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

Nathalie Béchon[1,2,#], Alicia Jiménez-Fernández[1,#], Jerzy Witwinowski[3], Emilie Bierque[1,4,7], Najwa Taib[3,5], Thomas Cokelaer[5,6], Laurence Ma[6], Jean-Marc Ghigo[1], Simonetta Gribaldo[3] and Christophe Beloin[1,*]

[1] Genetics of Biofilm Laboratory, Institut Pasteur, UMR CNRS2001, Paris, 75015, France

[2] Université de Paris, Sorbonne Paris Cité, Paris, France

[3] Evolution 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] Biomics, C2RT, Institut Pasteur, Paris, FrancePlate-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.

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, which compete with each other for binding either to cells or to surfaces, with strong consequences on *V. parvula* biofilm formation capacity. 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.

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 rather currently undescribed adhesion factors. We previously 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 found that the main *V. parvula* biofilm modulating determinants are T5SS adhesins, i.e. typical diderm

76 determinants. We also showed that a locus encoding a metal-dependent phosphohydrolase HD
77 domain protein is involved in biofilm formation, similarly to what was shown in the
78 prototypical monoderm *Bacillus subtilis* (18). Therefore, our results demonstrate that diderm
79 Firmicutes use a mixture of diderm/monoderm factors to modulate their ability to engage into
80 biofilm lifestyle, supporting the idea that monoderm and diderm molecular systems could have
81 co-evolved in these atypical Firmicutes.

82

RESULTS

PacBio sequencing of the genetically amenable *V. parvula* SKV38 strain

We sequenced the *V. parvula* SKV38 whole genome using PacBio sequencing and obtained 338,310 subreads with a mean length of 9,080 bp (Figure S1A). Assembly was performed with Canu and no gaps or drops of coverage was detected based on sequana_coverage output (Figure S1B) (19). One contig of 2,146,482 bp was generated after assembly closely matching the genome size (2.1422 Mbp) and GC content (38.7%, expected 38.6%) of the reference *V. parvula* DSM2008 strain (see supplementary material and methods for details). PROKKA annotation of the *V. parvula* SKV38 genome detected 1,912 predicted protein-encoding open reading frame (ORF), 12 rRNA, 49 tRNA and one tmRNA. Rapid Annotation using Subsystem Technology (RAST) analysis assigned 53% of the predicted proteins to known subsystems (Figure 1A). The two-way average nucleotide identity (ANI) between SKV38 and the reference strain DSM2008 is 95.43% (Figure 1B). A maximum likelihood tree of concatenated RpoB-RpoC-InfB sequences situated SKV38 strain among the Negativicutes, within the *V. parvula* clade (Figure 2).

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

In order to identify genes involved in biofilm formation, 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 *Clostridium difficile* (20), 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 3A). Whole genome sequencing localized the transposons in two loci putatively implicated in biofilm formation (Figure 3B). 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 3B).

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 other bacteria (21). *FNLLGLLA_00516* is a homolog of *V. parvula* DSM2008 *vpar_0464,* which was shown to encode a protein 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 (22). 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 Δ*vtaA* had no growth defect (Figure S2A) but displayed a marked reduction of biofilm formation in microtiter plate (Figure 4A). Moreover, while *V. parvula* SKV38 cultures strongly aggregated, Δ*vtaA* did not (Figure 4B and S3). We constructed the strain *pTet-vtaA*, where the chromosomal *vtaA* gene is placed under the control of a functional tetracycline/anhydrotetracycline (aTc) inducible promoter (Figure S4) and showed that its aggregation capacity and biofilm formation directly correlated with aTc concentration (Figure 4C-D), demonstrating that VtaA-mediated cell-to-cell interactions are critical for biofilm formation. Whereas the microtiter plate assay corresponds to a static biofilm assay, we also used continuous flow biofilm microfermentors to investigate the contribution of VtaA to biofilm formation in dynamic conditions. Surprisingly, Δ*vtaA* formed almost six times more biofilm than the WT strain in these conditions (Figure 4E). This suggests that auto-aggregation differentially contributes to biofilm formation in dynamic or static conditions.

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

The strong biofilm phenotype displayed by the Δ*vtaA* 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 ß-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 ß-domain but no putative α-domain that normally carries the function of the protein, and *FNLLGLLA_00036-37* and

147 *FNLLGLLA_00040-41*, which are homologs of *V. parvula* DSM2008 *Vpar_0041* and
148 *Vpar_0048*, respectively, and that appear to be split in SKV38 (Table 1). Among those, six
149 autotransporters plus *FNLLGLLA_00035, FNLLGLLA_00036-37* and *FNLLGLLA_00040-41*
150 form a potential genomic cluster coding for adhesins (Figure 1B and 5A), whereas the six others
151 are located in different areas of the genome (Figure 1B and Figure 5B).

152

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

155 To assess the function of the potential adhesins identified in the *V. parvula* SKV38
156 genome, we constructed -within the cluster of adhesin genes- independent deletion mutants for
157 the two first autotransporters (*vmaA* and *vtaB*) and a large deletion for the eight adjacent genes
158 encoding trimeric autotransporters or partial trimeric autotransporters, hereafter called Δ*8*
159 *(*Δ*FNLLGLLA_00036* to *vtaF*). We also generated independent individual mutants for each of
160 the six additional autotransporters located outside of the cluster. These mutants were all tested
161 for biofilm formation and aggregation capacities. While the previously mentioned Δ*vtaA* strain
162 was the only mutant involved in cell-to-cell interactions (Figure 6A), both Δ*vtaA* and Δ*8* led to
163 lower biofilm formation in microtiter plate (Figure 6B and C), suggesting that the adhesins of
164 this cluster could be involved in biofilm formation independently of cell-to-cell interactions.
165 However, we observed no significant difference with the WT when evaluating biofilm capacity
166 of the Δ*8* mutant in microfermentor (Figure 6D). To determine whether this was due to flow or
167 the nature of adhesion surfaces (plastic in microtiter plate *vs.* glass in microfermentor), we used
168 plastic microscopy coverslip slides to grow biofilms in microfermentor. Scanning electronic
169 microscopy (SEM) showed that the WT formed a biofilm displaying filaments and protein
170 deposits that could be part of *V. parvula* extracellular matrix, whereas Δ*vtaA* formed a much
171 thicker biofilm, although without filaments and proteins (Figure 6E). Δ*8*, however, only poorly
172 covered the plastic coverslip with sporadic aggregates of cells producing extracellular matrix,
173 consistent with the reduced biofilm formation observed in microtiter plate (Figure 6E). Initial
174 adhesion assay to glass spatula showed that both *vtaA* and Δ*8* displayed a lower percentage of
175 initial adhesion than WT, suggesting that VtaA-mediated auto-aggregation contributed to initial
176 adhesion while the adhesin cluster is directly involved in surface binding (Figure 6F). These
177 two contributions were additive since a Δ*vtaA*Δ*8* double mutant showed a reduced initial
178 adhesion on microfermentor spatula compared to either WT, Δ*vtaA* or Δ*8* (Figure 6F). In
179 addition, Δ*vtaA*Δ*8* formed 17 times less biofilm than Δ*vtaA* 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 6D).

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.

Prevalence of large adhesin clusters in *Veillonellaceae* genomes

The involvement of T5SS adhesins in the modulation of *V. parvula* biofilm formation prompted us to broaden our analysis and determine the prevalence of these adhesins in the *Veillonellaceae*. We plotted the results of a hidden Markov Model (HMM) search for four major types of bacterial domains found in T5SS adhesins, namely autotransporter (PF03797) typical of type Va transport, ShlB (PF03865) typical of two-partner system Vb, YadA_anchor (PF03895) typical of trimeric autotransporter type Vc transport, and iAT-ß typical of inverted autotransporter Ve (PF11924) (Figure 2). Trimeric autotransporters Vc are the most abundant potential adhesins in *Veillonellaceae,* followed by Va classical autotransporters, while type Vb two-partner system and Ve inverted autotransporters are only sporadic. Despite this trend, the numbers of each of these classes varied strongly within the same genus or even the same species of *Veillonellaceae*, indicating an important plasticity of the autotransporters repertoire.

Large clustering of potential adhesin encoding genes observed in SKV38 strain and defined as a group of two or more adhesin-coding genes immediately upstream of a conserved rRNA locus, is to our knowledge a peculiar genomic character. We found no evidence of the existence of this specific adhesin locus outside the *Veillonella* genus. 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 (Figure 7). 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.

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 2). The *FNLLGLLA_01127* gene is homologous to YqeK, a putative phosphatase required for pellicle formation and the development of biofilm in *B. subtilis* (18). A *FNLLGLLA_01127* deletion mutant (Δ*1127*) showed a moderate growth defect (Figure S3AB) and a moderate 1.5-fold decrease in biofilm formation in microtiter plate after correcting for the growth defect (Figure 8A). This mutant also displayed a slightly faster aggregation rate than the WT during early time points (Figure 8B). The strongest phenotype of this mutant was detected in microfermentor with a 9-fold increase in biofilm formation compared to WT (Figure 8C), which was reduced by expressing *FNLLGLLA_01127* gene *in trans* (plasmid *p1127*) (Figure 8D). Scanning electronic microscopy showed that Δ*1127*, similarly to Δ*vtaA*, formed a thick layered biofilm, although with fewer filaments and protein deposits compared to WT (Figure 8E). However, contrary to Δ*vtaA* or Δ*8* mutants, Δ*1127* showed no defect in initial adhesion to a glass spatula (Figure 8F). Interestingly, a Δ*1127*Δ*8* double mutant formed almost 20 times less biofilm than Δ*1127* in microfermentor (Figure 8C), suggesting that at least some of the autotransporters of the cluster were necessary for Δ*1127* 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. Although biofilm formation and adhesion are likely important in these niches, molecular studies in *Veillonella* have 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 (23), VtaA represents the first *Veillonella* protein involved in abiotic surface adhesion and auto-aggregation in diderm Firmicutes. Beyond the potential impact on *Veillonella* niche colonization, aggregation capacity contributes to bacterial protection from environmental stresses or host responses (24), promotion of host colonization (25), or pathogenesis (26). 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 (22), 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 (27–33). 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. We hypothesize that under continuous flow, VtaA-mediated aggregates are more sensitive to flow than individual cells and that adhesion to surfaces or to the biofilm extracellular matrix might be more important than cell-to-cell interactions.

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 (34–37). 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, Δ*vtaA* 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 diderm adhesins. None of them appeared to be involved in cell-to-cell interactions or biofilm formation on abiotic surfaces, and their function remains unclear. As *V. parvula* is present in different microbiota, it is expected that a large arsenal of adhesion factors might be 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* (38–40), that produce *Veillonella* favored carbon source, lactate (41), 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 (42). *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 (43), but we detected no classical protease domain in the *Veillonella* autotransporters.

Although we showed that genes encoding autotransporters are widespread in *Veillonellaceae* genomes, even closely related strains display different numbers, sizes and types of adhesins. For instance, *V. parvula* DSM2008 encodes three proteins with filamentous haemagglutinin domain (PF13332), whereas this class of proteins is completely absent in the closely related SKV38 genome. This suggests rapid evolutionary changes in the repertoire *Veillonella* adhesins, potentially eased by the presence of an adhesin cluster that facilitates recombination events between adjacent homologous adhesins and reduces the odds of inactivating genes involved in important cellular functions by duplication, deletion and recombination. Such a diversification might be selected for when the environmental pressure imposes to constantly adapt the adhesive properties, as it might be the case in the various niches colonized by *Veillonella*.

Trimeric autotransporters possess a characteristic YadA_anchor domain (PF03895) that can be found mainly in Proteobacteria, but also in Cyanobacteria, Verrumicrobia, Planctomycetes, Kiritimatiellaeota, Chlorobi, Synergistetes, Fusobacteria and in the diderm Firmicutes but only in Negativicutes (https://pfam.xfam.org/family/PF03895, Dec 2019 (44)). Interestingly, the YadA_anchor of 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 5). While the function of the coiled domain is unknown, in some bacteria, the periplasmic SLH domains bind peptidoglycan (45), 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 diderm phyla phylogenetically related to the Firmicutes such as Synergistetes and Fusobacteria (DataSet 1). Interestingly, in addition to possessing trimeric autotransporters with an extra coiled C-terminus domain, the Fusobacteria *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 (46) (DataSet 1). This suggests that a subset of phylogenetically close diderm 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). *FNLLGLLA_01127/yqeK* is found in a cluster of genes (namely *obg*, *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 S5 and S6, DataSet 2). *Staphylococcus aureus* YqeK was recently shown to be a nucleotidase hydrolyzing diadenosine-tetraphosphate (Ap4A) into ADP (47). 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 (48). c-di-GMP regulates biofilm formation by modulating production of a variety cell-surface appendages or exopolysaccharides in both monoderm and diderm bacteria (49–53). Interestingly, *B. subtilis* YqeK induces the

330 *epsA-O* operon, involved in the production of biofilm matrix-forming polysaccharides (54).
331 Deletion of *V. parvula FNLLGLLA_01127* only led to a minor decrease in biofilm formation in
332 96-well plate, but a strong increase in continuous-flow biofilm formation that was dependent
333 on the presence of the cluster of trimeric autotransporters. While we cannot exclude a direct
334 effect on the adhesins of the cluster, more likely FNLLGLLA_01127 participates to the
335 production/regulation of an unknown exopolysaccharide, which, contrary to *B. subtilis,* would
336 interfere with the function or exposure of the adhesins of the cluster rather than favor
337 community development. FNLLGLLA_01127 and its homologue YqeK are both involved in
338 biofilm formation, but the presence or absence of an outer membrane containing adhesins
339 changes the outcome of the HD phosphatase-mediated regulation of biofilm.

340

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

351

MATERIALS AND METHODS

Bioinformatic analyses

The *V. parvula* SKV38 genome was annotated using PROKKA (55) and complementary analyses were performed with RAST (56). Genetic comparison figures were generated with BRIG (57) or with SnapGene (GSL Biotech, www.snapgene.com) for visualization of genetic organization. The SKV38 annotated genome sequence was deposited in the ENA (European Nucleotide Archive) under the accession number ERZ1303164. Average Nucleotide Identity (ANI) was calculated using the following web tool http://enve-omics.ce.gatech.edu/ani/ (58). For protein domain visualization, PFAM domains (pfam.xfam.org, Pfam 32.0. (44)) were detected using HMMER (59). 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) (60).
For phylogenetic analyses, a local databank of the currently available 133 Negativicutes genomes was assembled and searched for InfB, RpoB and RpoC protein sequences using HMMER (59). Sequences were aligned using MAFFT (61), trimmed using BMGE with default parameters (62), concatenated, and used to generated a tree using IQTREE (63) with the LG+R5 model and ultrafast bootstrap approximation with 1000 replicates of the original dataset. Only the Veillonellaceae part of the tree was explored, the rest serving as an outgroup to root the tree. Different adhesin families were searched in the same local databank, using HMMER (59) and plotted onto the tree. *Veillonella* adhesin cluster synteny was analyzed using Easyfig for eight selected species (64) using tblastx.

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-

384 microaerophilic station. *Escherichia coli* was grown in Lysogeny Broth (LB) (Corning)
385 medium under aerobic conditions at 37°C. 20 mg L^{-1} chloramphenicol (Cm), 200 mg L^{-1}
386 erythromycin (Ery) or 2.5 mg L^{-1} tetracycline (Tc) were added to *V. parvula* cultures, 100 mg
387 L^{-1} carbenicillin (Cb) or 5 mg L^{-1} tetracycline (Tc) were added to *E. coli* cultures when needed.
388 100 µg L^{-1} anhydrotetracycline (aTc) was added to induce the pTet promoter unless stated
389 otherwise. All chemicals were purchased from Sigma-Aldrich unless stated otherwise.

390

391 **Natural transformation**

392 Cells were resuspended in 1 mL SK media adjusted to 0.4-0.8 OD_{600} and 10 µL were dotted on
393 SK-agar petri dishes. On each drop, 0.5-1 µg plasmid or 75-200 ng μL^{-1} linear dsDNA PCR
394 product was added, or water for the negative control. The plates were incubated 48 hours. The
395 biomass was resuspended in 500 µL SK medium and plated on SK-agar supplemented with the
396 corresponding antibiotic and incubated for another 48 hours. Colonies were streaked on fresh
397 selective plates and the correct integration of the construct was confirmed by PCR and
398 sequencing.

399

400 **Random mariner transposon mutagenesis**

401 Plasmid pRPF215, described for *Clostridium* mutagenesis (Addgene 106377) (20) was
402 transformed into *V. parvula* SKV38 by natural transformation and selected on Cm
403 supplemented SK-agar medium. An overnight culture of *V. parvula* SKV38-pRPF215 in
404 BHILC was then diluted to 0.1 OD_{600} in the same media, supplemented with aTc and grown 5
405 hours to induce the transposase. After induction, the culture was diluted and plated on BHILC
406 supplemented with Ery and aTc for selection and incubated for 48 hours. From the resulting
407 colonies, 940 were inoculated in Greiner Bio-one polystyrene flat-bottom 96-well plates
408 (655101) and grown in BHILC supplemented with either Ery and aTc, or Cm, to confirm both
409 the presence of the transposon and the loss of pRPF215 and then kept in 15 % glycerol at -
410 80°C. Selected transposon mutants were grown overnight and the genomic DNA was harvested
411 using DNeasy blood and tissue kit (Qiagen). The genomic DNA was then sent for whole
412 genome sequencing at the Mutualized platform for Microbiology of Institut Pasteur.

413

414 **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 (65)) 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 *Bam*HI (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, (65)) 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 ß-glucuronidase activity generated by the presence of pRPF185 transformed in *V. parvula* SKV38 (Figure S4).

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

449 Overnight cultures in BHILC medium were diluted to 0.05 OD_{600} and transferred to three
450 Greiner Bio-one polystyrene flat bottom 96-well plates adding 150 µL per well. After 24 hours
451 of static incubation, one of the three plates was resuspended by pipetting to measure OD_{600}
452 using a TECAN Infinite-M200-Pro spectrophotometer. The two other plates were used for
453 coloration: cultures were removed by pipetting carefully the supernatant out and biofilms fixed
454 with 150 µL Bouin solution (HT10132, Sigma-Aldrich) for 15 minutes. Bouin solution was
455 removed by inversion and the biofilms were washed once in water. The biofilms were stained
456 with 150 µL crystal violet 1 % (V5265, Sigma-Aldrich) for 15 minutes without shaking, then
457 washed in water twice and left to dry. All washes were made by flicking of the plate. After
458 drying the plate, crystal violet was dissolved with 200 µL absolute ethanol and transferred to a
459 clean 96-well plate for OD_{620} measurement (TECAN Infinite-M200-Pro spectrophotometer).
460
461 **Biofilm formation in microfermentor**
462 Continuous flow non-bubbled microfermentor containing a removable spatula were used as
463 described in (66, 67) (see https://research.pasteur.fr/en/tool/biofilm-microfermenters/). Briefly,
464 a glass spatula was dipped in an overnight culture diluted to 0.5 OD_{600} in 15 mL BHILC for 15
465 minutes and returned to the fermentor. Biofilm was grown on the spatula for 48 hours at 37°C.
466 BHILC was constantly supplied through a peristaltic pump at 4 rpm. During the last hour, the
467 speed was increased to 10 rpm to remove planktonic bacteria. A mix of filtered 90% nitrogen/5%
468 hydrogen/5% carbon dioxide was also constantly supplied to maintain anaerobic condition.
469 After 48 hours of growth, the spatula was removed, and the biofilm was resuspended by
470 vortexing in 15 mL BHILC. We measured OD_{600} of the resuspended biofilms with Smart Spec
471 Plus spectrophotometer (BioRad).
472
473 **Aggregation curve**
474 Overnight cultures were diluted to 0.8 OD_{600} in Brain-heart infusion (BHI) media in semi-micro
475 spectrophotometry cuvette (Fisherbrand) and left to sediment on the bench in presence of
476 oxygen, so no growth should occur. OD_{600} was measured every hour in a single point of the
477 cuvette using SmartSpec spectrophotometer (BioRad).
478
479 **Initial adhesion on glass**
480 Glass spatula from microfermentor (described above) were dipped in overnight cultures diluted

to 0.5 OD_{600} 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

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 5 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 (Ministère Français de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche) 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 S.G. 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, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch 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, Palaniappan K, Chain P, Rohde M, Göker M, Bristow J, Eisen JA, Markowitz V, Hugenholtz P, Kyrpides NC, Klenk HP, 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. 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.

20. 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

584 Clostridium difficile. MBio 6:e02383.

585 21. Łyskowski A, Leo JC, Goldman A. 2011. Structure and Biology of Trimeric
586 Autotransporter Adhesins, p. 143–158. *In* Advances in experimental medicine and
587 biology.

588 22. Ruiz-Ranwez V, Posadas DM, Estein SM, Abdian PL, Martin FA, Zorreguieta A.
589 2013. The BtaF trimeric autotransporter of Brucella suis is involved in attachment to
590 various surfaces, resistance to serum and virulence. PLoS One 8:e79770.

591 23. Zhou P, Liu J, Merritt J, Qi F. 2015. A YadA-like autotransporter, Hag1 in Veillonella
592 atypica is a multivalent hemagglutinin involved in adherence to oral streptococci,
593 Porphyromonas gingivalis, and human oral buccal cells. Mol Oral Microbiol 30:269–
594 279.

595 24. Trunk T, S. Khalil H, C. Leo J. 2018. Bacterial autoaggregation. AIMS Microbiol
596 4:140–164.

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

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

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

605 28. Ageorges V, Schiavone M, Jubelin G, Caccia N, Ruiz P, Chafsey I, Bailly X, Dague E,
606 Leroy S, Paxman J, Heras B, Chaucheyras-Durand F, Rossiter AE, Henderson IR,
607 Desvaux M. 2019. Differential homotypic and heterotypic interactions of antigen 43
608 (Ag43) variants in autotransporter-mediated bacterial autoaggregation. Sci Rep
609 9:11100.

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

614 30. Totsika M, Wells TJ, Beloin C, Valle J, Allsopp LP, King NP, Ghigo J-M, Schembri

615 MA. 2012. Molecular characterization of the EhaG and UpaG trimeric autotransporter
616 proteins from pathogenic Escherichia coli. Appl Environ Microbiol 78:2179–89.

617 31. Valle J, Mabbett AN, Ulett GC, Toledo-Arana A, Wecker K, Totsika M, Schembri
618 MA, Ghigo J-M, Beloin C. 2008. UpaG, a New Member of the Trimeric
619 Autotransporter Family of Adhesins in Uropathogenic Escherichia coli. J Bacteriol
620 190:4147–4161.

621 32. Paton AW, Srimanote P, Woodrow MC, Paton JC. 2001. Characterization of Saa, a
622 novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative
623 Shiga-toxigenic Escherichia coli strains that are virulent for humans. Infect Immun
624 69:6999–7009.

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

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

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

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

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

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

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

645 40. Mashima I, Nakazawa F. 2014. The influence of oral Veillonella species on biofilms

formed by Streptococcus species. Anaerobe 28:54–61.

41. 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.

42. 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.

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

44. 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.

45. Mesnage 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.

46. Park JS, Lee WC, Yeo KJ, Ryu K-S, Kumarasiri M, Hesek D, Lee M, Mobashery S, Song JH, Kim S Il, 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.

47. 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.

48. 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.

49. 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.

50. Peng X, Zhang Y, Bai G, Zhou X, Wu H. 2016. Cyclic di-AMP mediates biofilm

677 formation. Mol Microbiol 99:945–959.

678 51. Townsley L, Yannarell SM, Huynh TN, Woodward JJ, Shank EA. 2018. Cyclic di-
679 AMP Acts as an Extracellular Signal That Impacts Bacillus subtilis Biofilm Formation
680 and Plant Attachment. MBio 9.

681 52. Jenal U, Reinders A, Lori C. 2017. Cyclic di-GMP: second messenger extraordinaire.
682 Nat Rev Microbiol 15:271–284.

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

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

688 55. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics
689 30:2068–2069.

690 56. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S,
691 Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED
692 and the Rapid Annotation of microbial genomes using Subsystems Technology
693 (RAST). Nucleic Acids Res 42:D206–D214.

694 57. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image
695 Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12:402.

696 58. Rodriguez-R L, Konstantinidis K. 2016. The enveomics collection: a toolbox for
697 specialized analyses of microbial genomes and metagenomes.

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

700 60. Lupas A, Dyke M Van, Stock J. 1991. Predicting Coiled Coils from Protein Sequences.
701 Science (80-). American Association for the Advancement of Science.

702 61. Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software
703 Version 7: Improvements in Performance and Usability. Mol Biol Evol 30:772–780.

704 62. Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy):
705 a new software for selection of phylogenetic informative regions from multiple
706 sequence alignments. BMC Evol Biol 10:210.

63. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol 32:268–274.

64. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27:1009–1010.

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

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

67. Beloin C, Valle J, Latour-Lambert P, Faure P, Kzreminski M, Balestrino D, Haagensen JAJ, Molin S, Prensier 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.

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

69. 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.

70. Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: Rapid annotations using subsystems technology. BMC Genomics 9.

71. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206-14.

72. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes.

738 Sci Rep 5.

739

740

741 **TABLES**

742

743

Locus tag	PROKKA Gene name	Genome position		Gene size (kb)	Strand	Description	DSM2008 homolog	Name	Class
FNLLGLLA_00032	prn 1	39,354	41,723	2,37	forward	Autotransporter	fusion Vpar_0036-0037	VmaA	Va
FNLLGLLA_00034	btaE 1	42,345	43,754	1,41	reverse	Trimeric Autotransporter: YadA like	Vpar_0039	VtaB	Vc
FNLLGLLA_00035	hypothetical protein	44,146	45,189	1,04	forward	Autotransporter (partial)	Vpar_0040		Ve
FNLLGLLA_00036	hypothetical protein	45,453	46,883	1,431	forward	none	split Vpar_0041		?
FNLLGLLA_00037	omp-alpha	46,91	47,878	969	forward	Trimeric Autotransporter/ S-layer homology domain	split Vpar_0041		Vc?
FNLLGLLA_00038	upaG 1	48,397	56,829	8,433	forward	Trimeric Autotransporter: YadA like	Vpar_0042	VtaC	Vc
FNLLGLLA_00040	btaE 2	57,966	59,84	1,875	forward	Trimeric Autotransporter: YadA like (Partial)	split Vpar_0048		?
FNLLGLLA_00041	ata 1	59,837	63,463	3,627	forward	Trimeric Autotransporter: YadA like	split Vpar_0048		Vc?
FNLLGLLA_00044	ehaG 1	65,3	71,515	6,216	forward	Trimeric Autotransporter: YadA like	Vpar_0051	VtaD	Vc
FNLLGLLA_00045	upaG 2	71,995	81,42	9,426	forward	Trimeric Autotransporter: YadA like	Vpar_0052	VtaE	Vc
FNLLGLLA_00046	ata 2	81,941	91,519	9,579	forward	Trimeric Autotransporter: YadA like	Vpar_0053	VtaF	Vc
FNLLGLLA_00098	btaE 3	151,792	153,522	1,731	forward	Trimeric Autotransporter/ S-layer homology domain	Vpar_0100	VtaG	Vc
FNLLGLLA_00099	ata 3	154,024	158,982	4,959	forward	Trimeric Autotransporter/ S-layer homology domain	absent	VtaH	Vc
FNLLGLLA_00335	prn 2	414,666	416,888	2,223	forward	Autotransporter	Vpar_0330	VmaB	Va
FNLLGLLA_00516	btaF	581,236	590,358	9,123	forward	Trimeric Autotransporter: YadA like	Vpar_0464	VtaA	Vc
FNLLGLLA_00581	brkA	668,34	670,583	2,244	forward	Autotransporter	Vpar_1322	VmaC	Va
FNLLGLLA_01790	ehaG 2	1,943,661	1,946,159	2,499	reverse	Trimeric Autotransporter/ S-layer homology domain	Vpar_1664	VtaI	Vc

744 Table 1 : ***V. parvula*** **SKV38 autotransporters**

745

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 Δ*FNLLGLLA_00516::tetM*	This study
pTet-vtaA	*Veillonella parvula* SKV38 *catP-Term(fdx)-Ptet-FNLLGLLA_00516*	This study
Δ8	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00036-46::tetM*	This study
ΔvmaA	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00032::tetM*	This study
ΔvtaB	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00034::tetM*	This study
ΔvtaG	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00098::tetM*	This study
ΔvtaH	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00099::tetM*	This study
ΔvmaB	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00335::tetM*	This study
ΔvmaC	*Veillonella parvula* SKV38 Δ*FNLLGLLA_00581::tetM*	This study
ΔvtaI	*Veillonella parvula* SKV38 Δ*FNLLGLLA_01790::tetM*	This study
ΔvtaAΔ8	*Veillonella parvula* SKV38 Δ*FNLLGLLA_ 00516::catP* Δ*FNLLGLLA_00036-46::tetM*	This study
Δ1127	*Veillonella parvula* SKV38 Δ*FNLLGLLA_01127::tetM*	This study
Δ1127Δ8	*Veillonella parvula* SKV38 Δ*FNLLGLLA_01127::tetM* Δ*FNLLGLLA_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_{tet}-φ	SKV38- pRPF185*ΔgusA*	This study
P_{tet}-*gusA*	SKV38- pRPF185	This study
P_{Cwp2} -*gusA*	SKV38- pRPF144	This study

Plasmid	Description	Reference

pRPF215	*mariner* Tn delivery plasmid, P_{tet}::*Himar1*, ITR-*ermB*-ITR, *catP*, *tetR*	(20)
pRPF185	tetracycline inducible expression system fused with ß-glucuronidase *gusA* Term(fdx)- P_{tet} -gusA-Term(slpA), *catP*	(65)
pRPF185Δ*gusA*	pDIA6103, tetracycline inducible expression system Term(fdx)-P_{tet}-Term(slpA), *catP*	(68)
pRPF144	*carries a Clostridium* constitutitive promoter fused with *gusA* P_{Cwp2}-*gusA*	(65)
pBSJL2	*E. coli-Veillonella* shuttle plasmid, P_{gyrA}::*tetM*	(69)
pBSJL2-cat	*E. coli-Veillonella* shuttle plasmid, P_{cat}::*catP, pmdh* promoter	This study
p1127	pBSJL2-*catP-pmdh-FNLLGLLA_01127*	This study

Table 2: Strains and plasmids used in this study

750 **FIGURES AND FIGURES LEGEND**

751

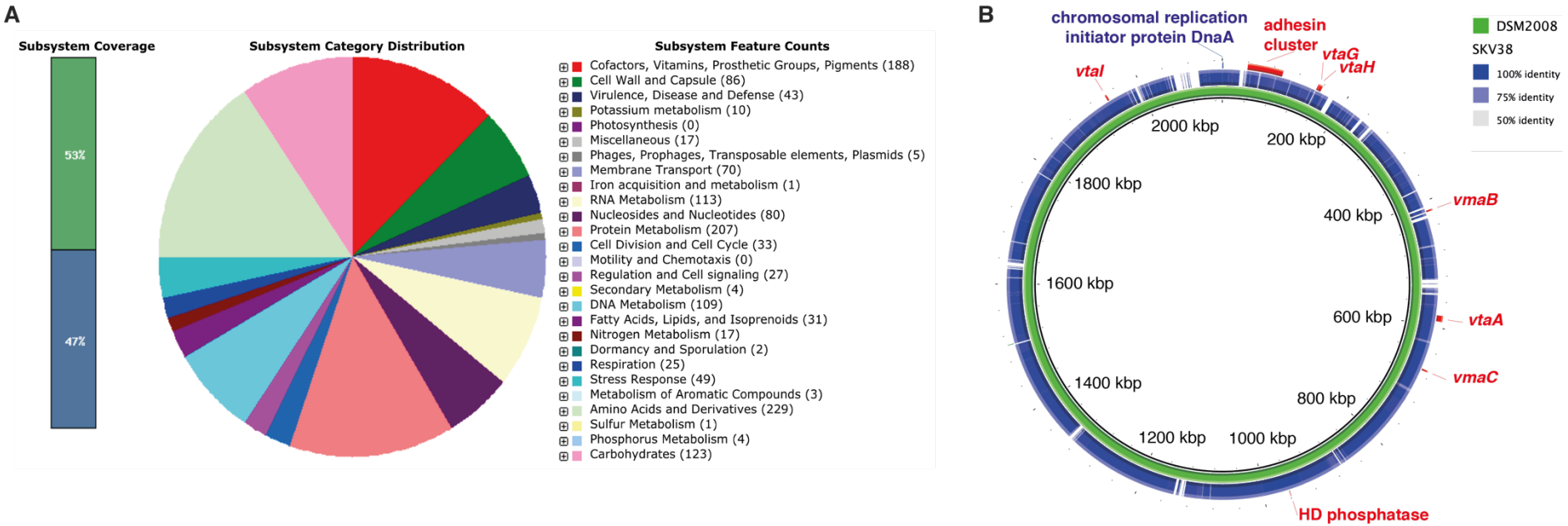


Figure 1: **V. parvula *SKV38 genome analysis.*** **A.** RAST functional annotation of the *V. parvula* SKV38 genome. Left panel: Subsystem coverage. Green color indicates the percentage of predicted proteins assigned to known subsystems, blue those that were not. Center and right panels: Subsystem category distribution: the number of proteins assigned to each category is depicted on the right panel. The analysis was performed on https://rast.theseed.org server (70–72) **B.** Comparison of the genetic identity between *V. parvula* SKV38 and the reference strain, DSM2008. The position of genes coding for potential biofilm determinants discussed in this study are indicated in the *V. parvula* genome.

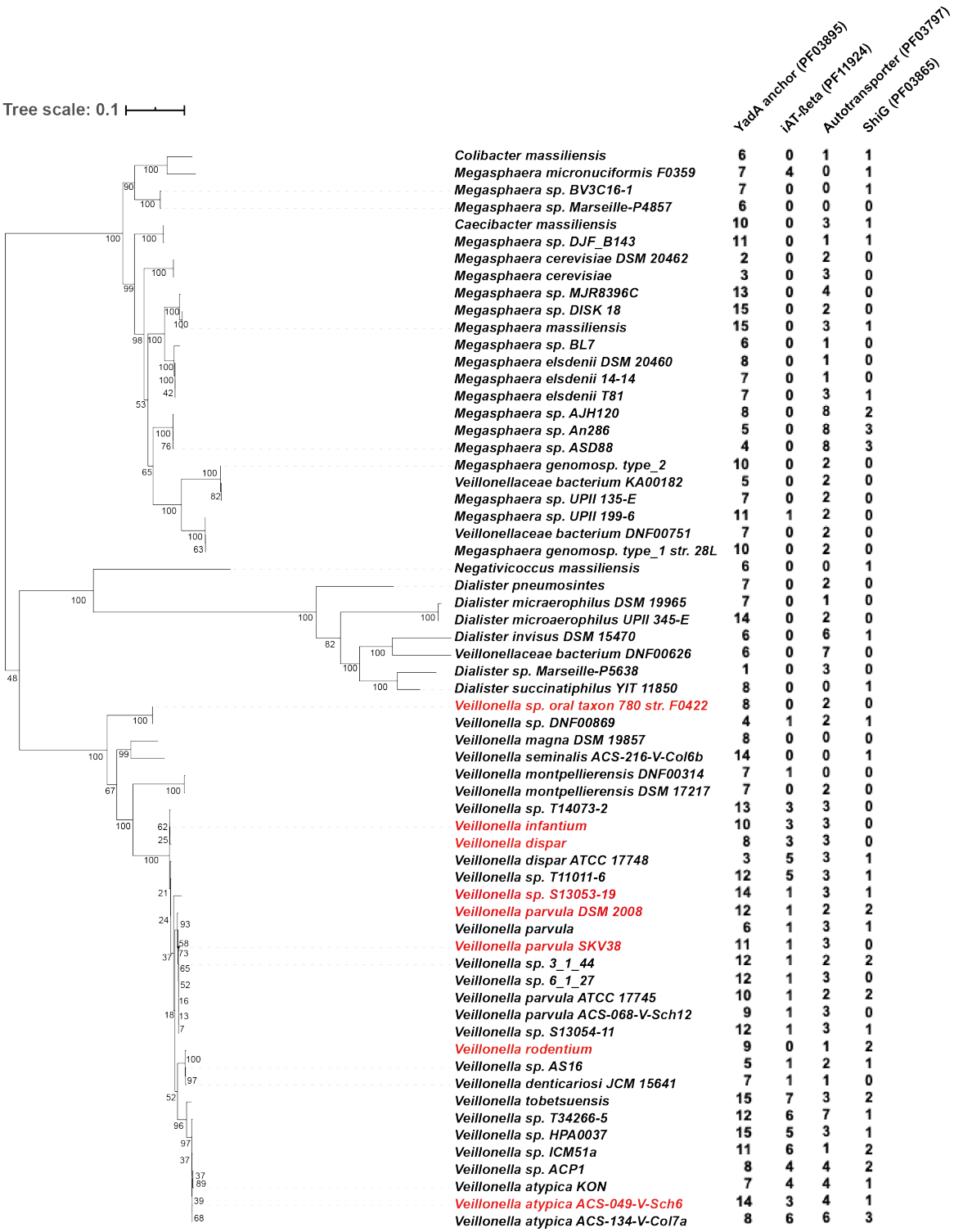


Figure 2**: Phylogenetic tree of Veillonellaceae and prevalence of the main type V secreted adhesins.** Maximum likelihood phylogenetic tree constructed from a concatenated dataset of RpoB, RpoC, InfB, comprising 3027 amino acid positions, and rooted by other Negativicutes (see material and methods). Numbers at nodes represent bootstrap values calculated based on 100 replicates of the initial character supermatrix. The number of the four main classes of potential type V secreted adhesins queried by an HMM search is indicated for every species. The scale bar indicates the average number of substitutions per site. Species whose adhesin gene clusters were analyzed on Figure 7 are colored in red.

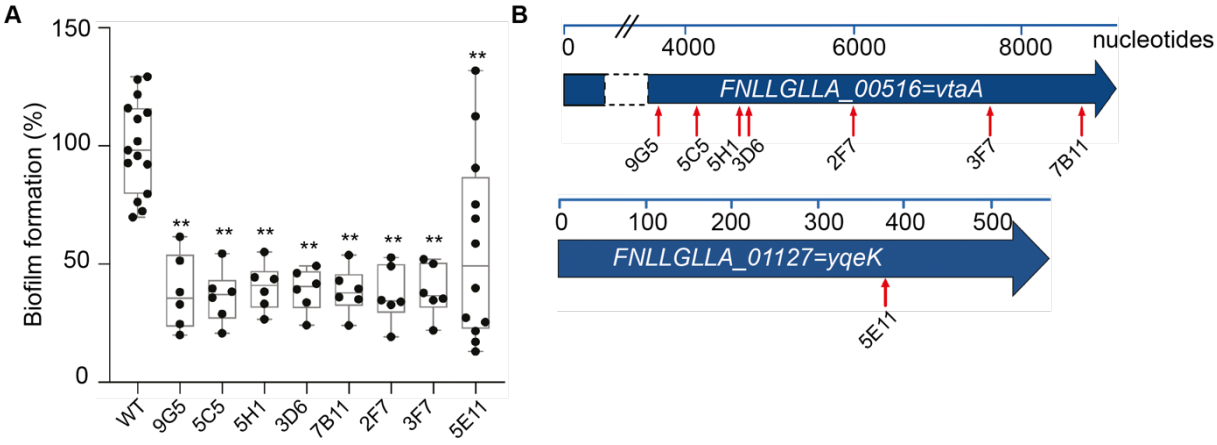


Figure 3: **Random transposon mutagenesis in *Veillonella parvula* SKV38 led to identification of mutants with reduced biofilm formation. A**. 96-well polystyrene plate biofilm assay after CV staining of nine transposon mutants identified by random mutagenesis grown 24h in BHILC. Mean of WT is adjusted 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.

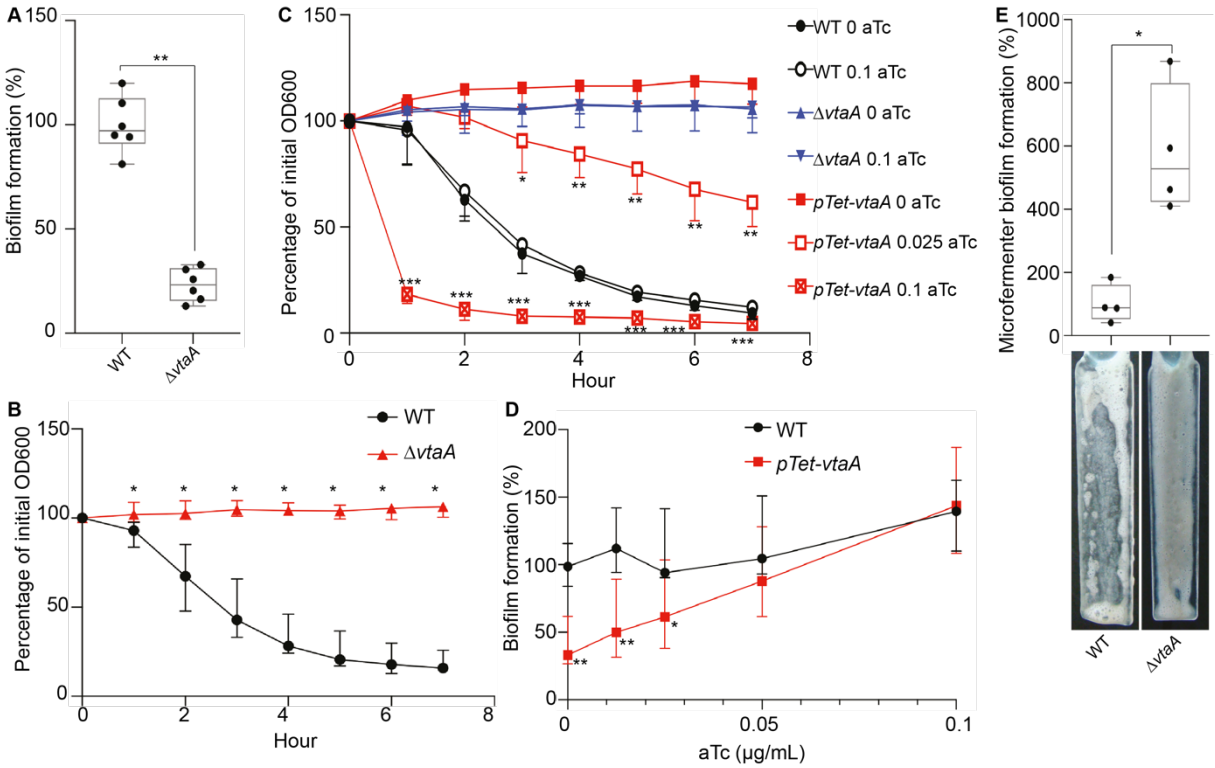
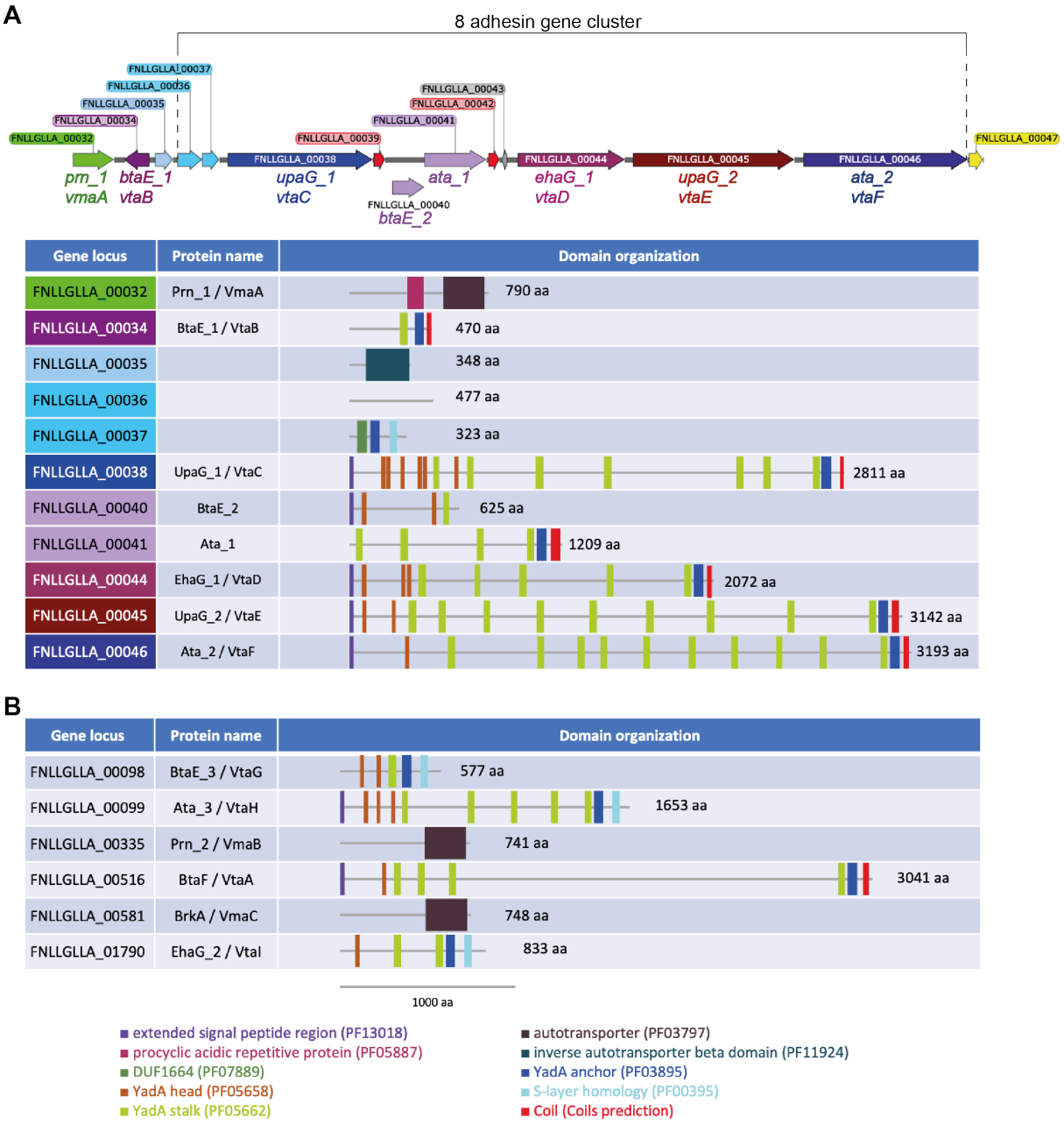


Figure 4: **VtaA is an adhesin involved in auto-aggregation and biofilm formation**. **A**. 96-well plate biofilm assay after 24h growth in BHILC. Mean of WT is adjusted to 100 %. Min-max boxplot of 6 biological replicates for each strain. * p-value<0.05, ** p-value <0.005, Mann-Whitney test between strains. **B.** and **C.** Aggregation curve in spectrophotometry cuvette of WT and Δ*vtaA* (**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 Δ*vtaA*, 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 histogram bar. * p-value<0.05, Mann-Whitney test.

800



801

802 Figure 5: ***Veillonella parvula*** **autotransporters domain organization. A.** Genetic
803 organization of the *V. parvula* SKV38 autotransporter adhesin gene cluster and the
804 corresponding adhesin domain organization. **B**. Domain organization of the six remaining *V.*
805 *parvula* SKV38 autotransporter adhesins encoded by genes located outside of the cluster.
806 Domains were detected with the HMMER package (59), only the domains with e-values lower
807 than 10^{-3} are shown. Presence of C-terminal coils structure was determined using the COILS
808 program (https://embnet.vital-it.ch/software/COILS_form.html).

809

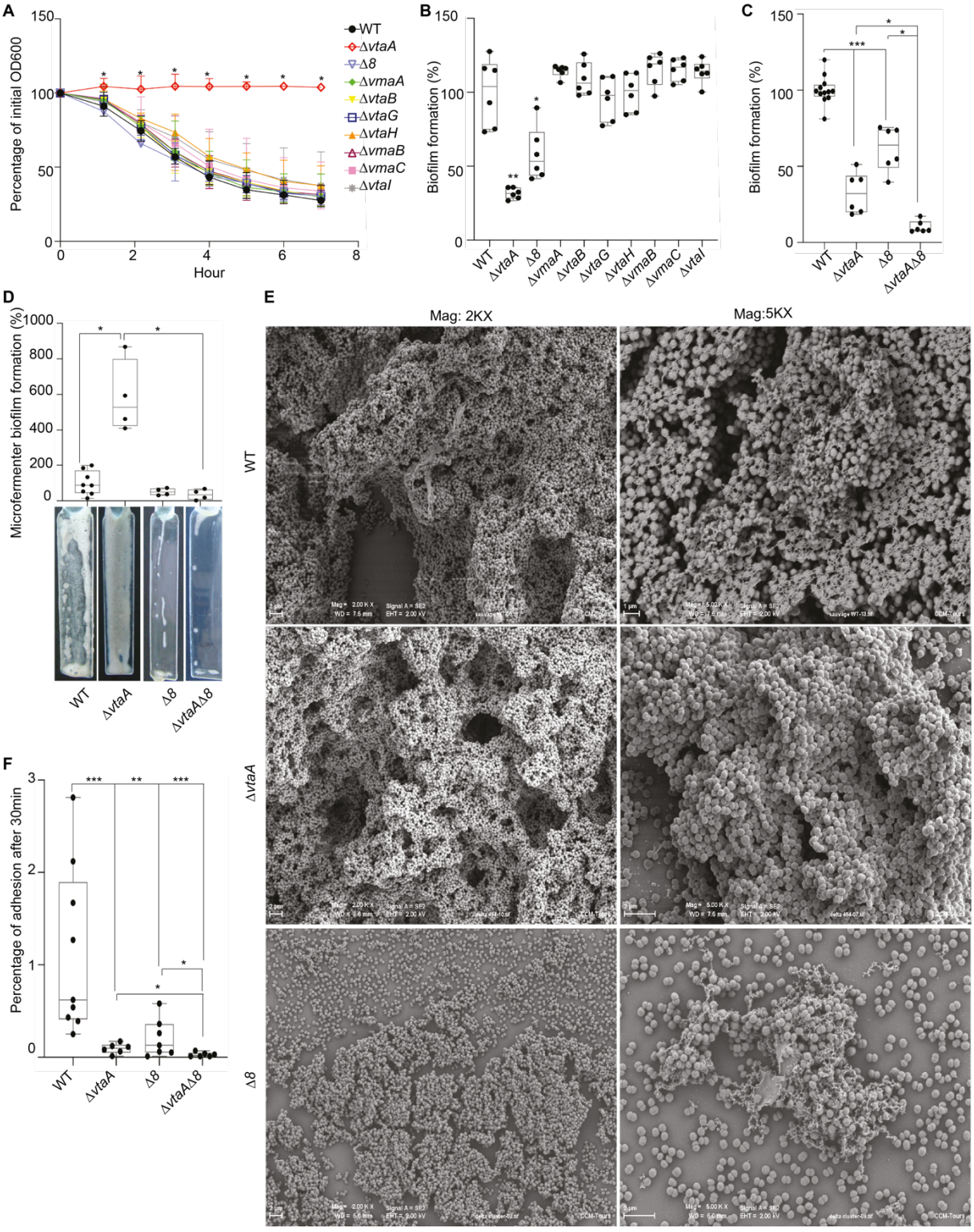


Figure 6: **A cluster of eight trimeric autotransporters is involved in surface binding.** **A.** Aggregation curve in spectrophotometry cuvette. 100 % represent lack of aggregation, 0 % complete sedimentation of the culture. Median of 6 biological replicates, error bars represent 95% confidence interval. * 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

819 for multiple testing: tests were called significant only if p-value<0.01: * p-value<0.01, ** p-
820 value <0.001, *** p-value <0.0001. **D.** Biofilm formation in continuous flow microfermentor
821 on glass spatula during 48h in BHILC. WT was adjusted to 100 %. Min-max boxplot of 4
822 biological replicates for each strain. * p-value<0.05, Mann-Whitney test. A picture of spatula
823 before resuspension is shown for each mutant on the right. **E.** Scanning electronic microscopy
824 (SEM) of biofilms grown under continuous flow of BHILC in microfermentor on plastic
825 microscopy slide, magnified 2K or 5K times. **F.** Initial adhesion on glass spatula. Percentage
826 of CFU that adhered to the spatula controlled by the number of CFU of the inoculation solution.
827 Min-max boxplot of 6-9 replicates per strain is represented. * p-value<0.05, ** p-value <0.005,
828 *** p-value <0.0005, Mann-Whitney test.

829

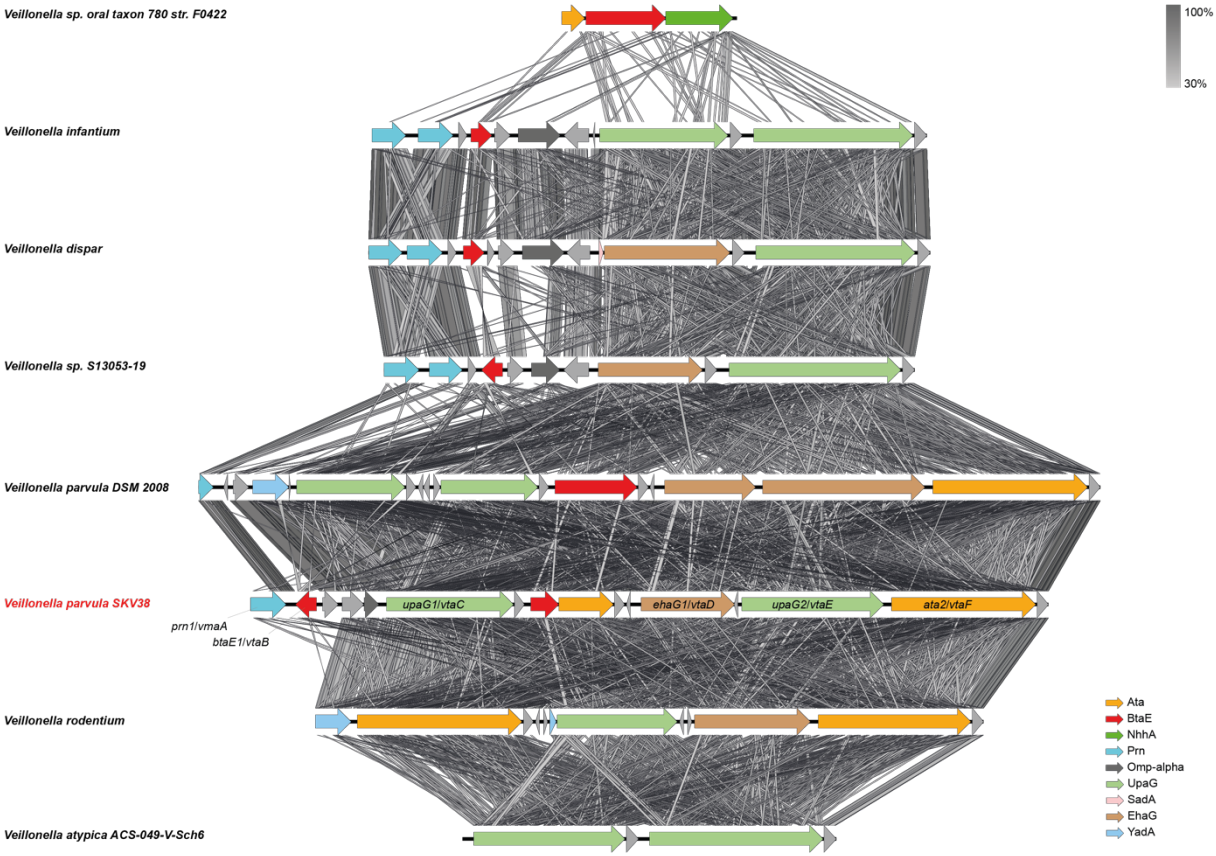
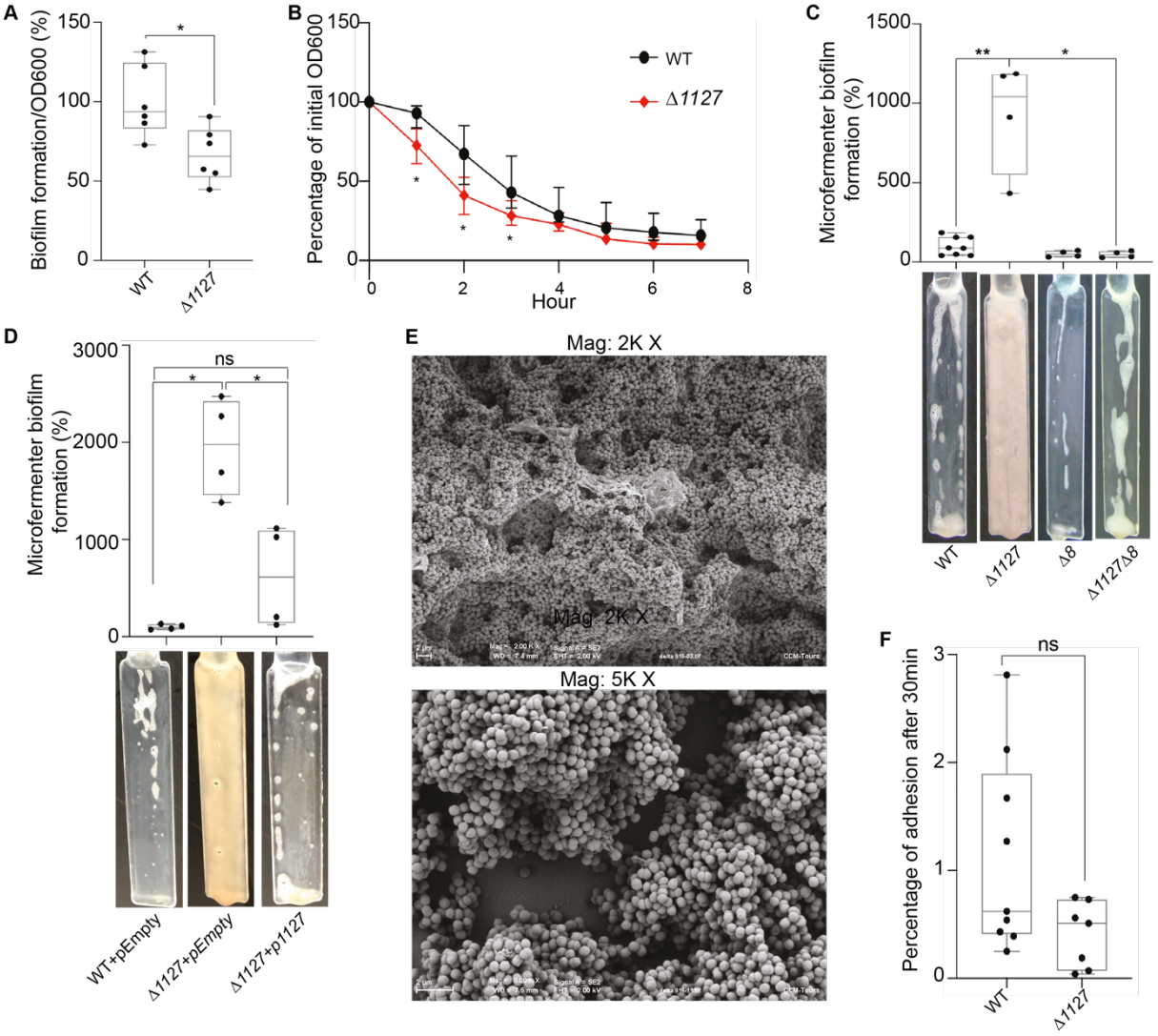


Figure 7: **Synteny of the adhesin gene cluster in a selection of *Veillonella* species**. The synteny of the proteins of the cluster between the closest relatives was assessed using EasyFig (64). Oblique lines between genes represent tblastx identities (program parameters: maximum e-value of 10^{12}, minimum length of 30, minimum identity of 30). The *V. parvula* SKV38 strain used in this study is presented in red. The functional genes of the cluster are indicated.



837

838 Figure 8: ***FNLLGLLA_01127* represses biofilm formation in microfermentor**. **A.** 96-well
839 plate biofilm assay after 24 h growth in BHILC corrected by OD_{600} after 24 h growth in plate.
840 Mean of WT is adjusted to 100 %. Min-max boxplot of 6 biological replicates for each strain,
841 each replicate is the mean of two technical replicates. * p-value < 0.05, Mann-Whitney test. **B.**
842 Aggregation curve in spectrophotometry cuvette. 100 % represent lack of aggregation, 0 %
843 complete sedimentation of the culture. Median of 6 biological replicates, error bars represent
844 95% confidence interval. * Mann-Whitney test, corrected for multiple testing with Bonferroni
845 correction: significance is achieved if p-value<0.007. **C.** Biofilm formation in continuous flow
846 microfermentor on glass spatula during 48h in BHILC. Mean of WT is adjusted to 100 %. Min-
847 max boxplot of 4 biological replicates for each strain. * p-value < 0.05, ** p-value<0.005,
848 Mann-Whitney test. A picture of a spatula before resuspension is shown for each strain below
849 the histogram. **D.** Biofilm formation in continuous flow microfermentor on glass spatula during
850 48h in BHILC+chloramphenicol. Mean of WT+pEmpty is adjusted to 100 %. Min-max boxplot
851 of 4 biological replicates for each strain. * p-value < 0.05, Mann-Whitney test. A picture of a
852 spatula before resuspension is shown for each strain below the histogram. **E.** Scanning
853 electronic microscopy of Δ*1127* biofilm grown under continuous flow of BHILC in
854 microfermentor on a plastic microscopy slide. Magnification 2K and 5K. **F.** Initial adhesion on
855 glass spatula. Percentage of CFU that adhered to the spatula in 30 min controlled by the number
856 of CFU of the inoculation solution. Min-max boxplot of 6-9 replicates per strain. * p-
857 value<0.05, Mann-Whitney test.