, Saunders E, Brettin T, Han C, Detter JC, Bruce D, Goodwin L, Land M, Hauser L, Chang YJ, Jeffries CD, Pati A, Mavromatis K, Mikhailova N, Chen A, Palaniapan K, Chain P, Rohde M, Göker M, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Kyrpides NC, Lucas S. 2010. Complete genome sequence of Veillonella parvula type strain (Te3 T). Stand Genomic Sci 2:57–65.

15. Poppleton DI, Duchateau M, Hourdel V, Matondo M, Flechsler J, Klingl A, Beloin C, Gribaldo S. 2017. Outer membrane proteome of Veillonella parvula: A diderm firmicute of the human microbiome. Front Microbiol 8:1–17.

16. Berne C, Ducret A, Hardy GG, Brun Y V. 2015. Adhesins Involved in Attachment to Abiotic Surfaces by Gram-Negative Bacteria. Microbiol Spectr 3.

17. Knapp S, Brodal C, Peterson J, Qi F, Kreth J, Merritt J. 2017. Natural Competence Is Common among Clinical Isolates of Veillonella parvula and Is Useful for Genetic Manipulation of This Key Member of the Oral Microbiome. Front Cell Infect Microbiol 7:1–12.

18. Branda SS, González-Pastor JE, Dervyn E, Ehrlich SD, Losick R, Kolter R. 2004. Genes involved in formation of structured multicellular communities by Bacillus subtilis. J Bacteriol 186:3970–9.

19. Dembek M, Barquist L, Boinett CJ, Cain AK, Mayho M, Lawley TD, Fairweather NF, Fagan RP. 2015. High-Throughput Analysis of Gene Essentiality and Sporulation in Clostridium difficile. MBio 6:e02383.

20. Łyskowski A, Leo JC, Goldman A. 2011. Structure and Biology of Trimeric Autotransporter Adhesins, p. 143–158. In Advances in experimental medicine and
21. Ruiz-Ranwez V, Posadas DM, Estein SM, Abdian PL, Martin FA, Zorreguieta A. 2013. The BtaF trimeric autotransporter of Brucella suis is involved in attachment to various surfaces, resistance to serum and virulence. PLoS One 8:e79770.

22. Zhou P, Liu J, Merritt J, Qi F. 2015. A YadA-like autotransporter, Hag1 in Veillonella atypica is a multivalent hemagglutinin involved in adherence to oral streptococci, Porphyromonas gingivalis, and human oral buccal cells. Mol Oral Microbiol 30:269–279.

23. Trunk T, S. Khalil H, C. Leo J. 2018. Bacterial autoaggregation. AIMS Microbiol 4:140–164.

24. Bongrand C, Ruby EG. 2019. The impact of Vibrio fischeri strain variation on host colonization. Curr Opin Microbiol 50:15–19.

25. Bonazzi D, Lo Schiavo V, Machata S, Djafer-Cherif I, Nivoit P, Manriquez V, Tanimoto H, Husson J, Henry N, Chaté H, Voituriez R, Duménil G. 2018. Intermittent Pili-Mediated Forces Fluidize Neisseria meningitidis Aggregates Promoting Vascular Colonization. Cell 174:143-155.e16.

26. Klemm P, Vejborg RM, Sherlock O. 2006. Self-associating autotransporters, SAATs: Functional and structural similarities. Int J Med Microbiol 296:187–195.

27. Ageorges V, Schiavone M, Jubelin G, Caccia N, Ruiz P, Chafsey I, Bailly X, Dague E, Leroy S, Paxman J, Heras B, Chaucyeras-Durand F, Rossiter AE, Henderson IR, Desvaux M. 2019. Differential homotypic and heterotypic interactions of antigen 43 (Ag43) variants in autotransporter-mediated bacterial autoaggregation. Sci Rep 9:11100.

28. Wells TJ, Sherlock O, Rivas L, Mahajan A, Beatson SA, Torpdahl M, Webb RI, Allsopp LP, Gobius KS, Gally DL, Schembri MA. 2008. EhaA is a novel autotransporter protein of enterohemorrhagic Escherichia coli O157:H7 that contributes to adhesion and biofilm formation. Environ Microbiol 10:589–604.

29. Totsika M, Wells TJ, Beloin C, Valle J, Allsopp LP, King NP, Ghigo J-M, Schembri MA. 2012. Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic Escherichia coli. Appl Environ Microbiol 78:2179–89.

30. Valle J, Mabbett AN, Ulett GC, Toledo-Arana A, Wecker K, Totsika M, Schembri
MA, Ghigo J-M, Beloin C. 2008. UpaG, a New Member of the Trimeric Autotransporter Family of Adhesins in Uropathogenic Escherichia coli. J Bacteriol 190:4147–4161.

Paton AW, Srimanote P, Woodrow MC, Paton JC. 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic Escherichia coli strains that are virulent for humans. Infect Immun 69:6999–7009.

Leo JC, Lyskowski A, Hattula K, Hartmann MD, Schwarz H, Butcher SJ, Linke D, Lupas AN, Goldman A. 2011. The Structure of E. coli IgG-Binding Protein D Suggests a General Model for Bending and Binding in Trimeric Autotransporter Adhesins. Structure 19:1021–1030.

Beloin C, Michaelis K, Lindner K, Landini P, Hacker J, Ghigo J-M, Dobrindt U. 2006. The transcriptional antiterminator RfaH represses biofilm formation in Escherichia coli. J Bacteriol 188:1316–31.

Korea C-G, Badouraly R, Prevost M-C, Ghigo J-M, Beloin C. 2010. Escherichia coli K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. Environ Microbiol 12:1957–1977.

Hasman H, Chakraborty T, Klemm P. 1999. Antigen-43-mediated autoaggregation of Escherichia coli is blocked by fimbriation. J Bacteriol 181:4834–41.

Schembri MA, Dalsgaard D, Klemm P. 2004. Capsule shields the function of short bacterial adhesins. J Bacteriol 186:1249–57.

Hughes C V, Andersen RN, Kolenbrander PE. 1992. Characterization of Veillonella atypica PK1910 adhesin-mediated coaggregation with oral Streptococcus spp. Infect Immun 60:1178–86.

Mashima I, Nakazawa F. 2015. Interaction between Streptococcus spp. and Veillonella tobetsuensis in the early stages of oral biofilm formation. J Bacteriol. American Society for Microbiology.

Mashima I, Nakazawa F. 2014. The influence of oral Veillonella species on biofilms formed by Streptococcus species. Anaerobe 28:54–61.

Periasamy S, Kolenbrander PE. 2010. Central role of the early colonizer Veillonella sp. in establishing multispecies biofilm communities with initial, middle, and late...
colonizers of enamel. J Bacteriol 192:2965–2972.

41. Hughes C V, Kolenbrander PE, Andersen RN, Moore L V. 1988. Coaggregation properties of human oral Veillonella spp.: relationship to colonization site and oral ecology. Appl Environ Microbiol 54:1957–63.

42. Wells TJ, Totsika M, Schembri MA. 2010. Autotransporters of Escherichia coli: A sequence-based characterization. Microbiology 156:2459–2469.

43. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47.

44. Mesnagge S, Fontaine T, Mignot T, Delepierre M, Mock M, Fouet A. 2000. Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. EMBO J 19:4473–4484.

45. Park JS, Lee WC, Yeo KJ, Ryu K-S, Kumarasiri M, Hesek D, Lee M, Mobashery S, Song JH, Kim S II, Lee JC, Cheong C, Jeon YH, Kim H-Y. 2012. Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the gram-negative bacterial outer membrane. FASEB J 26:219–228.

46. Minazzato G, Gasparrini M, Amici A, Cianci M, Mazzola F, Orsomando G, Sorci L, Raffaelli N. 2020. Functional characterization of COG1713 (YqeK) as a novel diadenosine tetraphosphate hydrolase family. J Bacteriol JB.00053-20.

47. Monds RD, Newell PD, Wagner JC, Schwartzman JA, Lu W, Rabinowitz JD, O’Toole GA. 2010. Di-adenosine tetraphosphate (Ap4A) metabolism impacts biofilm formation by Pseudomonas fluorescens via modulation of c-di-GMP-dependent pathways. J Bacteriol 192:3011–3023.

48. Gundlach J, Rath H, Herzberg C, Mäder U, Stülke J. 2016. Second Messenger Signaling in Bacillus subtilis: Accumulation of Cyclic di-AMP Inhibits Biofilm Formation. Front Microbiol 7:804.

49. Peng X, Zhang Y, Bai G, Zhou X, Wu H. 2016. Cyclic di-AMP mediates biofilm formation. Mol Microbiol 99:945–959.

50. Townsley L, Yannarell SM, Huynh TN, Woodward JJ, Shank EA. 2018. Cyclic di-AMP Acts as an Extracellular Signal That Impacts Bacillus subtilis Biofilm Formation.
51. Jenal U, Reinders A, Lori C. 2017. Cyclic di-GMP: second messenger extraordinaire. Nat Rev Microbiol 15:271–284.

52. Krasteva PV, Sondermann H. 2017. Versatile modes of cellular regulation via cyclic dinucleotides. Nat Chem Biol 13:350–359.

53. Yan F, Yu Y, Wang L, Luo Y, Guo J-H, Chai Y. 2016. The comER Gene Plays an Important Role in Biofilm Formation and Sporulation in both Bacillus subtilis and Bacillus cereus. Front Microbiol 7:1025.

54. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736.

55. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakhikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. PLoS One 9:e112963.

56. Cokelaer T, Desvillechabrol D, Legendre R, Cardon M, Cokelaer T, Desvillechabrol D, Legendre R, Cardon M. 2017. “Sequana”: a Set of Snakemake NGS pipelines. J Open Source Softw 2:352.

57. Desvillechabrol D, Bouchier C, Kennedy S, Cokelaer T. 2018. Sequana Coverage: Detection and Characterization of Genomic Variations using Running Median and Mixture Models. Gigascience 7.

58. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212.

59. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

60. Johnson LS, Eddy SR, Portugaly E. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics 11:431.

61. Lupas A, Dyke M Van, Stock J. 1991. Predicting Coiled Coils from Protein Sequences. Science (80- ). American Association for the Advancement of Science.
62. Abby SS, Néron B, Ménager H, Touchon M, Rocha EPC. 2014. MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. PLoS One 9:e110726.

63. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259.

64. Fagan RP, Fairweather NF. 2011. Clostridium difficile Has Two Parallel and Essential Sec Secretion Systems. J Biol Chem 286:27483–27493.

65. Ghigo J-M. 2001. Natural conjugative plasmids induce bacterial biofilm development. Nature 412:442–445.

66. Beloin C, Valle J, Latour-Lambert P, Faure P, Kzreminski M, Balestrino D, Haagensen JAJ, Molin S, Premsier G, Arbeille B, Ghigo J-M. 2003. Global impact of mature biofilm lifestyle on Escherichia coli K-12 gene expression. Mol Microbiol 51:659–674.

67. Soutourina OA, Monot M, Boudry P, Saujet L, Pichon C, Sismeiro O, Semenova E, Severinov K, Le Bouguenecc C, Coppée JY, Dupuy B, Martin-Verstraete I. 2013. Genome-Wide Identification of Regulatory RNAs in the Human Pathogen Clostridium difficile. PLoS Genet 9.

68. Liu J, Xie Z, Merritt J, Qi F. 2012. Establishment of a tractable genetic transformation system in Veillonella spp. Appl Environ Microbiol 78:3488–91.

69. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27:1009–1010.
| Locus tag       | Gene name | Genome position | Gene size (kb) | Strand | Description                                           | DSM2008 homolog | Name | Class |
|-----------------|-----------|-----------------|----------------|--------|-------------------------------------------------------|----------------|------|-------|
| FNLLGLLA_00032 | prn 1     | 39,354          | 2,37           | forward| Autotransporter                                      | fusion Vpar_0036-0037 | VmaA | Va    |
| FNLLGLLA_00034 | btaE 1    | 42,345          | 1,41           | reverse| Trimeric Autotransporter: YadA like                   | Vpar_0039       | VnaB | Vc    |
| FNLLGLLA_00035 | hypothetical protein | 44,146          | 1,04           | forward| Autotransporter (partial)                            | Vpar_0040       | Vc   |       |
| FNLLGLLA_00036 | hypothetical protein | 45,453          | 1,431          | forward| none                                                  | split Vpar_0041  | ?    |       |
| FNLLGLLA_00037 | omp alpha | 46,91           | 969            | forward| Trimeric Autotransporter/ S-layer homology domain    | split Vpar_0041  | Vc   |       |
| FNLLGLLA_00038 | spaG 1    | 48,397          | 8,433          | forward| Trimeric Autotransporter: YadA like                   | Vpar_0042       | VnaC | Vc    |
| FNLLGLLA_00040 | btaE 2    | 57,966          | 1,875          | forward| Trimeric Autotransporter: YadA like (Partial)        | split Vpar_0048  | ?    |       |
| FNLLGLLA_00041 | ata 1     | 59,837          | 3,627          | forward| Trimeric Autotransporter: YadA like                   | split Vpar_0048  | VvA  | Vc    |
| FNLLGLLA_00044 | vfrG 1    | 65,3            | 6,216          | forward| Trimeric Autotransporter: YadA like                   | Vpar_0051       | VnaD | Vc    |
| FNLLGLLA_00045 | spaG 2    | 71,995          | 9,426          | forward| Trimeric Autotransporter: YadA like                   | Vpar_0052       | VnaE | Vc    |
| FNLLGLLA_00046 | ata 2     | 81,941          | 9,579          | forward| Trimeric Autotransporter: YadA like                   | Vpar_0053       | VnaF | Vc    |
| FNLLGLLA_00098 | btaE 3    | 151,592         | 1,731          | forward| Trimeric Autotransporter/ S-layer homology domain    | Vpar_0100       | VnaG | Vc    |
| FNLLGLLA_00099 | ata 3     | 154,024         | 4,959          | forward| Trimeric Autotransporter/ S-layer homology domain    | absent          | VnaH | Vc    |
| FNLLGLLA_00335 | prn 2     | 414,666         | 2,223          | forward| Autotransporter                                      | Vpar_0330       | VmaB | Vc    |
| FNLLGLLA_00516 | btaF      | 581,236         | 9,123          | forward| Trimeric Autotransporter: YadA like                   | Vpar_0464       | VnaA | Vc    |
| FNLLGLLA_00581 | brkA      | 668,34          | 2,244          | forward| Autotransporter                                      | Vpar_1322       | VnaC | Vc    |
| FNLLGLLA_01790 | vfrG 2    | 1,943,661       | 2,499          | reverse| Trimeric Autotransporter/ S-layer homology domain    | Vpar_1664       | Vna  | Vc    |

Table 1: *V. parvula* SKV38 autotransporters
| Strain name | Description | Reference |
|-------------|-------------|-----------|
| WT          | *Veillonella parvula* SKV38 | (17) |
| 9G5         | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 5C5         | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 5H1         | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 3D6         | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 7B11        | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 2F7         | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 3F7         | *Veillonella parvula* SKV38 FNLLGLLA_00516::Transposon | This study |
| 5E11        | *Veillonella parvula* SKV38 FNLLGLLA_01127::Transposon | This study |
| ΔvtaA       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00516::tetM | This study |
| pTet-vtaA   | *Veillonella parvula* SKV38 catP-Term(fdX)-Ptet-FNLLGLLA_00516 | This study |
| Δ8          | *Veillonella parvula* SKV38 ΔFNLLGLLA_00036-46::tetM | This study |
| ΔvmaA       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00032::tetM | This study |
| ΔvtaB       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00034::tetM | This study |
| ΔvtaG       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00098::tetM | This study |
| ΔvtaH       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00099::tetM | This study |
| ΔvmaB       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00335::tetM | This study |
| ΔvmaC       | *Veillonella parvula* SKV38 ΔFNLLGLLA_00581::tetM | This study |
| ΔvtaI       | *Veillonella parvula* SKV38 ΔFNLLGLLA_01790::tetM | This study |
| ΔvtaAΔ8     | *Veillonella parvula* SKV38 ΔFNLLGLLA_00516::catP ΔFNLLGLLA_00036-46::tetM | This study |
| Δ1127       | *Veillonella parvula* SKV38 ΔFNLLGLLA_01127::tetM | This study |
| Δ1127Δ8     | *Veillonella parvula* SKV38 ΔFNLLGLLA_01127::tetM ΔFNLLGLLA_00036-46::catP | This study |
| WT+pEmpty   | *Veillonella parvula* SKV38-pBSJL2-catP-pmdh | This study |
| Δ1127+pEmpty | *Veillonella parvula* SKV38 ΔFNLLGLLA_01127::tetM-pBSJL2-catP-pmdh | This study |
| Δ1127+p1127 | *Veillonella parvula* SKV38 ΔFNLLGLLA_01127::tetM-pBSJL2-catP-pmdh-FNLLGLLA_01127 | This study |
| P_{ter-φ}   | SKV38- pRPF185ΔgusA | This study |
| P_{ter-gusA}| SKV38- pRPF185 | This study |
| PCwp2 -gusA | SKV38- pRPF144 | This study |

| Plasmid | Description | Reference |
|---------|-------------|-----------|
|         |             |           |
|   | Description                                                                 | Reference |
|---|-----------------------------------------------------------------------------|-----------|
| pRPF215 | mariner Tn delivery plasmid, \( P_{tet}: \text{Himar1, ITR-ermB-ITR, catP, tetR} \) | (19)      |
| pRPF185 | tetracycline inducible expression system fused with \( \beta\)-glucuronidase \( gusA \) Term(fdx)-\( P_{tet} \)-gusA-Term(slpA), \( catP \) | (64)      |
| pRPF185\( \Delta \)gusA | pDIA6103, tetracycline inducible expression system Term(fdx)-\( P_{tet} \)-Term(slpA), \( catP \) | (67)      |
| pRPF144 | carries a \( \text{Clostridium} \) constitutive promoter fused with \( gusA \) \( P_{Cwp2-gusA} \) | (64)      |
| pBSJL2  | \( \text{E. coli-Veillonella} \) shuttle plasmid, \( P_{gyrA}: \text{tetM} \) | (68)      |
| pBSJL2\( \cdot \)cat | \( \text{E. coli-Veillonella} \) shuttle plasmid, \( P_{cat}: \text{catP} \), \( pmdh \) promoter | This study |
| p1127   | \( \text{pBSJL2}\cdot\text{catP}\cdot pmdh\cdot FNLLGLLA_01127 \) | This study |

Table 2: Strains and plasmids used in this study
FIGURES LEGEND

Figure 1: Random transposon mutagenesis in *Veillonella parvula* SKV38 led to identification of mutants with reduced biofilm formation. A. 96-well polystyrene plate biofilm assay after CV staining of nine transposon mutants identified by random mutagenesis grown 24h in BHILC. Mean of WT is adjusted to 100 %. Min-max boxplot of 6-15 biological replicates for each strain are represented, each replicate is the mean of two technical replicates. * p-value<0.05, ** p-value <0.005, Mann-Whitney test. B. Schematic representation of the transposon insertion point identified (red arrow) for the 8 transposon mutants. Blue bar represents the size of the gene in nucleotides.

Figure 2: VtaA is an adhesin involved in auto-aggregation and biofilm formation. A. 96-well plate biofilm assay after 24h growth in BHILC. Mean of WT is adjusted to 100 %. Min-max boxplot of 6 biological replicates for each strain. * p-value<0.05, ** p-value <0.005, Mann-Whitney test between strains. B. and C. Aggregation curve in spectrophotometry cuvette of WT and ΔvtaA (B) and of an inducible vtaA with 0, 0.025 or 0.1 µg/mL of the inducer aTc (C). 100 % represent lack of aggregation, 0 % complete sedimentation of the culture. Median of 6 biological replicates, error bars represent 95% confidence interval. At each time points we computed the Mann-Whitney test between conditions. We applied Bonferroni correction for multiple testing: p-value are only considered significant if *p-value<0.004, **p-value<0.0004, ***p-value<0.00004. Indicated p-values were calculated by comparing in B, WT and ΔvtaA, and in C, pTet-vtaA without aTc and pTet-vtaA with different aTc concentrations. D. 96-well plate biofilm assay after 24h growth of an inducible vtaA in BHILC with different concentrations of aTc. WT without aTc is adjusted to 100 %. Median of 6 biological replicates, each replicate corresponds to the mean of two technical replicates, error bars represent 95% confidence interval. * p-value<0.05, ** p-value <0.005, Mann-Whitney test. E. Biofilm formation in continuous flow microfermentor on glass spatula during 48h in BHILC. WT was adjusted to 100 %. Min-max boxplot of 4 biological replicates for each strain. A picture of the spatula before resuspension is shown below each boxplot bar. * p-value<0.05, Mann-Whitney test.

Figure 3: *Veillonella parvula* autotransporters domain organization. A. Genetic organization of the *V. parvula* SKV38 autotransporter adhesin gene cluster and the corresponding adhesin domain organization. B. Domain organization of the six remaining *V. parvula* SKV38 autotransporter adhesins encoded by genes located outside of the cluster. Domains were detected with the HMMER package (60), only the domains with e-values lower than 10⁻³ are shown. Presence of C-terminal coils structure was determined using the COILS program (https://embnet.vital-it.ch/software/COILS_form.html). All *V. parvula* trimeric ATs display an additional C-terminal domain (a SLH or a coiled coil domain) following the YadA anchor domain as compared to classical trimeric autotransporters.

Figure 4: Synteny of the adhesin gene cluster in a selection of *Veillonella* species. The synteny of the proteins of the cluster between the closest relatives was assessed using EasyFig (69). Oblique lines between genes represent tblastx identities (program parameters: maximum e-value of 10¹², minimum length of 30, minimum identity of 30). The *V. parvula* SKV38 strain...
Figure 5: A cluster of eight trimeric autotransporters is involved in surface binding. A. Aggregation curve in spectrophotometry cuvette. 100 % represent lack of aggregation, 0 % complete sedimentation of the culture. Median of 6 biological replicates, error bars represent 95% confidence interval. * Mann-Whitney test, corrected for multiple testing with Bonferroni correction: significance is achieved if p-value < 0.007. B. and C. 96-well plate biofilm assay after 24 h growth in BHILC. Mean of WT is adjusted to 100 %. Min-max boxplot of 6 biological replicates for each strain, each replicate is the mean of two technical replicates. In B, we applied a Mann-Whitney; * p-value<0.05, ** p-value <0.005. In C, we applied Bonferroni correction for multiple testing: tests were called significant only if p-value<0.01; * p-value<0.01, ** p-value <0.001, *** p-value <0.0001. D. Biofilm formation in continuous flow microfermentor on glass spatula during 48h in BHILC. WT was adjusted to 100 %. Min-max boxplot of 4 biological replicates for each strain. * p-value<0.05, Mann-Whitney test. A picture of spatula before resuspension is shown for each mutant below the boxplot. E. Initial adhesion on glass spatula. Percentage of CFU that adhered to the spatula controlled by the number of CFU of the inoculation solution. Min-max boxplot of 6-9 replicates per strain is represented. * p-value<0.05, ** p-value <0.005, *** p-value <0.0005, Mann-Whitney test.

Figure 6: Occurrence and synteny of HD Phosphatase (YqeK) in diderm and monoderm bacteria. A. The presence of the cluster was investigated using MacSyFinder (62) and the results were plotted onto a schematic reference tree of 187 cultivable bacteria among the 390 of the analyzed databank. The cell wall status of each phylum is indicated as: (-) diderm with LPS, (+) monoderm, (atyp.) diderm without LPS, (?) unclear. For the Firmicutes, the diderm lineages are indicated in red (Negativicutes), blue (Halanaerobiales) and purple (Limnochordales).

Figure 7: FNLLGLLA_01127 represses biofilm formation in microfermentor. A. 96-well plate biofilm assay after 24 h growth in BHLC corrected by OD600 after 24 h growth in plate. Mean of WT is adjusted to 100 %. Min-max boxplot of 6 biological replicates for each strain, each replicate is the mean of two technical replicates. * p-value < 0.05, Mann-Whitney test. B. Aggregation curve in spectrophotometry cuvette. 100 % represent lack of aggregation, 0 % complete sedimentation of the culture. Median of 6 biological replicates, error bars represent 95% confidence interval. * Mann-Whitney test, corrected for multiple testing with Bonferroni correction: significance is achieved if p-value<0.007. C. Biofilm formation in continuous flow microfermentor on glass spatula during 48h in BHILC. Mean of WT is adjusted to 100 %. Min-max boxplot of 4 biological replicates for each strain. * p-value < 0.05, ** p-value<0.005, *** p-value <0.0001, Mann-Whitney test. A picture of a spatula before resuspension is shown for each strain below the histogram. D. Biofilm formation in continuous flow microfermentor on glass spatula during 48h in BHILC+chloramphenicol. Mean of WT+pEmpty is adjusted to 100 %. Min-max boxplot of 4 biological replicates for each strain. * p-value < 0.05, Mann-Whitney test. A picture of a spatula before resuspension is shown for each strain below the boxplot. E. Scanning electronic microscopy of ∆1127 biofilm grown under continuous flow of BHILC in microfermentor on a plastic microscopy slide. Magnification 2K and 5K. F. Initial adhesion on glass spatula.

Percentage of CFU that adhered to the spatula in 30 min controlled by the number of CFU of
the inoculation solution. Min-max boxplot of 6-9 replicates per strain. * p-value<0.05, Mann-Whitney test.
### A

| Gene locus     | Protein name | Domain organization |
|----------------|--------------|---------------------|
| FNLLGLA_00032 | Pmr_1/VmaA   | 790 aa              |
| FNLLGLA_00034 | BtaE_1/VtaB  | 470 aa              |
| FNLLGLA_00035 | Ata_1        | 348 aa              |
| FNLLGLA_00036 | Ata_2        | 477 aa              |
| FNLLGLA_00037 | UpaG_1/VtaC  | 323 aa              |
| FNLLGLA_00040 | BtaE_2       | 2811 aa             |
| FNLLGLA_00041 | Ata_1        | 625 aa              |
| FNLLGLA_00044 | EhaG_1/VtaD  | 1209 aa             |
| FNLLGLA_00045 | UpaG_2/VtaE  | 2072 aa             |
| FNLLGLA_00046 | Ata_2/VtaF   | 3142 aa             |

### B

| Gene locus     | Protein name | Domain organization |
|----------------|--------------|---------------------|
| FNLLGLA_00098 | BtaE_3/VtaG  | 3193 aa             |
| FNLLGLA_00099 | Ata_3/VtaH   | 1653 aa             |
| FNLLGLA_00335 | Pmr_2/VmaB   | 741 aa              |
| FNLLGLA_00516 | BtaF/VtaA    | 577 aa              |
| FNLLGLA_00581 | BrkA/VmaC    | 3041 aa             |
| FNLLGLA_01790 | EhaG_2/VtaI  | 748 aa              |

- **extended signal peptide region (PF13018)**
- **procytic acidic repetitive protein (PF05887)**
- **DUF1564 (PF07889)**
- **YadA head (PF05658)**
- **YadA stalk (PF05662)**
- **autotransporter (PF03797)**
- **inverse autotransporter beta domain (PF11924)**
- **YadA anchor (PF08395)**
- **S-layer homology (PF00395)**
- **Coil (Coils prediction)**
