



Published in final edited form as:

Mol Microbiol. 2022 May ; 117(5): 1023–1047. doi:10.1111/mmi.14887.

Dual Adhesive Unipolar Polysaccharides Synthesized by Overlapping Biosynthetic Pathways in *Agrobacterium tumefaciens*

Maureen C. Onyeziri¹, Gail G. Hardy¹, Ramya Natarajan¹, Jing Xu¹, Ian P. Reynolds¹, Jinwoo Kim^{1,a}, Peter M. Merritt¹, Thomas Danhorn^{1,b}, Michael E. Hibbing^{1,c}, Alexandra J. Weisberg², Jeff H. Chang², Clay Fuqua^{1,*}

¹Department of Biology, Indiana University, Bloomington, IN 47405

²Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331

Summary

Agrobacterium tumefaciens, is a member of the Alphaproteobacteria that pathogenizes plants, and associates with biotic and abiotic surfaces via a single cellular pole. *A. tumefaciens* produces the unipolar polysaccharide (UPP) at the site of surface contact. UPP production is normally surface-contact inducible, but elevated levels of the second messenger cyclic diguanylate monophosphate (cdGMP) bypass this requirement. Multiple lines of evidence suggest that the UPP has a central polysaccharide component. Using an *A. tumefaciens* derivative with elevated cdGMP and mutationally disabled for other dispensable polysaccharides, a series of related genetic screens have identified a large number of genes involved in UPP biosynthesis, most of which are Wzx-Wzy-type polysaccharide biosynthetic components. Extensive analyses of UPP production in these mutants have revealed that the UPP is composed of two genetically, chemically and spatially discrete forms of polysaccharide, and that each requires a specific Wzy-type polymerase. Other important biosynthetic, processing and regulatory functions for UPP production are also revealed, some of which are common to both polysaccharides, and a subset of which are specific to each type. Many of the UPP genes identified are conserved among diverse rhizobia, whereas others are more lineage specific.

Graphical Abstract

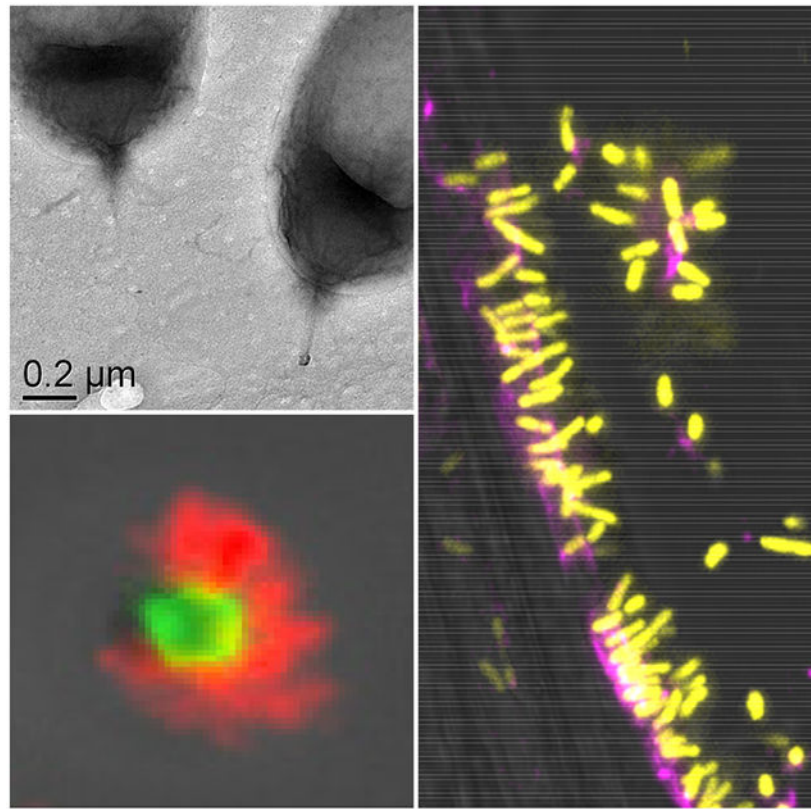
*Address for correspondence: Clay Fuqua, Dept. Biol., 1001 E. 3rd St., Biology Building 142, Indiana Univ., Bloomington, IN 47405-1847. Tel: 812-856-6005, FAX: 812-855-6705, cfuqua@indiana.edu.

^aDept. Plant Medicine, Gyeongsang National University, Jinju, Korea

^bUniversity of Colorado Cancer Center, Anschutz Medical Campus, Aurora, CO 80045

^cFimbrion Therapeutics, St. Louis MO

Author contributions: C.F. formulated the research plan; M.O., R.N., G.H., J.X., I.R., J.K., T.D., M.H. and C.F. performed the experiments, J.C. and A.W. performed genomic sequence analyses; C.F., M.O. and G.H. wrote the paper. The authors have no conflicts of interest to declare.



Introduction

Many pathogenic bacteria stably associate with the tissues of their hosts as an obligate component of infection. This physical association is usually mediated by adhesins elaborated from the bacterial surface, including pili, flagella, other externalized proteins and exopolysaccharides (Berne et al., 2018). Adhesins can also drive bacterial attachment to abiotic surfaces. On both abiotic surfaces and host tissues, adherent cells may assemble into multicellular biofilms, protected bacterial communities that are physically and chemically recalcitrant (Flemming et al., 2016). Biofilm formation is driven by a combination of cell-to-surface interactions, intercellular cohesion, and extrusion of exopolymeric materials to build an extracellular matrix.

We discovered a polysaccharide-based adhesin in *Agrobacterium tumefaciens* (Xu et al. 2012), a member of the Alphaproteobacteria (APB) and causative agent of crown gall disease, on a wide variety of dicotyledonous plants (Gelvin, 2003). Crown gall is a neoplastic disease that results from the transfer of a large segment of DNA (T-DNA) from the bacteria into plant cells at the site of infection via the activity of a type IV secretion system (Li & Christie, 2018). This interkingdom gene transfer process, both its regulation and the mechanism of transfer, has been the subject of intense study and T-DNA transfer has been extensively engineered for the generation of transgenic plants. There is however, limited understanding of the initial colonization events that occur on the tissue surface, both for benign interactions and those that lead to the initiation of T-DNA transfer (Fuqua, 2008).

Micrographs dating as early as the 1940s suggest that *A. tumefaciens* often attaches to surfaces via a single cellular pole (Braun & Elrod, 1946, Tomlinson & Fuqua, 2009). Surfaces to which *A. tumefaciens* can attach include plant tissues and cells, other bacterial cells to form multicellular rosettes, and abiotic materials. We discovered that *A. tumefaciens* produces an adhesive structure at one of its poles that is predominantly composed of polysaccharide, and hence we designated this the unipolar polysaccharide (UPP) (Tomlinson & Fuqua, 2009). Fluorescently-tagged lectins including wheat germ agglutinin (WGA) and *Dolichos biflorus* agglutinin (DBA) label the UPP as a small fluorescent focus at the pole that is in contact with abiotic surfaces, plant tissues, and other bacterial cells (Xu et al., 2012, Xu et al., 2013). The UPP structure shares facile and functional similarity to the adhesive holdfast localized to the tip of the cellular stalk for the prosthecate APB *Caulobacter crescentus* (Poindexter, 1981). In fact, production of the UPP is coordinated with the biphasic life cycle of *A. tumefaciens* in a similar fashion to the holdfast in *C. crescentus*, and is produced by mother cells, that are often adhered to surfaces, and will generate multiple progeny daughter cells (Brown et al., 2012). In both bacterial taxa, production of the unipolar adhesin is stimulated by surface contact. In *C. crescentus*, this accelerates holdfast production to earlier in the cell cycle, but in *A. tumefaciens* surface contact is required for cells to produce the UPP (Li et al., 2012). The smaller size and strict surface-contact dependence of UPP production led to the structure being overlooked for many years, relative to the more microscopically conspicuous holdfast of *C. crescentus*. Several *A. tumefaciens* mutants have been characterized which exhibit a bypass of surface-contact dependence and elevated UPP production due to increased levels of the cytoplasmic second messenger cyclic diguanylate monophosphate (cdGMP) (Feirer et al., 2017, Feirer et al., 2015, Xu et al., 2013).

A cluster of six predicted polysaccharide biosynthetic genes (ATU_RS06125-RS06100) was initially implicated based on adhesion defects of a small number of mutants in *A. tumefaciens* (Xu et al., 2012). These genes are homologous to six genes in a syntenic cluster required for synthesis of a glucomannan involved in recognition of host legumes during formation of symbiotic nodules in the related nitrogen-fixing APB *Rhizobium leguminosarum* bv. *viciae* (Williams et al., 2008). Deletion of the entire six-gene cluster in *A. tumefaciens* led to complete loss of biofilm formation on abiotic surfaces and abolished lectin labeling of the UPP (Xu et al., 2012). In-frame deletion of a single gene in this cluster (ATU_RS06105) encoding a polyisoprenylphosphate hexose-1-phosphate transferase (PHPT), homologous to the HfsE glycosyltransferase from *C. crescentus* required for holdfast synthesis (Toh et al., 2008), also abolished surface attachment and visible UPP production. In addition to this PHPT protein, the other five genes in this cluster encode homologs to polysaccharide biosynthesis proteins (Table S1), and we provisionally designated this gene cluster as *uppABCDEF* (Thompson et al., 2018).

The *uppA-F* gene cluster is completely or partially conserved among diverse members of the Rhizobiales, and mutations in several of these homologs have been shown to impact production of a unipolar polysaccharide and rosette formation in *Rhodopseudomonas palustris* (Fritts et al., 2017). Thus, in at least three members of the Rhizobiales, homologs of the *A. tumefaciens uppA-F* genes are required for synthesis of polar polysaccharides, and it seems likely that this function is conserved in these bacteria. However, notably absent from

these gene clusters are homologs of the Wzx-type polysaccharide flippase enzymes and the Wzy-type polysaccharide polymerases, two of the defining components of these pathways (Cuthbertson et al., 2009). In this study, we use targeted mutational analysis along with several distinct forward genetic approaches to expand and deepen our prior analyses on *A. tumefaciens* UPP biosynthesis, identifying over 15 additional genes in six genomic clusters that are required for UPP production and attachment to surfaces. We combine this genetic analysis with microscopy and functional assays to reveal that there are at least two distinct polysaccharides that both contribute to the *A. tumefaciens* UPP, and that the biosynthetic pathways are genetically separable, but also overlapping and co-regulated.

Results

Mutations in the *uppABCDEF* gene cluster disrupt attachment and UPP production

A. tumefaciens synthesizes the UPP adhesin when it attaches to surfaces, including plant tissues and other bacterial cells, that can be readily detected with fluorescently tagged lectins such as Alexa Fluor 594-tagged WGA (af-WGA) (Fig. 1). Null mutations in the *uppE* gene (ATU_RS06105) within the *uppABCDEF* cluster result in loss of lectin labeling and surface attachment (Xu et al. 2012; Fig. 2A), but the roles for the other genes in this cluster had not been examined. Using the general polysaccharide biosynthetic nomenclature, UppE encodes a WbaP initiating glycosyltransferase, UppB (ATU_RS06120) encodes a predicted Wzb-type periplasmic co-polymerase (PCP), and UppC (ATU_RS06115) encodes a putative Wza-type outer membrane polysaccharide secretin (OPX) (Table S1). The other three genes are not restricted to Wzx-Wzy systems, but are predicted polysaccharide biosynthesis genes; UppA (ATU_RS06125) encodes a putative polysaccharide acetyltransferase with similarity to BcsL involved in bacterial cellulose synthesis, UppD (ATU_RS06110) encodes a Type 4 glycosyltransferase (GT4), and UppF (ATU_RS06100) is similar to WaaL O-antigen ligases that often transfer polysaccharides to the core of lipopolysaccharide (LPS), in formation of the O-antigen or LPS-conjugated polysaccharides (Romling & Galperin, 2015, Ruan et al., 2018).

To test the other genes in the *uppABCDEF* cluster (ATU_RS06125-RS06100), we created in-frame deletions of each gene individually and observed significant decreases in biofilm formation on PVC coverslips (Fig. 3, Table S1). The *uppC* and *uppE* deletion mutants exhibited the most profound defects, but the *uppA*, *uppB*, *uppD* and *uppF* deletion mutants resulted in substantially decreased biofilm formation, although still detectable. Short term attachment assays with cells incubated on coverslips revealed weak attachment in all mutants and loss of labeling with af-WGA for all of the mutants (Fig. 4, Table S1). Attachment and biofilm formation deficiencies of deletion mutants were fully complemented by ectopic expression with plasmid-borne copies of these genes, with the exception of the *uppC* mutant, which was only partially complemented with the *P_{lac}-uppC* plasmid (Fig. 3). We realized that the *uppC* deletion also impacted the overlapping downstream *uppD* coding sequence, and provision of a plasmid expressing both *uppC* and *uppD* fully reversed the *uppC* mutant phenotype. As described above, the six *upp* genes are all homologous to genes that encode polysaccharide biosynthetic proteins (Table S1). However, it was clear that based on their predicted functions in a Wzx-Wzy-type pathway, they would

be insufficient and that there must be additional genes required for UPP synthesis. Most notably, homologs of both the Wzx flippase and the Wzy polymerase are not found in this gene cluster nor in proximal genes.

A facile genetic screen for UPP biosynthetic genes

In prior studies, we had identified a set of mutants in four different loci (*visR* ATU_RS02585, *claR* ATU_RS08010, *pruA* ATU_RS05580, and *dcpA* ATU_RS16190), that manifested elevated production of the UPP coincident with elevated cellular cdGMP levels (Feirer et al., 2017, Feirer et al., 2015, Xu et al., 2013). The screen was based on the polysaccharide-binding, azo-dye Congo red (CR, which can also bind to β -amyloid proteins), and in all cases elevated CR pigmentation correlated with increased UPP production and hyper-adherence to surfaces. Wild type *A. tumefaciens* forms very pale orange colonies when cultured on agar plates of a standard minimal medium (ATGN) supplemented with CR (ATGN-CR) (Xu et al., 2013). Elevated production of either the UPP or cellulose results in red pigmentation of bacterial colonies grown on ATGN-CR, which we designated the ECR (Elevated Congo Red) phenotype. In an *A. tumefaciens* mutant disabled for cellulose production (Cel⁻), the ECR phenotype faithfully reflects UPP production. In the current study, we utilized this same CR-based screen, but performed it with the four different ECR mutants, each with in-frame deletions in the relevant regulatory genes (*visR*, *claR*, *pruA*, and *dcpA*). These strains were also genetically disabled for other known dispensable polysaccharides (cellulose, curdlan, β ,1-2 glucans, succinoglycan) but retained UPP production (UPP⁺); for ease of description we generally designate these strains as ECR (CDGS⁻) derivatives (the specific deleted regulator resulting in the ECR phenotype is indicated where relevant).

Mariner (*Himar-I*) transposon mutant libraries of each strain were screened for colonies with a notable decrease in CR pigmentation (the DCR phenotype), reasoning that we would likely isolate mutations in UPP biosynthetic functions, as well as positive regulatory elements required for the ECR phenotype. Overall, we screened close to 50,000 colonies of these derivatives for the DCR phenotype. Once confirmed as DCR mutants, the transposon insertion site junctions were amplified and sequenced. From this initial screen 140 independent mutants were identified in 18 discrete annotated genes (Fig. 2, Table S2), most predicted to be involved in polysaccharide biosynthesis or clustered with such genes (other DCR mutants were in known or predicted regulatory functions, see Xu et al., 2013). All but one of the 18 disrupted loci were identified multiple times, and mutations in all but two genes were found in more than one of the mutants used in the screen, with nine genes identified in at least three different mutant backgrounds (Table S2). Subsequent analyses revealed that most of the genes disrupted by Mariner insertions in the DCR mutants, when deleted and tested in the wild-type genetic background, resulted in decreased or absent UPP production and decreased biofilm formation (Table S1). This validated the use of the ECR CDGS⁻ mutants to initially screen for UPP defects, and that these are consistently manifested when the corresponding genes are disrupted in a wild-type background. Of the genes identified, four were in the previously characterized *uppA-F* cluster (number of mutants); *uppB* (2), *uppC* (14), *uppD* (6), and *uppE* (16) for a total of 38 (27%) of the independent mutants (Fig. 2A). No DCR mutations were identified in *uppA* or *uppF*. The

identification of four of the genes in this cluster however, validated the ability of our DCR screen to detect mutants with diminished UPP biosynthesis.

Identification of two Wzx flippase homologs that impact UPP production

One of the required Wzx-Wzy pathway functions that had not been identified previously was a flippase (Wzx) homolog. The DCR screen identified two different Wzx-type genes (Fig. 2B, 2C; Tables S1, S2). Across all four mutant backgrounds used in the screen, 21 independent mutations were obtained in a Wzx-type flippase homolog, ATU_RS02360, that we now designate *uppX*. This gene encodes a 426 aa protein, with 10-12 predicted transmembrane helices, as is consistent with many flippase homologs (Hong et al., 2018). In-frame deletion of *uppX* in wild type C58 resulted in a dramatic loss of biofilm formation (Fig. 5A) and attachment, as well as UPP production as judged by WGA and DBA labeling (Fig. 5B). These phenotypes were fully complemented by expression of a plasmid-borne copy of *uppX*. Divergently oriented from *uppX* is ATU_RS02370 (Fig. 2B), encoding a Wzy-type protein predicted to function as a polysaccharide polymerase, but no DCR transposon mutants were obtained in this gene. Interestingly, DCR mutations in a second flippase homologue were also identified, ATU_RS16320, now designated *uppV* (Fig. 2C, Table S1). UppV is a 477 aa Wzx-type protein, also with 12 predicted transmembrane domains. Only two mutations in *uppV* were isolated, but one each in the *pruA* and *dcpA* mutants (Table S2). Deletion of *uppV* in wild-type C58 significantly decreased but did not entirely abolish biofilm formation (Fig. 5A, 6A), although weak attachment and no substantial WGA lectin binding was observed (Fig. 6B).

The *uppV* coding sequence terminates 8 bp upstream of another reading frame, ATU_RS16325, that we have designated *uppQ* (Fig. 2C). Fifteen independent DCR mutants were isolated with insertions in *uppQ*, which is oriented convergently to an adjacent unidentified reading frame (ATU_RS16330), strongly suggesting that *uppQ* itself is required for UPP production. UppQ is homologous to acetyltransferases of the GNAT family of enzymes (Vetting et al., 2005), and paralogous to UppA (ATU_RS06125) (Table S1). Deletion of *uppQ* in the wild type results in loss of WGA lectin binding and has a strong biofilm defect (Fig. 6A and 6B; Table S1) suggesting it plays a crucial role in UPP biosynthesis.

A gene cluster of multiple polysaccharide biosynthetic genes with roles in UPP biosynthesis and biofilm formation

Strikingly, 47 of the DCR mutants (33.6%) had transposon insertions in a contiguous gene cluster ATU_RS11550-RS11590, with mutations represented in all four screening backgrounds, and multiple independent mutations identified for all of these genes (Fig. 2D). Seven of these annotated genes have predicted polysaccharide-related functions, and the other genes encode hypothetical functions. Eight out of nine genes in this contiguous cluster were identified in the DCR screen, with the exception of ATU_RS11580, a hypothetical gene (Fig. 2D). The genes are organized into at least four different predicted transcription units. Multiple lines of evidence for most of these genes indicate roles in UPP biosynthesis and hence we have designated them *uppG* through *uppO* (Fig. 2D, 5, Table S1). Four of these genes are predicted to encode cytoplasmic glycosyl transferases (GTs) *uppG* is in the

Carbohydrate Active Enzymes (CAZy) group GT4, *uppI* and *uppJ* in CAZy group GT2, and *uppL* in CAZy group 26 (Lombard et al., 2014). The *uppH* gene encodes a predicted periplasmic Class 10 glycoside hydrolase, *uppM* encodes a periplasmic co-polymerase (PCP-2b) with a GumC domain, and *uppO* encodes a phosphomannomutase. One of the interrupted genes among the DCR transposon mutants, *uppK* encodes a hypothetical protein of unknown function.

Independent in-frame deletions were generated in all of these predicted genes in the wild-type C58 background and tested for biofilm formation defects. Profound defects were observed for *uppG*, *uppI*, *uppJ*, *uppL*, and *uppM* (Fig. 6A and Table S1). Ectopic expression of the deleted gene complemented the deletion mutants for *uppG*, *uppI*, *uppJ*, *uppL*, and *uppM* (Table S1, Fig. S1A). The *uppM* deletion was based on a prior genome annotation of *A. tumefaciens* C58 (BioProject Access No. PRJNA283) which has now been corrected for *uppM* (ATU_RS11585) extending its 5' end by 198 bp for a full length of 1629 bp (BioProject Access. No. PRJNA224116). The *uppM* deletion we generated removes the entire PCP-2b (GumC) domain retaining the 198 bp upstream segment, and results in a severe biofilm defect that is well complemented by ectopic expression of *uppM* (Table S1). In-frame deletion of *uppH* did not have an impact on attachment, but the gene overlaps by 4 bp with the *uppG* coding sequence and multiple lines of evidence suggest that the *uppH* transposon mutations are polar on *uppG*. However, subsequent analysis revealed an important activity of UppH relevant to the UPP (described below).

In frame deletion of *uppO* in the wild-type C58 background did not result in a biofilm defect and in fact seemed to result in elevated adherence (Fig. 6A), despite the isolation of four DCR transposon mutants with insertions in this gene. Similar to *uppO*, in-frame deletion of the hypothetical gene *uppK* in wild-type C58 did not decrease biofilm formation, despite the isolation of four transposon insertions that map to this gene and result in ECR phenotypes (Fig. 6A).

The UppG glycosyl transferase catalytic site is required for full UPP synthesis.

Careful examination of the *uppG* mutant revealed that it does in fact bind to the WGA lectin, but that the signal is significantly weaker than the wild type (Fig. S1A, S1B). We recreated this deletion in the *pruA* CDGS- mutant and found it to be fully complemented by a plasmid-borne copy of the gene as evaluated by CR colony staining (Fig. S1C). GT4-Class glycosyltransferases similar to UppG have a conserved DXD catalytic motif (Breton et al., 2006). Mutation of the predicted DXD active site motif (pos. 142-4) at the first D residue (DXD>AXD) abolished complementation (Fig. S1C), confirming a requirement for its GT activity in this and related phenotypes.

UppH encodes a glycoside hydrolase capable of destabilizing the UPP

The *uppH* gene (ATU_RS1155) was found to be largely dispensable for adhesion and UPP production (Fig. 6A, Table S1). However, it was predicted to be a periplasmic Class 10 glycoside hydrolase, and its position within the *uppG-O* cluster implicated it in UPP production. Periplasmic glycoside hydrolases are often encoded within polysaccharide biosynthetic gene clusters and are thought to hydrolyze glycosidic bonds of the associated

polysaccharide as it transits this cellular compartment (Oglesby et al., 2008). Recently it was reported that exogenous addition of the purified PelAh and PslGh periplasmic glycoside hydrolases from *Pseudomonas aeruginosa* can disrupt Pel and Psl-dependent biofilms through cleavage of these polysaccharides (Baker et al., 2016). A derivative of UppH with its N-terminal secretion signal replaced by a hexahistidinyl tag (His₆) was purified. His₆-UppH was added in standard culture media (1 μM final concentration) to coverslips on which *A. tumefaciens* (*pruA* CDGS⁻) had been allowed to adhere for 2 h, and incubated for an additional 1 h, followed by dual lectin labeling with both af-WGA and fl-DBA (DBA also effectively labels the UPP as described in more detail below and in Xu et al. 2012). In the untreated samples, many cells were polarly attached and visibly labeled with one or both lectins (Fig. 7A). Addition of the UppH enzyme to these attached populations did not dislocate the cells but removed visible lectin-labeled polysaccharide. In parallel, we generated and purified a site-specific UppH mutant in which two predicted catalytic glutamate codons were mutated to glutamine codons to create UppH** (E152Q E266Q). Addition of exogenous, purified UppH** at 1 μM had no impact on lectin labeling (Fig. 7A).

We also examined UppH activity using Transmission Electron Microscopy (TEM). The *A. tumefaciens pruA* CDGS⁻ strain was used for these experiments to eliminate additional polysaccharides that can obscure visualization of the UPP and because the surface-contact dependence of wild type *A. tumefaciens* makes TEM of adhered cells challenging. We found that pre-incubation of *A. tumefaciens pruA* CDGS⁻ strain with the same concentration of purified UppH, effectively removed unipolar fibrous material observed for the untreated cells (Fig. 7B), resulting in poles with no visible fibers, whereas incubation with the UppH** mutant did not affect these fibers. Similar to studies using the *P. aeruginosa* glycoside hydrolases (Baker et al., 2016), addition of increasing concentrations of purified His₆-UppH added at the beginning of cultivation inhibited *A. tumefaciens* biofilm formation (Fig. S2), whereas the catalytic site mutant had no effect (data not shown).

A predicted outer membrane protein required for UPP synthesis

Nine independent DCR mutants had transposon insertions within or immediately upstream of the ATU_RS1765 gene, and we have designated this gene as *uppP* (Fig. 2E). One of the nine transposon insertion mutants mapped to the short annotated 318 bp upstream hypothetical coding sequence ATU_RS1760 which terminates 6 bp upstream of the predicted *uppP* start codon. This single transposon insertion is likely polar on *uppP*. UppP is 497 aa in length, has an N-terminal secretion signal and encodes an outer membrane β-barrel-type II porin protein (Table S1). UppP is much larger than the predicted polysaccharide secretin UppC (190 aa), does not match well with Wza homologues, and unlike UppC, lacks a lipidation motif. The *uppP* gene is convergently oriented to a GalU homologue (ATU_RS17570), predicted to encode a UTP-glucose-1-phosphate uridylyltransferase, that could synthesize a UDP-glucose precursor, but no DCR mutants were isolated that had insertions in this gene (although it is implicated by our subsequent findings, described below). An in-frame deletion of *uppP* in wild type resulted in complete loss of biofilm formation and was fully complemented by plasmid-borne expression of the

gene (Fig. 6A, Table S1). The *uppP* mutant does not produce UPP as evaluated with af-WGA staining (Fig. 6B).

Mutations in the *exoC* gene result in the DCR phenotype

Eight independent DCR mutants were found to have transposon insertions within ATU_RS19035 (Fig. 2F), encoding a gene previously designated as *exoC* (Kamoun et al., 1989). The *exoC* gene product encodes a predicted phosphoglucomutase that would catalyze conversion of glucose-6-phosphate to glucose-1-phosphate, a precursor of sugar nucleotide substrates for polysaccharide biosynthesis. The *exoC* mutant has pleiotropic defects in multiple polysaccharide biosynthetic pathways as well as motility, in addition to the defects in UPP production documented here (see Supplementary Results for additional details).

Apparent redundancy for Wzy-type polymerase function(s)

The DCR screen described above revealed a large number of genes required for UPP biosynthesis. Given the wide range of polysaccharide-related functions identified, it was surprising that no DCR mutants were isolated with mutations in Wzy-type polymerase homologs, despite the presence of a gene encoding a Wzy-type protein (ATU_RS02370) oriented divergently from *uppX* (Fig. 2B). We reasoned that the crucial Wzy polymerase activity required for a Wzx-Wzy-type pathway, might be fulfilled by functionally redundant proteins. Examination of the *A. tumefaciens* C58 genome revealed a Wzy-type protein paralogous to ATU_RS02370, ATU_RS11495, to be the closest match, (37% identity, 57% similarity). As described below, we have designated these genes as *uppY* (ATU_RS02370) and *uppW* (ATU_RS11495) (Fig. 8A). Single, in-frame deletions of *uppY* and *uppW* did not result in any discernable change in CR colony pigmentation in the *pruA* CDGS– mutant nor adhesion defects when deleted in wild type C58 based on *in vitro* biofilm assays (Fig. 8B, 8C). However, combination of the *uppY* and *uppW* deletions resulted in the DCR phenotype in the *pruA* CDGS– mutant (Fig. 8B) and abolishment of biofilm formation in the wild-type strain (Fig. 8C). Independent expression of plasmid-borne copies of either gene complemented the *uppY uppW* mutant for adhesion.

The UPP adhesin is composed of two chemically and genetically distinct polysaccharides

The large number of putative UPP genes we identified from the DCR screen was striking given an apparent duplication of function for many of the requisite steps involved in Wzx-Wzy-type polysaccharide synthesis (Fig. 2, Table S1). Previously, *A. tumefaciens* UPP synthesis was examined using multiple lectins with different oligosaccharide specificities (Xu et al., 2012). In addition to GlcNAc residues recognized by WGA, the lectin from the legume *Dolichos biflorus* (DBA) has specificity for *N*-acetylgalactosamine residues (GalNAc) and effectively labels the *A. tumefaciens* UPP. We performed dual labeling experiments with af-WGA and fl-DBA of *A. tumefaciens* (*pruA* CDGS–) cells adhered to coverslips. Both lectins labeled the poles for a large proportion of adhered cells (Fig. 9). The pattern of labeling was striking with the fl-DBA fluorescence forming a tight focus at the point of contact, and the af-WGA labeling more diffusely bound in a second ring around the DBA-labeled sector. This suggested that there were two discrete populations of polysaccharide at the pole, one identifiable using the GlcNAc-specific WGA and a second

labeled with the GalNAc-specific DBA. We tentatively designated these two polysaccharide species as UPP_{GlcN} and UPP_{GalN}, respectively.

Our findings suggested that the two Wzy-type proteins, UppY and UppW, were functionally redundant for adhesion and for CR labeling. Lectin labeling of the *uppY* and *uppW* mutants background revealed reciprocal specificity for each polymerase and a corresponding polysaccharide (Fig. 10A). In short-term coverslip binding assays probed simultaneously with both lectins, the *uppY* mutation caused loss of all fl-DBA labeling, but did not affect that of af-WGA, whereas the *uppW* mutation caused loss of the af-WGA labeling but retained that of fl-DBA. The double mutant, which abolished elevated CR binding and biofilm formation, was unable to bind either lectin. This suggested that production of the DBA-labeled UPP_{GalN} requires UppY, whereas elaboration of the WGA-labeled UPP_{GlcN} requires the function of UppW.

UPP polysaccharide production was also examined using transmission electron microscopy (TEM) and lectins conjugated with colloidal gold beads. The appropriate cultures were grown and probed with each lectin preparation individually, and then incubated with the general TEM reagent uranyl acetate. In the *A. tumefaciens pruA* CDGS- mutant with functional copies of both *uppY* and *uppW*, the UPP could be observed as fibrous material at a single pole of the cell, that was absent in the *uppY uppW* mutant (Fig. 10B). The gold-labeled WGA (20 nm) bound this fibrous material quite effectively and no binding was observed for the *uppY uppW* mutant. Fibrous polar material was visible for the single *uppY* and *uppW* mutants, but the labeled WGA only bound efficiently to the fibers observed in the *uppY* mutant, and not those on the *uppW* mutant (Fig. 10B). Initial experiments with the gold-labeled DBA lectin, revealed it to bind very inconsistently to the UPP (data not shown). We found that the lectin from peanut (PNA) that also binds to polysaccharides with GalNAc residues, gave more consistent results (Fig. 10B). The gold-conjugated (10 nm) PNA lectin bound to the fibrous material for the cells that were wild type for both *uppY* and *uppW*, and the *uppW* mutant, but there was no binding to the *uppY* mutant or the double mutant. Although this pattern was observed consistently, it was clear that the PNA lectin binds much less efficiently than does WGA to the UPP.

A second genetic screen to distinguish between UppY- and UppW-dependent UPP biosynthetic pathways

Our findings suggest that there are two distinguishable polysaccharides that comprise the UPP, UPP_{GalN} and UPP_{GlcN}, and that each polysaccharide species requires a specific Wzy-type protein, UppY and UppW, respectively. The individual *uppY* and *uppW* mutants do not manifest pronounced deficiencies in CR staining nor in standard biofilm assays. However, there are a large number of other single mutants in presumptive UPP biosynthetic genes identified that do manifest these deficiencies. For example, deletion mutants in *uppC* (Wza, OPX, outer membrane secretin), *uppX* (Wzx, flippase), *uppE* (WbaP, PHPT initiating glycosyltransferase) and *uppQ* (GNAT acetyltransferase), all have very pronounced phenotypes. This suggests that certain UPP genes have functions that are required for production of both polysaccharides, whereas others, most notably UppY and UppW, are specific to each polysaccharide. To identify UPP-related functions specific to

each polysaccharide and those that play a role for both, a variation of the initial DCR screen was developed. Neither the individual *uppY* or the *uppW* mutant has a DCR phenotype in the *pruA* CDGS⁻ background (Fig. 8B). We reasoned that additional loss of the other UPP polysaccharide in these single mutants would cause a significant DCR colony phenotype as observed for the *uppY uppW* mutant (Fig. 8B). Two strains were transposon mutagenized in parallel: the *A. tumefaciens pruA* CDGS⁻ strain deleted for either the *uppY* gene or the *uppW* gene (Fig. S4). We then screened these transposon mutant libraries for DCR mutants, examining roughly 30,000 colonies per lineage. We obtained and confirmed about 100 independent DCR mutants from each lineage. These mutants were pooled, total genomic DNA extracted and barcoded to specify each screening background, and then sequenced *en masse* using *Himar-I* transposon-specific amplification and Illumina sequencing ([dataset] Onyeziri and Fuqua, 2021). We expected to identify DCR mutations in the *uppY* gene for the *uppW* derivative, and mutations in *uppW* in the *uppY* derivative, and multiple mutants in these genes were isolated as predicted (Fig. 11, Table S3). Other previously identified genes from the initial DCR screen were also re-isolated. For example, multiple independent mutants in *uppE* (Atu1236) and *uppC* (Atu1238) were isolated in both mutant backgrounds. However, no transposon insertions in any of the other genes in the *uppABCDEF* cluster were isolated in either screen, despite the isolation of mutants in *uppB* and *uppD* in the prior screen (Fig. 11). The two flippase homologs, *uppX* and *uppV* were isolated in both screening backgrounds, although as observed before, considerably more mutants were isolated in *uppX*. Overall, multiple transposon mutations were re-isolated in both the *uppY* and *uppW* screening backgrounds in 12 different *upp* genes that had been identified in the prior screen, (Fig. 11, paired green and red bars; Table S3). Several other genes were identified in both screens that had not been previously associated with UPP biosynthesis, notably a *galE/exoB* homolog (ATU_RS19470), a *traA* plasmid conjugation gene homolog (ATU_RS22800, located on the linear chromosome), several predicted regulators (see below), and two genes of unknown function (ATU_RS15030 and ATU_RS22740).

Genes identified in only one of the polymerase mutants

The *uppG* and *uppH* genes were the only previously identified genes restricted to one of the two pathway-specific screens other than *uppY* and *uppW*. They were identified multiple times in the *uppY* mutant but not once in the *uppW* mutant (Fig. 11, single red bars, Table S3). As described above, UppG encodes a glycosyltransferase in the GT4 family, whereas *uppH* is a Class 10 glycoside hydrolase. Given the polarity of transposon insertions in *uppH* on *uppG* that we had already confirmed, it is very likely that the DCR phenotype in the *uppH* transposon mutants was due to their effect on *uppG* expression. Thus, *uppG* is probably the only previously identified UPP biosynthesis function in addition to the *uppW* polymerase that is required specifically for UPP_{GlcN} production and does not appear to play a prominent role in UPP_{GalN} production (at least in the *uppY* background).

Another polysaccharide related gene we have now designated *uppZ* (ATU_RS11500) encoded convergently from the *uppW* Wzy-type gene, was disrupted in three independent mutant sequences. UppZ is an Acyltransferase-3 (AT-3) homolog similar to ExoZ from *S. meliloti*, required for production of acetylated succinoglycan (Reuber & Walker, 1993). Two

additional genes were identified only in the *uppY* screen. Four mutants were isolated in ATU_RS05430, predicted to be a fatty acid biosynthesis gene. Nine independent mutations were isolated in the *podJ* gene (ATU_RS02460), a polarity determination factor known to be required for normal polar localization of the *C. crescentus* holdfast (Hinz et al., 2003) and required for functions related to polar cell division in *A. tumefaciens* (Anderson-Furgeson et al., 2016). No *podJ* mutants were identified in the *uppW* background.

Transposon mutants in 11 different genes were isolated specifically in the *uppW* mutant that resulted in loss of the CR phenotype imparted by the UPP_{GalN} polysaccharide (Fig. 11, single green bars, Table S3). Although these genes were independently disrupted in multiple DCR transposon mutants, the majority were isolated in fewer than five mutants. Two genes with transposon insertions from the *uppW*-specific screen are adjacent, and annotated as *rfbA* (ATU_RS21635) and *rfbD* (ATU_RS21640), encoding a glucose-1-phosphate thymidyltransferase and a dTDP-4-dehydrorhamnose reductase. They are located in an operon of genes predicted to be involved in synthesis of the dTDP-L-rhamnose, often a component of the LPS core or O-antigen in members of the *Rhizobiaceae* (De Castro et al., 2008). The *nuoL* gene (ATU_RS06310) and *clpA* (ATU_RS06730) encode a NADH ubiquinone oxidoreductase subunit, and the ATP-binding subunit of the Clp protease. Seven of the genes identified are hypothetical reading frames.

Genes involved in sugar-nucleotide precursor synthesis

Multiple independent mutants were isolated in the *exoC* gene (ATU_RS19035), in both screening backgrounds (Fig. 11, Table S3), consistent with the prior screen (Fig. 2F). As described earlier, *exoC* encodes a phosphoglucomutase, predicted to convert glucose-6-phosphate to glucose-1-phosphate, a precursor of sugar nucleotide substrates for polysaccharide biosynthesis (Kamoun et al., 1989). A gene annotated as *galE/exoB* (ATU_RS19470) was isolated with five or more independent transposon mutants in both pathway-specific screens but had not been isolated in the initial mutant screen. This gene is predicted to encode a UDP-glucose epimerase, similar to GalE, that catalyzes the reversible epimerization reaction of the sugar nucleotide precursor UDP-GlcNAc to UDP-GalNAc, and has an NADP Rossmann domain (Holden et al., 2003).

Additional regulators identified in the pathway-specific screen

Several probable regulatory genes were identified in the UPP_{GlcN} and UPP_{GalN} specific screen (Fig. 11, Table S3). Mutations in three of these were identified in both screens. The *exoR* gene (ATU_RS08400), encodes a periplasmic regulator of the ChvG-ChvI two-component system (Heckel et al., 2014). In *A. tumefaciens*, *exoR* mutations dysregulate hundreds of genes, including those for synthesis of the polysaccharide succinoglycan. Mutations in *exoR* were also isolated as DCR mutants in the original screen (Xu and Fuqua, unpublished results). The two other putative regulators identified had not been implicated from the earlier screen. In both the *uppY* and *uppW* backgrounds, mutations were isolated for a homolog of the *relA/spoT* gene (ATU_RS05105). Multiple mutations in both backgrounds were also isolated in a gene annotated as the two-component type sensor kinase *ntrY* (ATU_RS07130).

Conservation of UPP biosynthesis genes

We used translated sequences of the UPP genes identified in this study as queries to search for homologs among genome sequences from members of the *Rhizobiaceae*. It is recognized that rhizobia and agrobacteria are polyphyletic and intermixed (Slater et al., 2009). Members of the *Agrobacterium* genus were historically classified into three primary biovars (I, II, and III), based on multiple attributes, and many of the wide host range *A. tumefaciens* strains including C58 are in Biovar I (Morton & Fuqua, 2012b). Recent phylogenetic studies have also divided the *A. tumefaciens* group itself into over 20 separate genomospecies (Lassalle et al., 2011). *A. tumefaciens* C58, the type-strain we use in this study, is in Genomospecies 8 (G8, also given the species designation *A. fabrum*). There is now a wealth of genomic data for comparison across the agrobacteria and among the wider *Rhizobiaceae* that allows great insights into the genetic conservation and ecological context among these bacteria (Weisberg et al., 2020).

A comparison of the UPP genes identified in the current study across representative genome sequences reveals multiple interesting trends. Several of the genes are widely conserved in sequence and synteny (Figs. 12, S5, S6). First, the *uppA-F* gene cluster is conserved largely throughout the agrobacterial genomospecies, with *uppA* less well conserved but still present (Fig. 12 and S5). This gene cluster is fully or partially conserved across the *Rhizobiaceae*, including among bacteria such as *Rhizobium leguminosarum* and *Rhodopseudomonas palustris* that are both recognized to make a UPP-type structure (Fritts et al., 2017, Williams et al., 2008). Beyond the *Rhizobiaceae* the gene cluster is also partially conserved more widely in *Brucella abortus* and other mammalian pathogens in the Alphaproteobacteria group. The *uppA-F* cluster shares limited similarity to the holdfast synthesis (*hfs*) gene cluster from *C. crescentus* but is clearly divergent between bacteria in the Orders *Rhizobiales* and *Caulobacterales* (Fritts et al., 2017).

Other *upp* genes that are well conserved across the agrobacteria genomospecies are *uppV* and *uppQ*, the flippase and acetyltransferase encoded within a presumptive operon, and *uppP*, that encodes the outer membrane protein. The sugar nucleotide enzyme ExoC and a GalU homolog (ATU_RS17570) encoded convergently with *uppP* (but not identified in our genetic screens) are also well conserved across the agrobacteria and the rhizobia. Strikingly, the *uppG-O* and *uppXY* gene clusters, and *uppW* gene are well conserved within the genomospecies considered Biovar I (G1, G2, G3, G4, G5, G7, G7-2, G8, G9, G13), with a few exceptions for specific genes within the clusters, but are largely absent outside of these taxa.

Discussion

In this study, we identified and analyzed a large number of genes required for biosynthesis of the *A. tumefaciens* unipolar polysaccharide known as the UPP, and in so doing, have discovered that this structure is composed of two genetically and biochemically discrete polysaccharides.

Genetic screens provisionally identify UPP biosynthetic genes

One of the two separable biosynthetic pathways is UppW-dependent and drives synthesis of the UPP species that labels with the WGA lectin specific for GlcNAc residues (UPP_{GlcN}). The other pathway is UppY-dependent and directs production of the UPP species that labels with the DBA and PNA lectins specific for GalNAc residues (UPP_{GalN}). Our genetic approach identified multiple components of a presumptive Wzx/Wzy pathway including formation of sugar nucleotide precursors (Fig. 13), several regulators and a collection of genes of unknown function. These screens were remarkably specific for UPP deficiencies, and creation of in-frame deletions in many of the genes identified from the screen revealed clear defects in UPP-dependent phenotypes in wild-type and re-constructed screening backgrounds. For our first DCR screen in which the *uppY* and *uppW* genes were intact, the majority of genes identified were predicted polysaccharide-related functions (Fig. 2 and Table S1), with the exception of *uppP*, which has the sequence of an outer membrane protein with no predicted role in polysaccharide biosynthesis, and *uppK*, a gene of unknown function. Only three of the disrupted genes identified in DCR mutants within the genetic clusters predicted to encode Wzx-Wzy-type pathway components, failed to diminish UPP-dependent phenotypes when the corresponding in-frame deletions were generated (*uppK*, *uppO*, *uppH*) in otherwise wild-type *A. tumefaciens*. Transposon insertions in *uppH* were polar on the downstream glycosyltransferase *uppG*, although the evidence suggests the UppH glycoside hydrolase recognizes the UPP structure. The incongruity between the *uppK* and *uppO* transposon mutants and in-frame deletion mutants in these genes is however not explained by simple polarity, and the basis for these differences requires further investigation.

Initial steps in UPP precursor synthesis

We have identified all of the requisite components of a complete Wzx/Wzy-type polysaccharide biosynthesis pathway (Whitfield et al., 2020), including homologs to flippases (Wzx-type proteins, UppX and UppV) and polymerases (Wzy-type proteins, UppY and UppW) (Fig. 13). Our preliminary model is that UppE functions as the WbaP-initiating glycosyltransferase for UPP biosynthesis, or PHPT (polyisoprenylphosphate hexose-1-phosphate transferase), adding the first monosaccharide to an undecaprenol pyrophosphate (Und-PP) lipid carrier. Our prior studies found that under phosphorus (P_i) limitation, the requirement for *uppE* could be supplanted by another PHPT-type protein ATU_RS00495 (Xu et al., 2012), but under P_i-replete conditions as used in the current study, the role for UppE is clear. Interestingly, multiple PHPT homologs are redundantly functional in holdfast biogenesis in *C. crescentus* (Toh et al., 2008).

As with other polysaccharide biosynthetic pathways, precursors for this and subsequent monosaccharide additions are supplied by cytoplasmic sugar nucleotide pathways that in *A. tumefaciens* and its relatives, depend upon ExoC to convert glucose-6-phosphate (Glc-6-P) to glucose-1-phosphate (Glc-1-P) (Kamoun et al., 1989). As GlcNAc appears to be a major component of the UPP, we predict Glc-1-P is acetylated to form *N*-acetyl glucosamine-1-phosphate (GlcN-1-P) by one of the putative acetyl transferases UppA or UppQ (Fig. 13), mutations in either of which cause severe UPP defects. The GlcNAc would be conjugated to UDP, to form UDP-GlcNAc, although it is possible that UDP linkage occurs

prior to acetylation. Our genetic screens did not identify any UTP-glucose-1-phosphate uridylyltransferases as DCR mutants, but the genes annotated *exoN* (ATU_RS18920) and *galU* (ATU_RS16570, encoded adjacent to *uppP*), are both predicted to catalyze this reaction and are non-essential (Curtis & Brun, 2014), so it is possible they are functionally redundant for this activity. The GalE homolog (ATU_RS19470) identified with a general UPP defect in our pathway-specific screen (Fig. 11, 13) is predicted to reversibly epimerize UDP-GlcNAc to UDP-GalNAc, an interesting activity since these sugar moieties are likely to be defining residues for the UPP_{GlcN} and UPP_{GalN} polysaccharides, respectively. Furthermore, sugar-nucleotide precursors for UPP synthesis are likely to be generated through these basal sugar-nucleotide pathways, although no additional genes annotated to function in this capacity were identified.

Following the Wzx-Wzy general model, after addition of the initiating monosaccharide to the Und-PP carrier to form monoglycosylated Und-PP, other glycosyltransferases utilize sugar-nucleotide precursors to add monosaccharides to form the mature subunit (Whitfield et al., 2020). Generally, the first of these monosaccharide additions is mediated by a glycosyl transferase homologous to WecG/TagA, generating the diglycosylated lipid II intermediate, which can then be acted upon by additional glycosyl transferases (Ginsberg et al., 2006). Five predicted glycosyltransferases were identified in one or both of our genetic screens (*uppD*, *uppG*, *uppI*, *uppJ* and *uppL*). UppL is predicted to encode a glycosyltransferase (CAZy Group GT26) homologous to WecG and TagA, and is strictly required for UPP synthesis. It is also the best match to the HfsJ glycosyl transferase of *C. crescentus*, a WecG/TagA type protein required for holdfast synthesis and a point in the pathway at which holdfast production is regulated (Fiebig et al., 2014). We hypothesize that UppL generates the Lipid II diglycosyl intermediate for UPP biosynthesis. The four additional GTs can then act to assemble the UPP repeating subunit (Fig. 13).

The *uppI*, *uppJ*, and *uppL* genes clearly all play roles in UPP synthesis, as they were consistently identified in our DCR mutant screens and in-frame deletions have significant impacts on UPP production and attachment. The *uppD* gene was identified as a DCR mutant in our initial screen with *uppY* and *uppW* intact (in all four ECR screening backgrounds) and the *uppD* mutant is clearly biofilm deficient and greatly diminished for UPP production. However, it was not identified in the pathway-specific screen, suggesting that the requirement for UppD may be less stringent. Given our evidence for a UPP_{GlcN}-specific role for UppG, we propose that it adds a particular residue that acts to divert a portion of the growing subunits to the UppW-dependent pathway, whereas UPP subunits in which this additional residue is not incorporated proceed through the UppY-dependent pathway. Additional modifications to the UPP subunits may include acetylation given the three potential acetyl transferases we have identified, UppA, UppQ, and UppZ.

Translocation, polymerization and export of the UPP.

A defining step in Wzx-Wzy pathway is translocation of the mature repeating subunit linked to Und-PP from the cytoplasmic face of the inner membrane to its periplasmic face (Islam and Lam, 2014). Mutants in *uppX*, encoding a predicted flippase with 10-12 transmembrane domains, have a severe UPP deficiency and were identified repeatedly in our screens. In

addition, the *uppV* gene also encodes a flippase homolog, that was consistently identified (albeit less frequently than *uppX*) and also manifests a pronounced UPP defect. Our initial prediction that each flippase would be specific for one of the two polymerases did not hold true, as each flippase mutant manifests defects for both UppW- and UppY-dependent polysaccharides. We therefore tentatively assign both UppX and UppV as flippases in our model, with perhaps UppX playing the primary role (Fig. 13).

Once in the periplasm, the UPP subunits initiate or add to the growing polysaccharide chain by the activity of either UppW or UppY, which are strictly required for UPP_{GlcN} and UPP_{GalN}, respectively (Fig. 13). We predict that the specific monosaccharide composition of the subunit dictates which polymerase drives chain elongation. Addition of the subunit to the growing chain associated with its cognate polymerase, releases the Und-PP which is then available to serve for the multiple other pathways that depend on it, such as for synthesis of peptidoglycan and other polysaccharides.

Based on the Wzx/Wzy paradigm, the UPP polysaccharide then associates with a periplasmic co-polymerase (PCP, Wzc) and is ushered to the outer membrane polysaccharide secretin (OPX, Wza) for export to the cell exterior (Whitfield et al., 2020). The details of this process and the specific role for the PCP remain ill-defined for any biosynthetic pathway, but it is clear that the PCP and OPX proteins form a large multimeric complex that spans the outer membrane, and the periplasm, projecting into the cytoplasm (Collins et al., 2007). Our genetic analysis suggests that the predicted PCP-2a proteins, UppM and UppB, fulfill the PCP (Wzc) requirement for both UPP polysaccharides and that UppC functions as the OPX (Wza) secretin (Fig. 13). Other important processing may occur in the periplasm, such as the activity of the periplasmic glycoside hydrolase UppH (although this is not required for the UPP-dependent phenotypes we analyzed) and perhaps anchoring of the UPP polysaccharide to an outer membrane component, via the O-antigen ligase homolog UppF. What retains the externalized UPP_{GlcN} and UPP_{GalN} at discrete locations at a single pole remains to be determined. Our preliminary results suggest that several of the UPP biosynthetic proteins identified here localize to the same pole as the UPP (Natarajan and Fuqua, unpublished) suggesting that at least some of the steps of biosynthesis occur proximal to the eventual site of UPP elaboration, similar to what has been observed for holdfast biogenesis in *C. crescentus* (Javens et al., 2013). A role for the PodJ polar localization factor in proper positioning of the UPP to the pole is supported by its identification in our genetic screens, also consistent with its function in *C. crescentus* holdfast localization. The reason that it was only identified in the UPP_{GlcN}-specific screen in *A. tumefaciens* is the subject of future studies.

Identification of new polysaccharide genes and regulators.

The pathway-specific screen identified a significant number of genes in which multiple independent transposon insertions resulted in the DCR phenotype indicative of UPP deficiencies, that had not been identified in the initial screen in which *uppY* and *uppW* were intact. Some of these genes are annotated to be related to polysaccharide biosynthesis (*rfbB*, *rfbD*, and *uppZ*). It is unclear what role the other genes with annotations either that predict functions outside of polysaccharide synthesis or hypothetical genes, might play in UPP

synthesis. Multiple transposon mutants were isolated in the predicted regulators *relA* and *ntrY* in both pathway-specific screens but not the original screen. This pattern suggests that their roles may be related to the specific defects caused by the *uppY* and *uppW* polymerase mutations. RelA is a (p)ppGpp synthetase/hydrolase best known from its role in the stringent response to nutrient limitation (Ronneau and Hallez, 2019). NtrY is a two component-type sensor kinase thought to function with the NtrX response regulator, which in *A. tumefaciens* is encoded immediately downstream of the *ntrY* gene (Stein et al. 2021). Both of these regulators control multiple functions in other APB, and it seems likely that UPP production overlaps with their regulatory networks in *A. tumefaciens*. Although the isolation of multiple independent mutants in these genes supports their impact on UPP synthesis, more extensive analysis is required to confirm their mechanisms and the extent to which they impact UPP production.

Anchoring and cohesiveness of the UPP

Interestingly, individual in-frame deletions of all the *uppA-F* genes result in UPP deficiencies although no mutations were identified for *uppA* or *uppF* in any of the DCR genetic screens performed. UppA encodes a cytoplasmic GNAT-type acetyl transferase. For several bacterial polysaccharides, the degree and distribution of acetylation is thought to alter the adhesiveness of the polysaccharide (Vuong et al., 2004, Wan et al., 2013). UppA may modulate adhesiveness of the UPP, but its loss may not visibly impact the CR-staining properties of the UPP. There is a strict requirement for the other cytoplasmic GNAT-type protein UppQ for the ECR phenotype, and so we speculate that UppA performs a non-critical role such as UPP modification, but we do not know with certainty, so both UppA and UppQ are depicted to modify the UPP (Fig. 13). The *uppZ* gene identified in the pathway-specific screen performed with the *uppY* mutant, also encodes a predicted acetylase similar to ExoZ that may function on periplasmic polysaccharide but this was only identified in the *uppY* mutant suggesting specificity for the UPP_{GlcN} polysaccharide.

UppF has homology to O-antigen ligases that function in attaching the O-antigen portion of LPS to the LPS core sugars during biogenesis of this outer membrane component (Ruan et al., 2018). The dramatic biofilm and attachment deficiencies for the *uppF* mutant suggest a profound UPP defect, but it was not identified in either of the DCR screens. In *C. crescentus*, the holdfast is attached to the tip of the stalk via the holdfast anchor (Hfa) proteins, that are thought to form a β -amyloid-type fiber (Hardy et al., 2010). However, there are no clear homologs of the Hfa proteins encoded by *A. tumefaciens* and we speculate that instead the UPP may be linked to LPS or another component on the cell surface. Given its combination of phenotypes, UppF is a promising candidate for linking the UPP to the cell surface.

The polysaccharide component of the *A. tumefaciens* UPP

As with the *C. crescentus* holdfast, the UPP structure was first detected using lectin reagents that are generally specific for polysaccharides (Tomlinson & Fuqua, 2009). Lectin labeling remains the most facile and reliable way to identify the *A. tumefaciens* UPP microscopically, although it is also visible in TEM and SEM without specific labeling (Xu et al., 2013). In addition to the genes encoding functions related to polysaccharide synthesis, we have also identified several regulators and the polar localization factor PodJ as

important for UPP biosynthesis. Our pathway-specific screen identified a significant number of genes of unknown function, which may or may not be related to polysaccharides, and remain to be investigated further. The UppP protein, predicted to encode an outer membrane porin (distinct from polysaccharide secretins such as UppC), is not a typical component of a Wzx/Wzy pathway, but is required for UPP production. Despite these few exceptions, the prominent role for Wzx/Wzy proteins and other polysaccharide related proteins (e.g. sugar nucleotide precursors) strongly support the conclusion that the UPP is predominantly comprised of polysaccharide.

We have added to this evidence by finding that the extracellular addition of the purified UppH glycoside hydrolase degrades all visible lectin-binding UPP material on pre-attached cells and can inhibit biofilm formation when co-incubated with planktonic *A. tumefaciens*. The UppH protein is annotated in the *A. tumefaciens* C58 genome as a xylosidase (Goodner et al., 2001, Wood et al., 2001), and a recent study, tentatively designating the gene *xynA*, found that it was required for growth on minimal medium with exogenous xylan as a carbon source (Mathews et al., 2019). Our observations on susceptibility of both UPP_{GlcN} and UPP_{GalN} to UppH (XynA) activity suggest that the enzyme hydrolyzes important glycosidic linkages common to both UPP polysaccharides. Based on its activity on exogenous xylan and its annotation, we hypothesize that these vulnerable positions within the UPP contain xylose residues. A recent chemical characterization of the *C. crescentus* holdfast revealed that it contains xylose residues, among other monosaccharides (Hershey et al., 2019). The *uppH* mutant does not have a detectable attachment phenotype, but this is a relatively common observation among periplasmic glycoside hydrolases associated with polysaccharide synthesis pathways (Baker et al., 2016). The roles for these enzymes in polysaccharide biosynthesis remain ill-defined despite their utility as purified reagents with high specificity for their cognate polysaccharides, and their anti-biofilm potential. We propose that UppH is a component of the UPP biosynthetic pathway that is active on the polysaccharide component of this structure. Interestingly, although exogenous UppH incubation clearly removes the UPP as visualized with lectin labeling and TEM, pre-attached cells remain on surfaces after treatment, and there is incomplete biofilm inhibition in the presence of UppH. This may reflect a curing process of the UPP, as has been observed for the *C. crescentus* holdfast (Hernando-Perez et al., 2018) to a glycosidase resistant, sterically blocked form, or alternatively the function of an unidentified component that is resistant to the UppH enzyme.

Given the number of genes encoding Wzx-Wzy components that are required for both UPP_{GlcN} and UPP_{GalN} synthesis, and their mutual susceptibility to the UppH glycoside hydrolase, we hypothesize that they have closely related structures and perhaps a common core motif. Two related UPP polysaccharides would be similar to the *Vibrio cholerae* VPS polysaccharide which occurs as two distinct species, one which contains α -D-Glucose among other monosaccharides, and another less abundant species where this residue is α -D-GlcNAc (Yildiz et al., 2014). The lectin specificity of the two UPP species suggests that one species contains GlcNAc and the other GalNAc, but this remains to be proven by chemical characterization of the two different polysaccharides. Interestingly, the only annotated biosynthetic genes other than the polymerase genes *uppW* and *uppY* that are specific for each polysaccharide, are (i) *uppG*, encoding a glycosyltransferase, (ii) *uppZ*,

an *exoZ* homolog (both of which are involved in UPP_{GlcN} synthesis but not UPP_{GalN}), and (iii) potentially the *rfaA* and *rfaD* homologs specific for UPP_{GalN} and not UPP_{GlcN}. The observation that the *uppG* mutant synthesizes a very small UPP, but still detectable with WGA, suggests that it is not solely responsible for addition of GlcNAc residues. We speculate that UppG may either contribute to addition of GlcNAc residues, or other monosaccharide residues that then dictate the efficiency of polymerization by UppW, and eventual elaboration of UPP_{GlcN}. Wzy-type polymerases can be highly specific, and subtle differences in the linkages and monosaccharides can dictate recognition between particular Wzy-type proteins (Hong et al., 2015). The UppZ protein is similar to ExoZ from *S. meliloti* and is predicted to be an AT-3 (Pfam PF01757), acyltransferase with 10 transmembrane domains. Proteins with the AT-3 domain are known to be required for acetylation of O-antigens in the periplasm during LPS biosynthesis (Pearson et al., 2020), and we speculate that UppZ may be acting to modify the UPP_{GlcN} polysaccharide in the periplasm during its synthesis (Fig. 13).

Physical separation of UPP_{GlcN} and UPP_{GalN} at cell poles

The sectored, bullseye pattern observed for lectin labeling when cells producing the UPP are incubated with both the DBA and WGA lectins suggests that the UPP_{GlcN} and UPP_{GalN} polysaccharide species are localized differently relative to each other (Fig. 9). The DBA labeling is usually proximal to the tip of the cell, whereas the WGA labeling, while still unipolar, is more diffuse and visibly surrounding the DBA-labeled material. This pattern suggests that these polysaccharides may be localized to specific positions around the pole. Alternatively, they may be extruded from the same point on the pole but temporally synthesized in stages. How the roles for these multiple polysaccharide species are integrated during the surface attachment process, and how their biophysical properties may differ are current topics of investigation.

Mixed conservation among the UPP biosynthesis genes

The six gene *uppABCDEF* cluster we initially identified is fully or partially conserved across members of the *Rhizobiales* (Fritts et al., 2017). Among the expanded UPP biosynthesis genes we have identified here, some show similar wide distribution, such as the *uppVQ* genes. However, several genes are restricted to the Biovar I group including the G8 genomospecies of *A. tumefaciens* C58. This is intriguing as many of the genes in these clusters encode functions that are required for both UPP_{GlcN} and UPP_{GalN} such as the *uppX* flippase and the *uppL* glycosyltransferase. Conversely, the *uppY* and *uppW* genes encode proteins that have specificity for UPP_{GalN} and UPP_{GlcN}, respectively. *R. palustris* produces a UPP-type structure that labels with WGA and presumably contains GlcNAc residues, but only the *uppA-F* cluster is conserved (with the *uppC*-type secretin in a different position), and a flippase has yet to be identified (Fritts et al., 2017). In *C. crescentus*, the holdfast was initially defined by WGA binding and is now known to contain GlcNAc, xylose, and other monosaccharides (Hershey et al., 2019), and *A. tumefaciens* homologs to the *hfs* genes reside in the *uppA-F* cluster, as well as others outside, such as the glycosyl transferases *uppJ* and *uppL*, which are similar to *hfsL* and *hfsJ*, respectively. In *R. leguminosarum*, the *uppA-F* cluster (designated as *gms* genes for glucomannan synthase) is conserved and syntenous with that from *A. tumefaciens*, but the glucomannan unipolar polysaccharide it directs

production of is composed of glucose and mannose and is recognized by host pea plants via a specific lectin (Laus et al., 2006, Williams et al., 2008). Clearly, diversification of these genes has led to production of different polysaccharides, with niche-specific functions. Over evolutionary history, several *A. tumefaciens* lineages within the Biovar I type strains have apparently bifurcated their UPP into two different chemical species, UPP_{GlcN} and UPP_{GalN}, by the activity of an expanded and intertwined gene set that encodes their synthesis. The UPP genes in general provide a fascinating example for diversification of cell surface polysaccharides, features that are arguably very much exposed to environmental selective pressures. The specific roles for the expanded UPP genes in the ecological context of the *A. tumefaciens* Biovar I lineages, including but not restricted to plant interactions, is an area of future interest.

Experimental Procedures

Strains, plasmids, and reagents.

The bacterial strains and plasmids used in this study are listed in Table S4. Oligonucleotide sequences are available upon request. *A. tumefaciens* strains were grown at 28°C or 30°C on liquid or solid LB (10 g of Bacto-tryptone, 5 g of yeast extract and 5 g of NaCl per liter, pH 7.0) or AT minimal medium (Tempé et al., 1977) with 0.5% (w/v) glucose and 15 mM ammonium sulfate to make ATGN. For biofilm assays, 22 µM FeSO₄ was added to the liquid ATGN medium immediately before inoculation, and for *sacB* counter-selection, 2.5-5% (w/v) sucrose replaced glucose as the sole carbon source to make ATSN. For ATGN-CR, the CR dye was dissolved in Milli-Q water (Millipore Sigma, Burlington, MA) to a concentration of 20 mg/ml, then passed through a 0.2 µm filter to remove aggregates. 250 µl of this stock solution was added per 100 ml ATGN solid medium to make a final concentration of 50 µg/ml ATGN-CR. *E. coli* strains were grown at 37°C in LB medium. Motility assays were performed in 0.3% ATGN agar in 60 mm plates, inoculated in the center of the medium with the tested strain, and incubated at 30° C. Plasmids were introduced into chemically competent cell preparations of *E. coli*, and into *A. tumefaciens* via conjugation or electroporation (Morton & Fuqua, 2012a). Oligonucleotide primers were obtained from Integrated DNA Technologies, Coralville, IA, and single-primer extension DNA sequencing was performed by ACGT, Inc., Wheeling, IL. Chemicals, antibiotics, and culture media were obtained from Fisher Scientific and Sigma-Aldrich. When necessary, antibiotics were added to the medium as follows: 100 µg/ml ampicillin (Ap), 50 µg/ml gentamycin (Gm) and 50 µg/ml kanamycin (Km) for *E. coli*, and 300 µg/ml Gm and Km for *A. tumefaciens*. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to 250-500 µM to induce expression of plasmid borne genes. All fluorescently conjugated lectins were obtained from Invitrogen; FITC-labeled WGA (fl-WGA), Alexa Fluor 594-labeled WGA (af-WGA), FITC-labeled DBA (fl-DBA),.

Controlled expression plasmids.

Wild type coding sequences of genes to be expressed were amplified from *A. tumefaciens* C58 genomic DNA (unless otherwise stated), and cloned into the LacI^Q encoding, IPTG inducible plasmid pSRK carrying either a gentamycin (pSRKGm) or kanamycin (pSRKKm) resistance cassette (Khan et al., 2008) using primers corresponding to each gene either by

amplification with Phusion or Q5 polymerases (NEB) and ligation or isothermal assembly. Amplicons were introduced into the appropriately cleaved plasmid and verified by PCR, restriction enzyme digestion and sequencing.

Site specific mutagenesis.

A slightly modified QuikChange mutagenesis protocol from Stratagene was used to engineer site specific mutations in *uppG* and *uppH*. Briefly, two complementary primers were designed with the required base pair mutation(s) flanked by ~15 nt wildtype sequences on either side. A PCR reaction (16 cycles) using NEB Phusion polymerase was used to generate linearized plasmid derivatives with the modified sequences and the amplicons were recircularized by ligation. Correct amplification was verified via gel electrophoresis. 1 μ l of the methylation-dependent restriction enzyme *DpnI* was used to digest the methylated wildtype plasmid. Digested product was transformed into *E. coli* and plasmid clones were purified and sequenced to confirm the presence of the desired mutation.

In-frame markerless deletions.

In-frame markerless deletions were constructed as previously described (Morton & Fuqua, 2012a, Xu et al., 2013). Upstream and downstream ~ 500 bp regions flanking the genes to be deleted were amplified from *A. tumefaciens* C58 genomic DNA using specific primers (P1 and P2 for upstream regions, P3 and P4 for downstream regions) and NEB Phusion polymerase. Care was taken to ensure the 5' and 3' ends of the genes were not altered, and that translational coupling between linked genes was retained. Primers P2 and P3 were designed with reverse complementarity in the 5' sequences allowing for single overlap extension (SOE) amplification of both genes into a single amplified product as described (Morton & Fuqua, 2012a). The combined flanking regions were ligated into the appropriately cut suicide vector pNPTS138 using isothermal assembly with the NEBuilder HiFi DNA Assembly Cloning Kit to create a deletion construct which was subsequently transformed into chemically competent S17-1/ λ pir *E. coli*. The pNPTS138 vector confers kanamycin resistance (Km^R) and sucrose sensitivity (Suc^S) due to the presence of *sacB*. Introduction of this plasmid into *A. tumefaciens* was via conjugation with S17-1/ λ pir *E. coli*. The ColE1 origin of pNPTS138 does not replicate in *A. tumefaciens*, and a single crossover event allows plasmid integration into the chromosome, which was confirmed by patching transformants on to ATGN-Km and ATSN to screen for $Km^R Suc^S$ strains. Strains that had excised the plasmid were isolated by parallel patching on to ATSN and ATGN-Km to screen for $Km^S Suc^R$ derivatives. Deletion of the target genes was confirmed by diagnostic PCR using primers (P5 and P6) flanking the site of deletion.

Congo red growth assays.

A. tumefaciens strains were grown in ATGN until mid-exponential phase, normalized to an OD_{600} of 0.05, and 2 μ l spotted on to ATGN-CR plates with or without IPTG. Plates were incubated at 30°C. Photographs of spots were taken following incubation for 48 h.

Biofilm assays.

Strains of *A. tumefaciens* were grown and analyzed as previously described (Morton & Fuqua, 2012c). Briefly, PVC coverslips were inserted vertically into the wells of 12-well polystyrene plates (Corning Inc.) and UV sterilized. The wells were subsequently filled with 3 ml mid-exponential phase cultures (with or without IPTG) normalized to OD₆₀₀ 0.05. The 12-well plates were incubated in static conditions for 48 or 72 h at room temperature or 30°C inside a plastic Tupperware containing an open bottle of saturated K₂SO₄ solution to maintain humidity. Following incubation, the coverslips were rinsed with sterile water and stained with 0.1% w/v crystal violet (CV) solution 0.1% w/v for 20 min. Stained coverslips were rinsed with sterile water and then submerged in weigh boats containing 1 ml 33% acetic acid to solubilize the CV. Quantification was performed by measuring soluble A₆₀₀ CV absorbance using a Biotek Synergy HT microplate reader and Gen 5 software (Biotek, Winooski, VT). Values were normalized to culture growth by dividing by final culture OD₆₀₀.

UPP production assays.

UPP was visualized using the WGA lectin conjugated to Alexa Fluor 594 (af-WGA, Invitrogen) and with DBA lectin labeled with FITC (fl-DBA). 1 ml of mid-exponential phase *A. tumefaciens* cultures were spun down (7000 x g), washed twice with ATGN, pelleted, and resuspended in 100 µl ATGN. 1 µl of 1 mg/ml the fluorescent lectin stock was added to the suspensions and incubated at 28°C for 20-40 min. 1-2 µl labeled suspensions were placed on microscope slides containing agarose pads (1-1.5% agarose, 100-200 µl). Coverslips were placed on top of the pads, and samples were visualized by fluorescence microscopy using a Nikon E800 epifluorescence microscope with 100X oil immersion objective with a Nikon E800 camera or with Nikon Eclipse Ti-E microscope and a Hamamatsu ORCAFlash 4.0 camera and a 60X oil immersion objective. Images were analyzed with NIS Elements software and Fiji (Schindelin et al., 2012).

Short term attachment and lectin binding assays.

Short term binding assays were carried out as previously described (Morton & Fuqua, 2012c). Briefly, glass coverslips were placed horizontally inside the wells of 6-well plates and submerged in 2 ml mid-exponential phase cultures of *A. tumefaciens* strains. Plates were incubated statically 2 h at room temperature to allow cells to swim to and bind the coverslip. Following incubation, cultures were aspirated, and coverslips washed twice in ATGN to remove unattached cells. Coverslips were then placed over 100 µl of 1:100 dilution (10 µg/ml final) fl-WGA, afWGA, or fl-DBA, or two in combination for dual lectin binding. Coverslips were allowed to stain 1 h, then rinsed twice in ATGN before placing them either on a microscope slide or an agarose pad (1-1.5% agarose, 100-200 µl). Visualization was performed as with standard UPP production assays. For high resolution single cell microscopy, spinning disk confocal microscopy (SDCM) was used (Yokogawa CSU10 confocal, Nikon TE2000U microscope, Photometrics Cascade II 512B camera) with Metamorph software (Molecular Devices Corp., Sunnyvale, CA). Fluorescence microscopy was performed with a Nikon Eclipse Ti-E microscope with a Hamamatsu ORCAFlash 4.0 and a 60X oil immersion objective. Excitation and emission wavelengths were specific for

af-WGA (excitation 510-560 nm, emission >610 nm) and fl-WGA or fl-DBA (excitation 460-500 nm, emission 510-560 nm). Images were processed using NIS-Elements Software and analyzed with Fiji (Schindelin et al., 2012).

Electron microscopy.

Electron microscopy was carried out at the Indiana Molecular Biology Institute Microscopy Center. For Transmission Electron Microscopy (TEM) and colloidal gold-labeled lectin, 1 ml of mid-exponential phase culture was pelleted and resuspended in 200 μ l ATGN. 5 μ l of 20 mg/ml suspensions of 20 nm colloidal gold conjugated WGA or 10 nm gold conjugated PNA (EY Laboratories) were added to the suspension and incubated at room temperature for 20 min. The labeled suspension was pelleted at 6000 x g and washed twice with 1 ml 10 mM Tris-Cl, pH 8.0 before resuspending in 10 mM Tris-Cl, pH 8.0. 10 μ l of the resuspension were placed on a formvar-coated copper grid for 5 min and stained with 10 μ l of 1% uranyl acetate for another 5 min. Filter paper was used to remove the excess uranyl acetate, and the grid was air dried for 5 min before viewing using the JEOL JEM-1010 Transmission Electron Microscope at 80 kV.

Purification of a His₆-UppH protein.

The *uppH* coding sequence (ATU_RS11555) was amplified from genomic DNA and fused downstream of the His₆-tag coding sequence in pET-15b cleaved at the *Nde*I and *Bam*HI restriction sites. The resulting plasmid pRN39 was transformed into chemically competent *E. coli* BL21-CodonPlus (DE3)-RIL cells (Stratagene) and grown at 37°C with shaking in 200 ml LB supplemented with Ap (100 μ g/ml). At mid-exponential phase, His₆-UppH expression was induced by addition of 200 μ M IPTG and incubating for 4 h. Cells were collected by centrifugation at 5,600 x g for 20 minutes and stored at -20°C. 2 g of cells were thawed on ice, resuspended in 5 ml extraction buffer (50 mM TES [(N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid)], 10 mM MgCl₂, 10 mM DTT, pH 7.0) and lysed using an M-110L Microfluidizer Processor (Microfluidics, Westwood, MA). Following 10 min centrifugation (9,900 x g), a significant fraction of the overexpressed tagged UppH was insoluble, but a fraction remained soluble based on SDS-PAGE analysis. The supernatant with the soluble His₆-UppH was loaded on to a 3 ml TALON Metal Affinity Resin (Clontech) equilibrated in binding buffer (50 mM Na₂HPO₄, 300 mM NaCl, 10% glycerol, 10 mM imidazole, pH 8). 4 ml of binding buffer containing increasing concentrations of imidazole (50 mM, 250 mM, 500 mM, and 1M) were used to carry out step gradient elution. UppH was present in the 250 mM imidazole fraction based on SDS-PAGE analysis and MALDI-MS of the excised band. The purified protein was dialyzed into 50 mM Na₂HPO₄, 100 mM NaCl, 10% glycerol, pH 8. Protein concentration was determined by the molar absorptivity using bovine serum albumin as a standard and the enzyme was aliquoted and stored at -80°C for future use. The same approach was used to purify His₆-UppH E152Q E266Q.

Transposon mutagenesis and *en-masse* sequencing.

Mutagenesis: Donor *E. coli* SM10- λ pir carrying the Himar1 plasmid pFD1 was grown in LB overnight until turbid. Recipient *A. tumefaciens* strains CDGS- Upp+ *pruA uppY*

and CDGS– Upp+ *pruA uppW* were also grown in LB until turbid. Turbid cultures were subcultured 1:10 in fresh LB and incubated until they reached mid-exponential phase (~4-6 h). 1 ml each of the donor and recipient strains were collected in Eppendorf tubes, spun down at 5,600 x g for 1 min, and the pellets resuspended in 50 µl LB. Subsequently, 50 µl of the donor strain was combined with 50 µl of each of the recipient strains. Entire matings (100 µl) were spotted on 0.2 µm cellulose acetate filter paper on LB plates. Plates were incubated overnight at 28°C to allow for conjugative transfer. Cells were resuspended from the filter paper in 30% glycerol, serially diluted in 1/10 increments, and frozen at –80°.

Screening and gDNA extraction: To optimize for visualization of CR phenotypes, solid medium was prepared using 1.5 ml of 2% stock solution of CR per 100 ml ATGN. Plates also contained 300 µg/ml Km for selection. 100 µl of cell suspensions were plated on ATGN-CR-Kan plates at 1:100 dilution for CDGS– Upp+ *pruA uppY*, and 1:1000 dilution for CDGS– Upp+ *pruA uppW*. Plates were incubated 48-72 h and screened for decreased CR (DCR) colony phenotypes. DCR colonies from the different strain backgrounds were picked, streaked or patched on fresh ATGN-CR-Kan plates and incubated at 28°C. PCR was used to confirm that the DCR mutants were *A. tumefaciens* and not Km^R contaminants with a DCR phenotype. Biofilm assays on randomly selected DCR mutants showed deficiencies to varying degrees. Confirmed DCR strains were inoculated in 2 ml ATGN-Kan, and 250 µl from each stationary-phase culture was pooled together into 1 ml batches. This was done separately for the two *A. tumefaciens* strain backgrounds. gDNA was extracted from pooled DCR cultures from the two strain backgrounds using (Promega gDNA extraction kit) and stored at 4°C. Over 50,000 colonies were screened collectively in both strain backgrounds, and roughly 0.5% were DCR mutants in each background.

Sequencing library preparation: Illumina library construction was carried out for the two separate gDNA pools using the terminal deoxynucleotidyl transferase (TdT) tailing method (Deng and Wu, 1983). Briefly, a Covaris instrument was used to shear gDNA to 200-700 bp fragments, and a Tape Station instrument was used to estimate sheared gDNA concentration (for this protocol, gDNA concentration was at least 69 ng/µl). Several consecutive PCR reactions were carried out to perform TdT tailing, to append adaptors sequences needed for Illumina sequencing, and to append adaptors with unique barcodes to the separate libraries corresponding to the two *A. tumefaciens* strain backgrounds. Final reaction products were combined to be multiplexed and the library was further purified by passing through a PCR purification column (Edge Biosystems).

The library was submitted to the IU Center for Genomics and Bioinformatics (CGB) for sequencing. Sequencing was performed on a NextSeq 500 using a 75-cycle single-end run. Reads were then mapped to the *Agrobacterium tumefaciens* C58 reference genome using bowtie2 (v2.3.5.1) with the following special parameters: --local -k 2. Only reads that originated within the transposon (and started with AACCTGTTA allowing a single mismatch) and had uniquely mapped were kept and tallied to quantify transposon insertion rates.

Multi-locus sequence analysis (MLSA).

Publicly available genome sequences from the *Agrobacterium/Rhizobium* group as well as select *Sinorhizobium*, *Ensifer*, and *Kaistia* strains were downloaded from NCBI GenBank on September 12, 2019. AutoMLSA with the default parameters and IQ-TREE v. 1.6.12 with the parameters “-bb 1000 -alrt 1000” were used to generate a multi-locus sequence analysis (MLSA) tree (Davis et al., 2016, Nguyen et al., 2015). Twenty-three translated gene sequences (*dnaK*, *prfC*, *hemN*, *hamI*, *truA*, *glyS*, *coxC*, ATU_RS03855, ATU_RS07710, *hom*, *rplB*, *rpoC*, *rpoB*, *murC*, *hemF*, ATU_RS12850, *acnA*, *leuS*, *cgtA*, *secA*, *plsC*, *aroB*, *lysC*) from the strain *A. fabrum* C58 (NCBI accession: GCA_000092025.1) were used as reference sequences for autoMLSA (Zhang et al., 2012). Protein sequences from the genome of *A. tumefaciens* (fabrum) C58 for the loci ATU_RS02360, ATU_RS02370, ATU_RS06100, ATU_RS06105, ATU_RS06110, ATU_RS06115, ATU_RS06120, ATU_RS06125, ATU_RS06130, ATU_RS11495, ATU_RS11500, ATU_RS11550, ATU_RS11555, ATU_RS11560, ATU_RS11565, ATU_RS11570, ATU_RS11575, ATU_RS11580, ATU_RS11585, Atu2378 (no annotation in NCBI), ATU_RS11590, ATU_RS16315, ATU_RS16320, ATU_RS17560, ATU_RS17565, ATU_RS17570, ATU_RS19030, ATU_RS19035, ATU_RS19040, ATU_RS26530, and ATU_RS16325 were used as queries in BLASTP v. 2.6.0+ searches of individual genomes; and the best blast hit was taken (Camacho et al., 2009). Top hit sequences were then used as queries in BLASTP searches against a database of proteins from the C58 reference genome. Blast hits were filtered to those with >50% query coverage and >40% amino acid ID, and the best hit to C58 were identified. If the best hits of both searches were reciprocal best hits, the presence of that gene was confirmed in a genome. Phylogenies and heat maps were visualized using the R package ggtree (Yu et al., 2017). Gene synteny figures were generated using clinker v. 0.0.9 with the default options (Gilchrist & Chooi, 2021).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We wish to acknowledge and Lorilyn Cozy and Megan Lohmiller for help with initial work on the UPP, and Yves Brun and Ellen Quardokus for helpful discussions and suggestions. Ankur Dalia generously provided protocols and troubleshooting advice for transposon sequencing. Barry Stein was a great help with electron microscopy, performed at the Indiana University (IU) Electron Microscopy Center. High throughput transposon sequencing was performed by the Indiana University Center for Genomics and Bioinformatics. We thank the Department of Botany and Plant Pathology at Oregon State University for its support of computing infrastructure. T.D. was supported through an Indiana University-Purdue University Collaborations in Life Science and Informatics Research grant. M.E.H. received support from the IU Genetics, Cellular and Molecular Sciences NIH Training Grant (T32 GM007757). This project was supported by grants to C.F. from the NIH (RO1 GM120337), and the Indiana University META-Cyt Program funded in part by a major endowment from the Lilly Foundation, and to J.H.C. by the National Science Foundation (DEB-1738028). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors have no conflicts of interest to declare.

Data availability:

High-throughput, Illumina-based transposon sequencing data has been deposited in the GEO database Accession No. GSE173827 ([dataset] Onyeziri and Fuqua, 2021). All other data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Anderson-Furgeson JC, Zupan JR, Grangeon R, and Zambryski PC (2016) Loss of PodJ in *Agrobacterium tumefaciens* leads to ectopic polar growth, branching, and reduced cell division. *J Bacteriol* 198: 1883–1891. [PubMed: 27137498]
- Baker P, Hill PJ, Snarr BD, Alnabelseya N, Pestrak MJ, Lee MJ, Jennings LK, Tam J, Melnyk RA, Parsek MR, Sheppard DC, Wozniak DJ, and Howell PL (2016) Exopolysaccharide biosynthetic glycoside hydrolases can be utilized to disrupt and prevent *Pseudomonas aeruginosa* biofilms. *Science advances* 2: e1501632. [PubMed: 27386527]
- Berne C, Ellison CK, Ducret A, and Brun YV (2018) Bacterial adhesion at the single-cell level. *Nat Rev Microbiol* 16: 616–627. [PubMed: 30008468]
- Braun AC, and Elrod RP (1946) Stages in the life history of *Phytomonas tumefaciens*. *J Bacteriol*. 52: 695–702. [PubMed: 16561237]
- Breton C, Snajdrova L, Jeanneau C, Koca J, and Imberty A (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology* 16: 29R–37R. [PubMed: 16049187]
- Brown PJ, de Pedro MA, Kysela DT, Van der Henst C, Kim J, De Bolle X, Fuqua C, and Brun YV (2012) Polar growth in the Alphaproteobacterial order Rhizobiales. *Proc Natl Acad Sci U S A* 109: 1697–1701. [PubMed: 22307633]
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421. [PubMed: 20003500]
- Collins RF, Beis K, Dong C, Botting CH, McDonnell C, Ford RC, Clarke BR, Whitfield C, and Naismith JH (2007) The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc Natl Acad Sci U S A* 104: 2390–2395. [PubMed: 17283336]
- Curtis PD, and Brun YV (2014) Identification of essential alphaproteobacterial genes reveals operational variability in conserved developmental and cell cycle systems. *Mol Microbiol* 93: 713–735. [PubMed: 24975755]
- Cuthbertson L, Mainprize IL, Naismith JH, and Whitfield C (2009) Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* 73: 155–177. [PubMed: 19258536]
- Davis EW, Weisberg AJ, Tabima JF, Grunwald NJ, and Chang JH (2016) Gall-ID: tools for genotyping gall-causing phytopathogenic bacteria. *PeerJ* 4: e2222. [PubMed: 27547538]
- De Castro C, Molinaro A, Lanzetta R, Silipo A, and Parrilli M (2008) Lipopolysaccharide structures from *Agrobacterium* and *Rhizobiaceae* species. *Carbohydr Res* 343: 1924–1933. [PubMed: 18353297]
- Deng GR, & Wu R (1983). Terminal transferase: Use in the tailing of DNA and for in vitro mutagenesis. *Methods in Enzymology*, 100(C), 96–116. 10.1016/0076-6879(83)00047-6. [PubMed: 6312266]
- Feirer N, Kim D, Xu J, Fernandez N, Waters CM, and Fuqua C (2017) The *Agrobacterium tumefaciens* CheY-like protein ClaR regulates biofilm formation. *Microbiology* 163: 1680–1691. [PubMed: 29068284]
- Feirer N, Xu J, Allen KD, Koestler BJ, Bruger EL, Waters CM, White RH, and Fuqua C (2015) A pterin-dependent signaling pathway regulates a dual-function diguanylate cyclase-phosphodiesterase controlling surface attachment in *Agrobacterium tumefaciens*. *mBio* 6: e00156. [PubMed: 26126849]

- Fiebig A, Herrou J, Fumeaux C, Radhakrishnan SK, Viollier PH, and Crosson S (2014) A cell cycle and nutritional checkpoint controlling bacterial surface adhesion. *PLoS Genet* 10: e1004101. [PubMed: 24465221]
- Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, and Kjelleberg S (2016) Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* 14: 563–575. [PubMed: 27510863]
- Fritts RK, LaSarre B, Stoner AM, Posto AL, and McKinlay JB (2017) A Rhizobiales-specific unipolar polysaccharide adhesin contributes to *Rhodopseudomonas palustris* biofilm formation across diverse photoheterotrophic conditions. *Appl Environ Microbiol* 83:e03035–16. [PubMed: 27986718]
- Fuqua C, (2008) *Agrobacterium*-host attachment and biofilm formation. In: *Agrobacterium: From Biology to Biotechnology*. Tzfira T & Citovsky V (eds). New York, NY: Springer Science + Business Media LLC, pp. 243–277.
- Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev* 67: 16–37. [PubMed: 12626681]
- Gilchrist CLM, and Chooi YH (2021) Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. *Bioinformatics*.
- Ginsberg C, Zhang YH, Yuan Y, and Walker S (2006) In vitro reconstitution of two essential steps in wall teichoic acid biosynthesis. *ACS Chem Biol* 1: 25–28. [PubMed: 17163636]
- Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, Qurollo B, Goldman BS, Cao Y, Askenazi M, Halling C, Mullin L, Houmiel K, Gordon J, Vaudin M, Iartchouk O, Epp A, Liu F, Wollam C, Allinger M, Doughty D, Scott C, Lappas C, Markelz B, Flanagan C, Crowell C, Gurson J, Lomo C, Sear C, Strub G, Cielo C, and Slater S (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* 294: 2323–2328. [PubMed: 11743194]
- Hardy GG, Allen RC, Toh E, Long M, Brown PJ, Cole-Tobian JL, and Brun YV (2010) A localized multimeric anchor attaches the *Caulobacter* holdfast to the cell pole. *Mol Microbiol* 76: 409–427. [PubMed: 20233308]
- Heckel BC, Tomlinson AD, Morton ER, Choi JH, and Fuqua C (2014) *Agrobacterium tumefaciens* *exoR* controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. *J Bacteriol* 196: 3221–3233. [PubMed: 24982308]
- Hernando-Perez M, Setayeshgar S, Hou Y, Temam R, Brun YV, Dragnea B, and Berne C (2018) Layered structure and complex mechanochemistry underlie strength and versatility in a bacterial adhesive. *mBio* 9:e02359–17.
- Hershey DM, Porfirio S, Black I, Jaehrig B, Heiss C, Azadi P, Fiebig A, and Crosson S (2019) Composition of the holdfast polysaccharide from *Caulobacter crescentus*. *J Bacteriol* 201:e00276–19. [PubMed: 31209074]
- Hinz AJ, Larson DE, Smith CS, and Brun YV (2003) The *Caulobacter crescentus* polar organelle development protein PodJ is differentially localized and is required for polar targeting of the PleC development regulator. *Mol Microbiol*. 47: 929–941. [PubMed: 12581350]
- Holden HM, Rayment I, and Thoden JB (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J Biol Chem* 278: 43885–43888. [PubMed: 12923184]
- Hong Y, Liu MA, and Reeves PR (2018) Progress in our understanding of Wzx flippase for translocation of bacterial membrane lipid-linked oligosaccharide. *J Bacteriol* 200:e00154–17. [PubMed: 28696276]
- Hong Y, Morcilla VA, Liu MA, Russell ELM, and Reeves PR (2015) Three Wzy polymerases are specific for particular forms of an internal linkage in otherwise identical O units. *Microbiology (Reading)* 161: 1639–1647. [PubMed: 25987464]
- Islam ST & Lam JS (2014) Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Canadian Journal of Microbiology*, 60, 697–716. 10.1139/cjm-2014-0595 [PubMed: 25358682]
- Javens J, Wan Z, Hardy GG, and Brun YV (2013) Bypassing the need for subcellular localization of a polysaccharide export-anchor complex by overexpressing its protein subunits. *Molecular Microbiology* 89: 350–371. [PubMed: 23714375]

- Kamoun S, Cooley MB, Rogowsky PM, and Kado CI (1989) Two chromosomal loci involved in production of exopolysaccharide in *Agrobacterium tumefaciens*. *J Bacteriol* 171: 1755–1759. [PubMed: 2921249]
- Khan SR, Gaines J, Roop RM 2nd, and Farrand SK (2008) Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl Environ Microbiol* 74: 5053–5062. [PubMed: 18606801]
- Lassalle F, Campillo T, Vial L, Baude J, Costechareyre D, Chapulliot D, Shams M, Abrouk D, Lavire C, Oger-Desfeux C, Hommais F, Gueguen L, Daubin V, Muller D, and Nesme X (2011) Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. *Genome Biol Evol* 3: 762–781. [PubMed: 21795751]
- Laus MC, Logman TJ, Lamers GE, Van Brussel AA, Carlson RW, and Kijne JW (2006) A novel polar surface polysaccharide from *Rhizobium leguminosarum* binds host plant lectin. *Mol Microbiol* 59: 1704–1713. [PubMed: 16553877]
- Li G, Brown PJ, Tang JX, Xu J, Quardokus EM, Fuqua C, and Brun YV (2012) Surface contact stimulates the just-in-time deployment of bacterial adhesins. *Mol. Microbiol* 83: 41–51. [PubMed: 22053824]
- Li YG, and Christie PJ (2018) The *Agrobacterium* VirB/VirD4 T4SS: mechanism and architecture defined through *in vivo* mutagenesis and chimeric systems. *Curr Top Microbiol Immunol* 418: 233–260. [PubMed: 29808338]
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, and Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42: D490–495. [PubMed: 24270786]
- Mathews SL, Hannah H, Samagaio H, Martin C, Rodriguez-Rassi E, and Matthyse AG (2019) Glycoside hydrolase genes are required for virulence of *Agrobacterium tumefaciens* on *Bryophyllum daigremontiana* and tomato. *Appl Environ Microbiol* 85:e00603–19. [PubMed: 31126942]
- Morton ER, and Fuqua C (2012a) Genetic manipulation of *Agrobacterium*. *Curr Protoc Microbiol*: Unit 3D 2. [PubMed: 22549164]
- Morton ER, and Fuqua C (2012b) Laboratory maintenance of *Agrobacterium*. *Curr Protoc Microbiol* **Chapter 1**: Unit 3D 1. **Chapter**
- Morton ER, and Fuqua C (2012c) Phenotypic analyses of *Agrobacterium*. *Curr Protoc Microbiol* **Chapter 3**: Unit 3D 3. [PubMed: 22549164] **Chapter**
- Nguyen LT, Schmidt HA, von Haeseler A, and Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32: 268–274. [PubMed: 25371430]
- Oglesby LL, Jain S, and Ohman DE (2008) Membrane topology and roles of *Pseudomonas aeruginosa* Alg8 and Alg44 in alginate polymerization. *Microbiology (Reading)* 154: 1605–1615. [PubMed: 18524915]
- Onyeziri and Fuqua (2021). Transposon sequencing for pathway-specific unipolar polysaccharide biosynthesis screen in *Agrobacterium tumefaciens* C58 has been deposited in the GEO database Accession No. GSE173827.
- Pearson CR, Tindall SN, Herman R, Jenkins HT, Bateman A, Thomas GH, Potts JR, and Van der Woude MW (2020) Acetylation of surface carbohydrates in bacterial pathogens requires coordinated action of a two-domain membrane-bound acyltransferase. *mBio* 11:e01364–20. [PubMed: 32843546]
- Poindexter JS (1981) The caulobacters: ubiquitous unusual bacteria. *Microbiol. Rev* 45: 123–179. [PubMed: 7012570]
- Reuber TL, and Walker GC (1993) The acetyl substituent of succinoglycan is not necessary for alfalfa nodule invasion by *Rhizobium meliloti* Rm1021. *J Bacteriol* 175: 3653–3655. [PubMed: 8501069]
- Romling U, and Galperin MY (2015) Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. *Trends Microbiol* 23: 545–557. [PubMed: 26077867]
- Ronneau S and Hallez R (2019) Make and break the alarmone: regulation of (p)ppGpp synthetase/hydrolase enzymes in bacteria. *FEMS Microbiol. Rev* 43:389–400. [PubMed: 30980074]

- Ruan X, Monjaras Feria J, Hamad M, and Valvano MA (2018) *Escherichia coli* and *Pseudomonas aeruginosa* lipopolysaccharide O-antigen ligases share similar membrane topology and biochemical properties. *Mol Microbiol* 110: 95–113. [PubMed: 30047569]
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, and Cardona A (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9: 676–682. [PubMed: 22743772]
- Slater SC, Goldman BS, Goodner B, Setubal JC, Farrand SK, Nester EW, Burr TJ, Banta L, Dickerman AW, Paulsen I, Otten L, Suen G, Welch R, Almeida NF, Arnold F, Burton OT, Du Z, Ewing A, Godsy E, Heisel S, Houmiel KL, Jhaveri J, Lu J, Miller NM, Norton S, Chen Q, Phoolcharoen W, Ohlin V, Ondrusek D, Pride N, Stricklin SL, Sun J, Wheeler C, Wilson L, Zhu H, and Wood DW (2009) Genome sequences of three agrobacterium biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol* 191: 2501–2511. [PubMed: 19251847]
- Stein BJ, Fiebig A, and Crosson S. (2021). The ChvG-ChvI and NtrY-NtrX two-component systems coordinately regulate growth of *Caulobacter crescentus*. *J. Bacteriol* 203:e001921.
- Tempé J, Petit A, Holsters M, Van Montagu M, and Schell J (1977) Thermosensitive step associated with transfer of the Ti plasmid during conjugation: possible relation to transformation in crown gall. *Proc. Natl. Acad. Sci. USA* 74: 2848–2849. [PubMed: 16592419]
- Thompson MA, Onyeziri MC, and Fuqua C (2018) Function and regulation of *Agrobacterium tumefaciens* cell surface structures that promote attachment. *Curr Top Microbiol Immunol* 418: 143–184. [PubMed: 29998422]
- Toh E, Kurtz HD Jr., and Brun YV (2008) Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. *J Bacteriol.* 190: 7219–7231. [PubMed: 18757530]
- Tomlinson AD, and Fuqua C (2009) Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. *Curr Opin Microbiol* 12: 708–714. [PubMed: 19879182]
- Vetting MW, LP S.d.C., Yu M, Hegde SS, Magnet S, Roderick SL, and Blanchard JS (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys* 433: 212–226. [PubMed: 15581578]
- Vuong C, Kocianova S, Voyich JM, Yao Y, Fischer ER, DeLeo FR, and Otto M (2004) A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion, and virulence. *J Biol Chem* 279: 54881–54886. [PubMed: 15501828]
- Wan Z, Brown PJ, Elliott EN, and Brun YV (2013) The adhesive and cohesive properties of a bacterial polysaccharide adhesin are modulated by a deacetylase. *Mol Microbiol* 88: 486–500. [PubMed: 23517529]
- Weisberg AJ, Davis EW 2nd, Tabima J, Belcher MS, Miller M, Kuo CH, Loper JE, Grunwald NJ, Putnam ML, and Chang JH (2020) Unexpected conservation and global transmission of agrobacterial virulence plasmids. *Science* 368:eaba5256. [PubMed: 32499412]
- Whitfield C, Wear SS, and Sande C (2020) Assembly of bacterial capsular polysaccharides and exopolysaccharides. *Annu Rev Microbiol* 74: 521–543. [PubMed: 32680453]
- Williams A, Wilkinson A, Krehenbrink M, Russo DM, Zorreguieta A, and Downie JA (2008) Glucomanan-mediated attachment of *Rhizobium leguminosarum* to pea root hairs is required for competitive nodule infection. *J Bacteriol.* 190: 4706–4715. [PubMed: 18441060]
- Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NFJ, Woo L, Chen Y, Paulsen IT, Eisen JA, Karp PD, Dovee DS, Chapman P, Clendenning J, Deatherage G, Gillet W, Grant C, Kutayin T, Levy R, Li M-J, McClelland E, Palmieri A, Raymond C, Rouse G, Saenphimmachak C, Wu Z, Romero P, Gordon D, Zhang S, Yoo H, Tao Y, Biddle P, Jung M, Krespan W, Perry M, Gordon-Kamm B, Liao L, Kim S, Hendrick C, Zhao Z-Y, Dolan M, Chumley F, Tingey SV, Tomb J-F, Gordon M, Olson MV, and Nester EW (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* 294: 2317–2323. [PubMed: 11743193]
- Xu J, Kim J, Danhorn T, Merritt PM, and Fuqua C (2012) Phosphorus limitation increases attachment in *Agrobacterium tumefaciens* and reveals a conditional functional redundancy in adhesin biosynthesis. *Res Microbiol* 163: 674–684. [PubMed: 23103488]

- Xu J, Kim J, Koestler BJ, Choi JH, Waters CM, and Fuqua C (2013) Genetic analysis of *Agrobacterium tumefaciens* unipolar polysaccharide production reveals complex integrated control of the motile-to-sessile switch. *Mol Microbiol* 89: 929–948. [PubMed: 23829710]
- Yildiz F, Fong J, Sadovskaya I, Grard T, and Vinogradov E (2014) Structural characterization of the extracellular polysaccharide from *Vibrio cholerae* O1 El-Tor. *PLoS One* 9: e86751. [PubMed: 24520310]
- Yu G, Smith DK, Zhu H, Guan Y, and Lam TT-Y (2017) ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Meth. Ecol. Evol* 8: 28–36.
- Zhang YM, Tian CF, Sui XH, Chen WF, and Chen WX (2012) Robust markers reflecting phylogeny and taxonomy of rhizobia. *PLoS One* 7: e44936. [PubMed: 23028691]

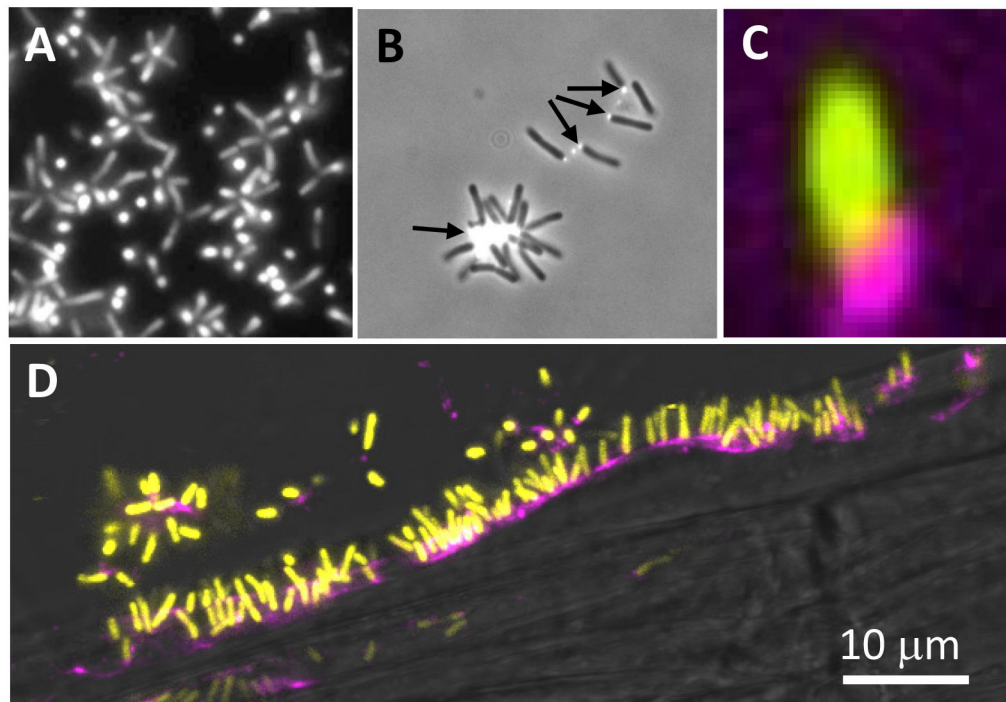


Figure 1. UPP-mediated polar attachment in *A. tumefaciens*.

A. tumefaciens UPP on abiotic surfaces and plant tissues. Fluorescence microscopy using wheat germ agglutinin lectin labeled with FTIC (fl-WGA) and Alexa Fluor 594 (af-WGA). (A-C) Attachment assays on 22 x 22 mm glass coverslips. *A. tumefaciens* C58 cells incubated for the indicated time in ATGN minimal medium with coverslips prior to lectin labeling with 10 $\mu\text{g}/\text{ml}$ of either fl-WGA or af-WGA for 20 min. (A) fl-WGA labeling of a 12 h biofilm. Merged image from bright field and fluorescence microscopy. (B) af-WGA-labeling of a short-term attachment assay after 2 h incubation. Merged image from phase contrast and fluorescence. (C) af-WGA labeling of a single *A. tumefaciens* cell attached to glass expressing plasmid-borne $P_{tac}::gfpmut3$ (pJZ383). Scanning Disk Laser Confocal fluorescence microscopy; Merged green and red channels, pseudo-colored yellow and magenta, respectively. (D) Binding of *A. tumefaciens* (pJZ383) to *Arabidopsis thaliana* WS 1 cm root cutting incubated for 3 h in ATGN minimal medium. Merged green and red channels pseudo-colored yellow and magenta, respectively plus bright field. Microscopes: Nikon E800 fluorescence microscope and Photometrics Cascade cooled CCD camera; Spinning Disk Nikon TE2000U microscope Yokagawa CSU10 confocal scanner unit.

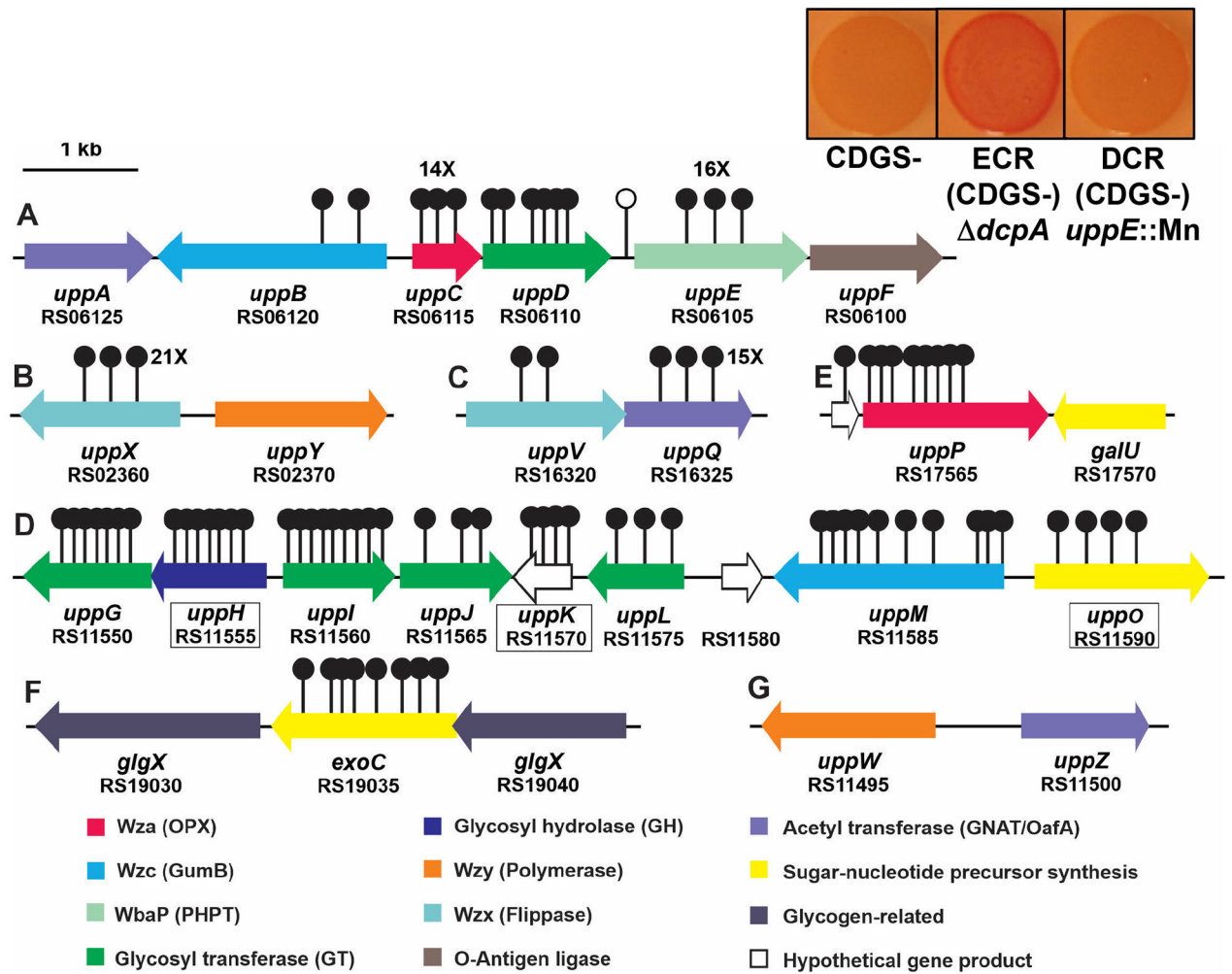


Figure 2. UPP biosynthesis gene clusters and Congo red colony phenotypes in *A. tumefaciens*. Gene clusters important for UPP biosynthesis that were identified via transposon screens for mutants with decreased Congo red (DCR) phenotypes. Stemmed circles represent independent, sequenced transposon insertion sites. Letters A-G designate specific genetic loci. The single unfilled marker upstream of *uppE* represents a transposon mutant isolated based on a biofilm deficiency (Xu et al., 2012). Numbers with X indicate 14 or more independent transposon insertions. Boxed gene names and designations indicate genes for which in-frame-deletions in wild type do not lead to attachment deficiencies. Predicted functions based in sequence similarity are described in the key below the gene map and are color-coded. Inset boxes on the upper right provide CR phenotypes of *A. tumefaciens* strains lacking all the known dispensable exopolysaccharides except UPP (CDGS-), the same strain elevated for cdGMP via a *dcpA* deletion and thus exhibiting an elevated Congo red (ECR) phenotype (middle), and the same strain with a Mariner transposon insertion in *uppE* conferring a DCR phenotype.

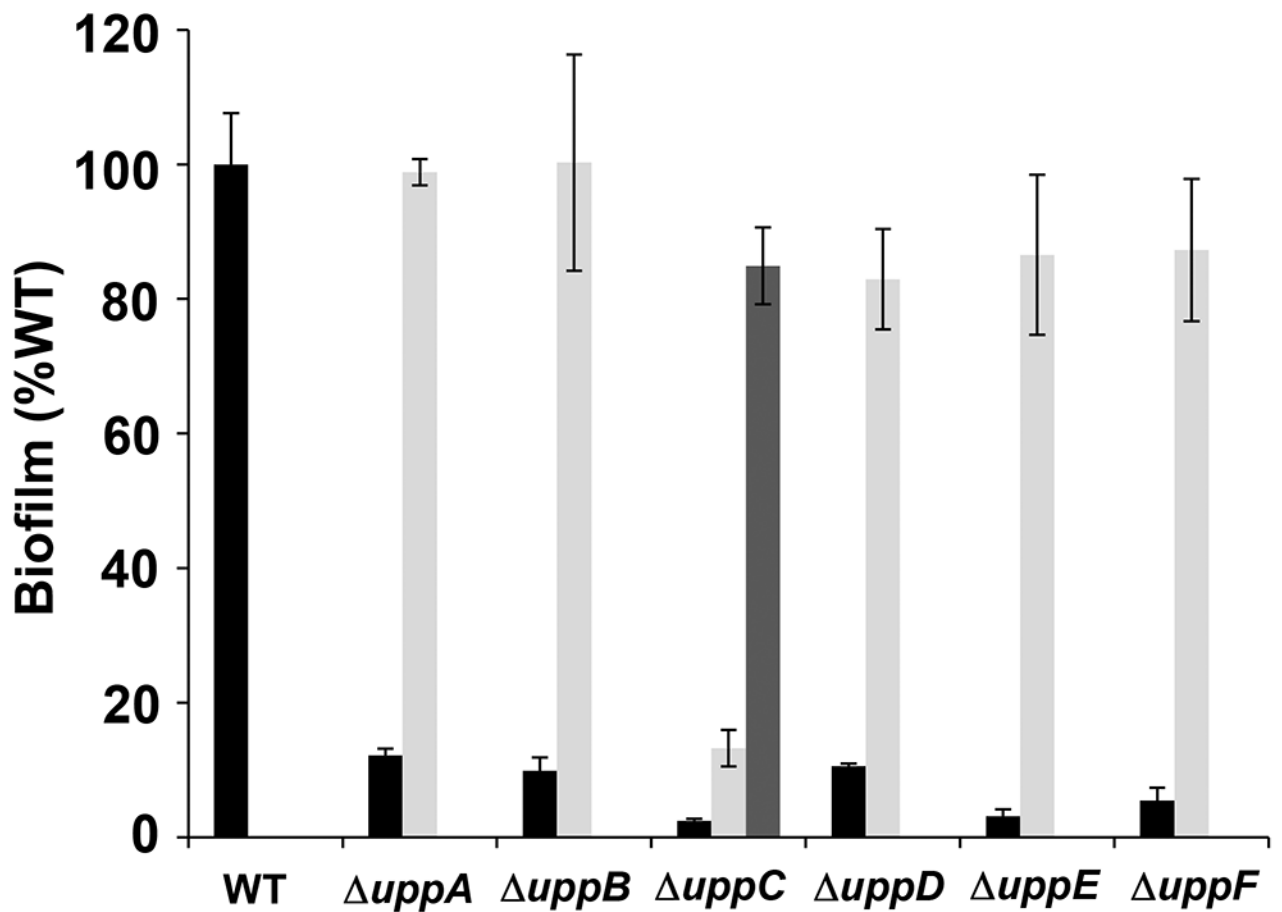


Figure 3. Biofilm phenotypes of *uppA-F* mutants.

Ratio of acetic acid-solubilized CV absorbance (A_{600}) from 48 h biofilm assays normalized to the OD_{600} turbidity measured from the planktonic phase of the same culture. Ratio for *A. tumefaciens* C58 wild type set to 100% and mutant derivatives expressed as percent wild type. Black bars are the strains harboring the pSRK-Gm vector, light grey bars are expressing pSRK-Gm derivatives with the corresponding gene, and the single dark grey bar is the *uppC* strain expressing a pSRK-Gm derivative with *uppC* and *uppD*. All assays were performed with 500 μ M IPTG. Bars represent the mean of triplicate assays and error bars are standard deviation.

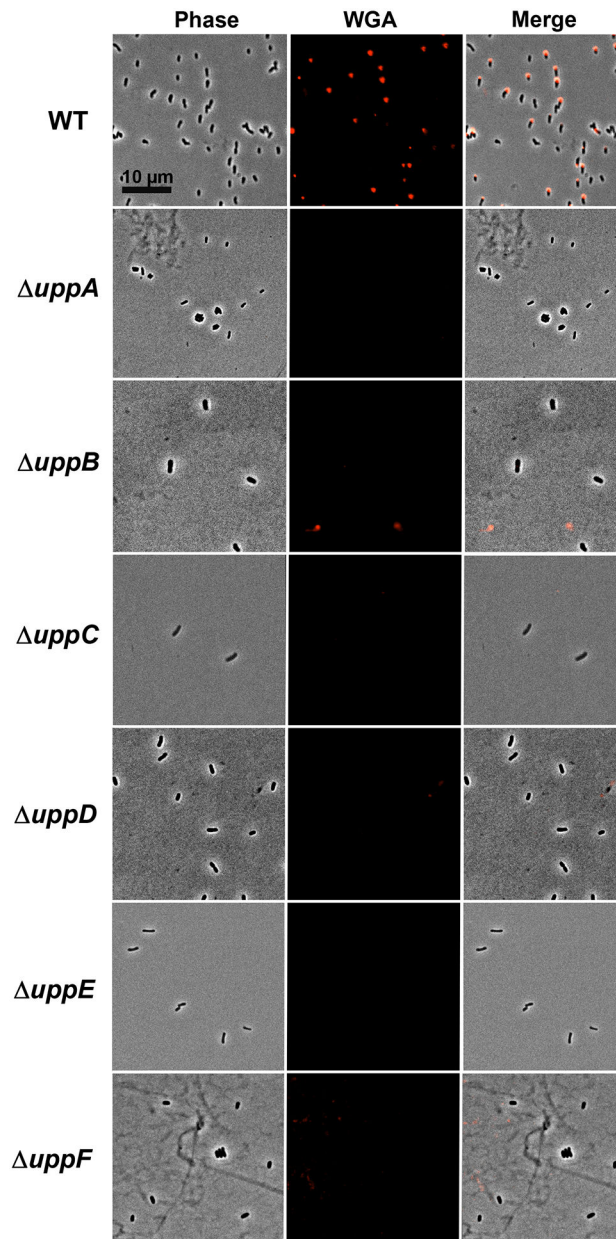


Figure 4. af-WGA labeling of UPP in *uppA-F* mutants.

Short term binding assays of *A. tumefaciens* C58 derivatives on glass coverslips stained with af-WGA. Phase images, fluorescence images and merged images. Microscopy was performed as described in experimental procedures on a Nikon Eclipse Ti-E microscope with a Hamamatsu ORCAFlash 4.0 with the 60X oil immersion objective.

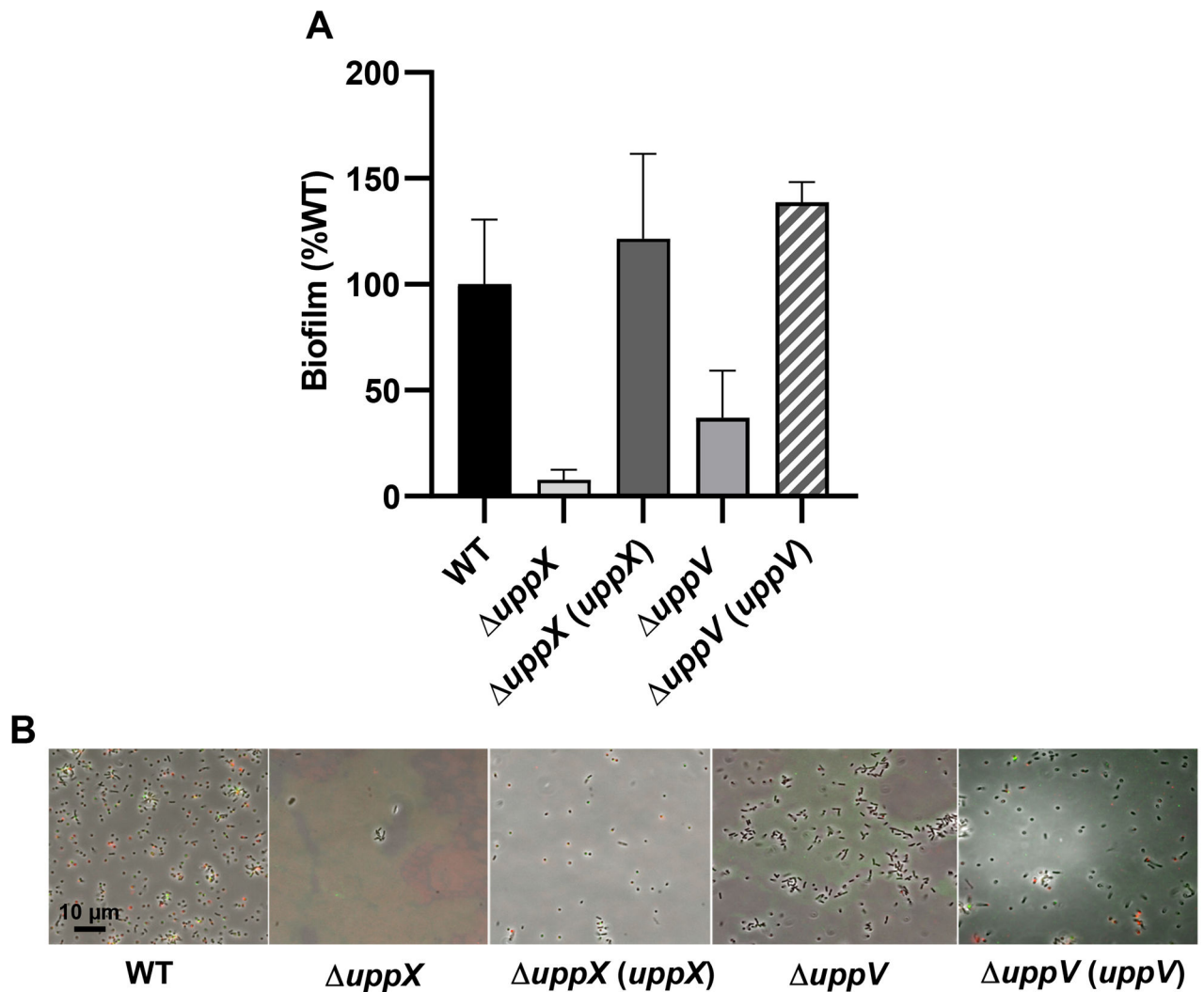


Figure 5. Biofilm and lectin-binding phenotypes of putative flippase deletion mutants.

(A) Biofilm assays performed for flippase deletion mutants and complementation strains with *A. tumefaciens* wild type (WT) biofilm was set to 100%. Ratio of acetic acid-solubilized CV absorbance (A_{600}) from 48 h biofilm assays normalized to the OD_{600} turbidity measured from the planktonic phase of the same culture. Plasmid-borne copies of genes were expressed from the P_{lac} promoter in pSRK derivatives induced with 500 μ M IPTG. Bars represent the mean of triplicate assays and error bars are standard deviation. (B) Short term attachment assays as described in the experimental procedures, co-incubated with af-WGA (red) and fl-DBA (green). Only merged images shown. Fluorescence microscopy was performed as described in experimental procedures on a Nikon Eclipse Ti-E microscope with a Hamamatsu ORCAFlash 4.0 camera with the 60X oil immersion objective.

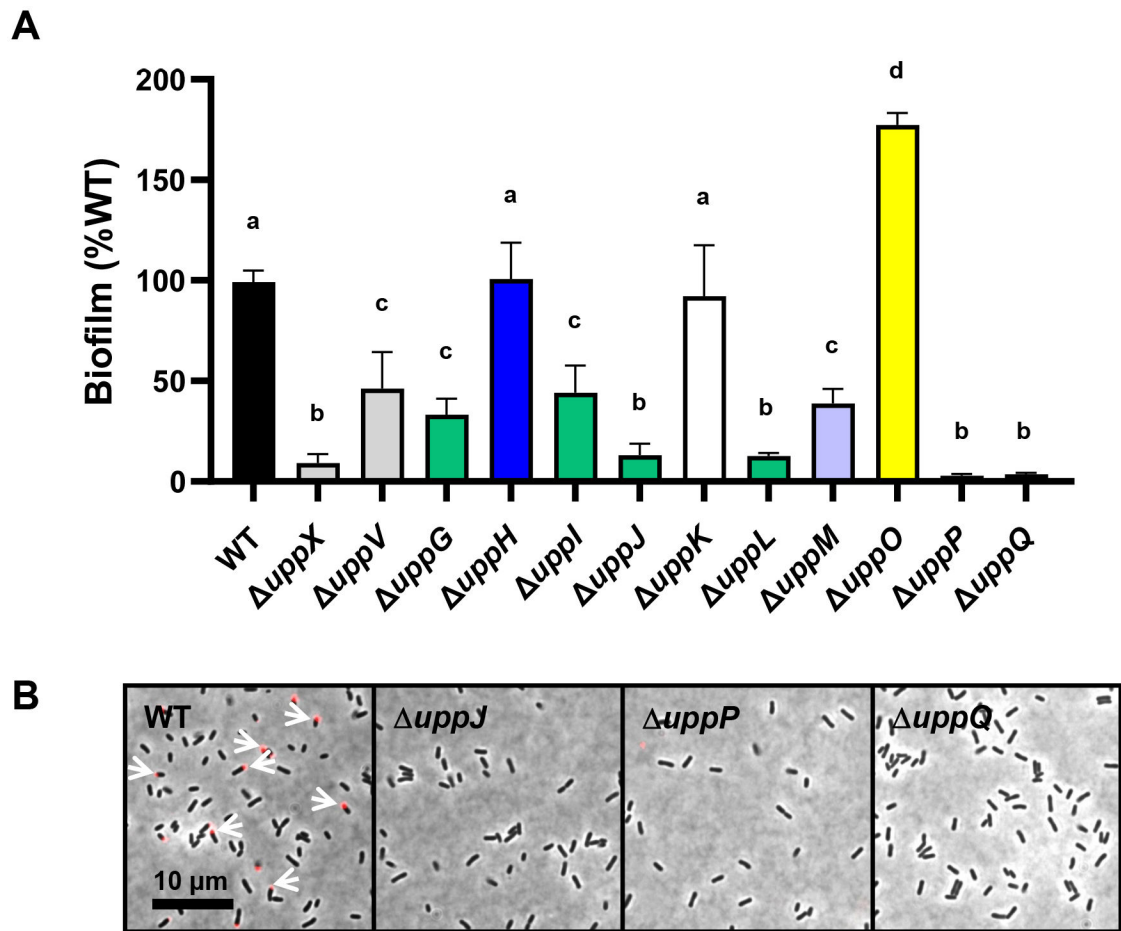


Figure 6. Biofilm formation and lectin labeling of genes identified in the DCR genetic screen. (A) A_{600} of acetic acid-solubilized CV absorbance from 48 h biofilm assays as a ratio to OD_{600} from the planktonic phase of the same culture. Ratio for wild type set to 100% and mutant derivatives expressed as percent wild type. Bars represent the mean of triplicate assays and error bars are standard deviation. Bars with the same letter have means that are not significantly different from: a, WT C58; b, *uppX* c, *uppV*; d, *uppO* (Ordinary one-way ANOVA, P value < 0.05). Color coding of bars matches gene map (Fig. 2) and the model (Fig. 13). (B) af-WGA labeling of *A. tumefaciens* C58 WT and selected mutant derivatives on 1.5% agar pads using a Nikon E800 epifluorescence microscope under 100x oil immersion objective.

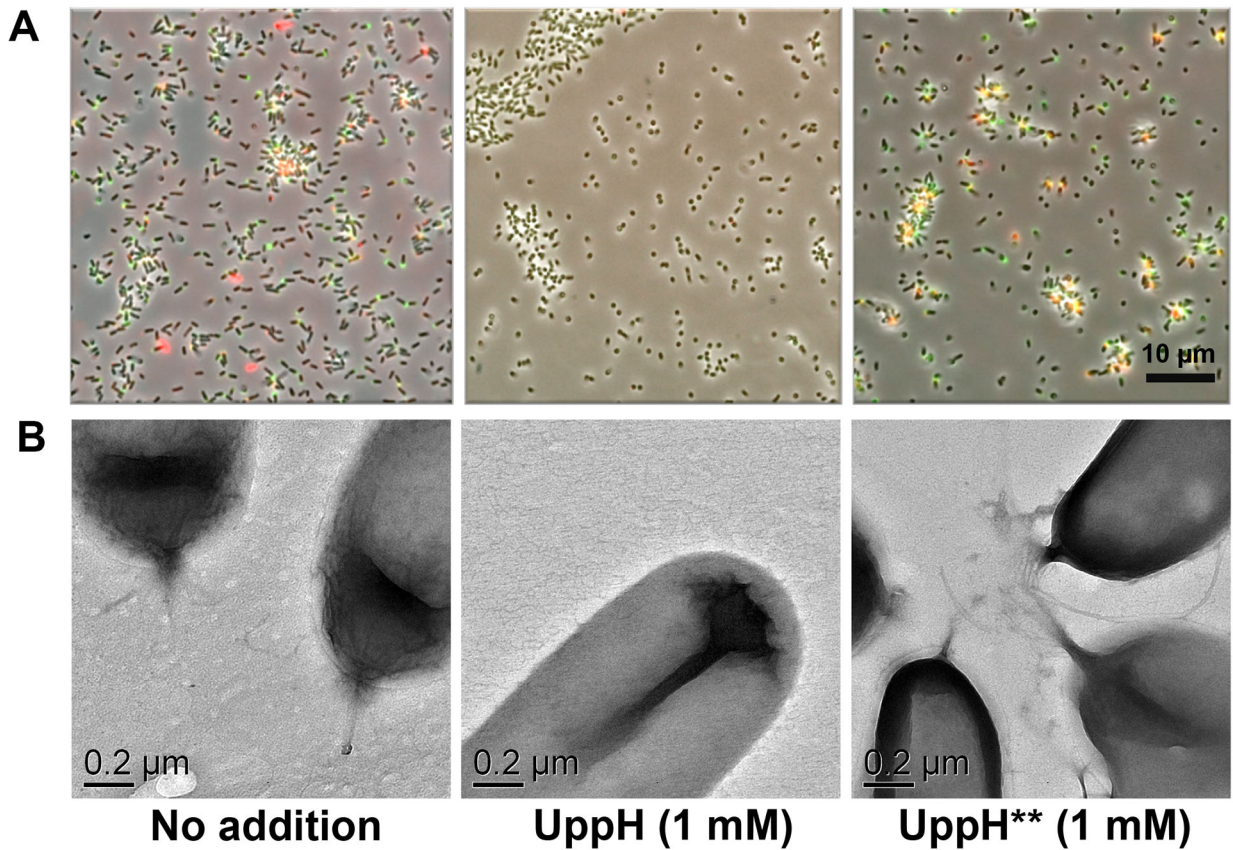


Figure 7. Purified UppH breaks down UPP in *A. tumefaciens*.

Purified His₆-UppH and His₆-UppH** at 1 mM were incubated for 1 h with *A. tumefaciens pruA* CDGS-. (A) Cells were pre-attached to glass coverslips prior to incubation with UppH preparations, and subsequently labeled with af-WGA and fl-DBA. Fluorescence microscopy was performed as described in experimental procedures with only the merged phase contrast and fluorescence images shown. (B) TEM performed on cells following incubation with UppH preparations on a JEOL JEM-1010 Transmission Electron Microscope at 80 kV.

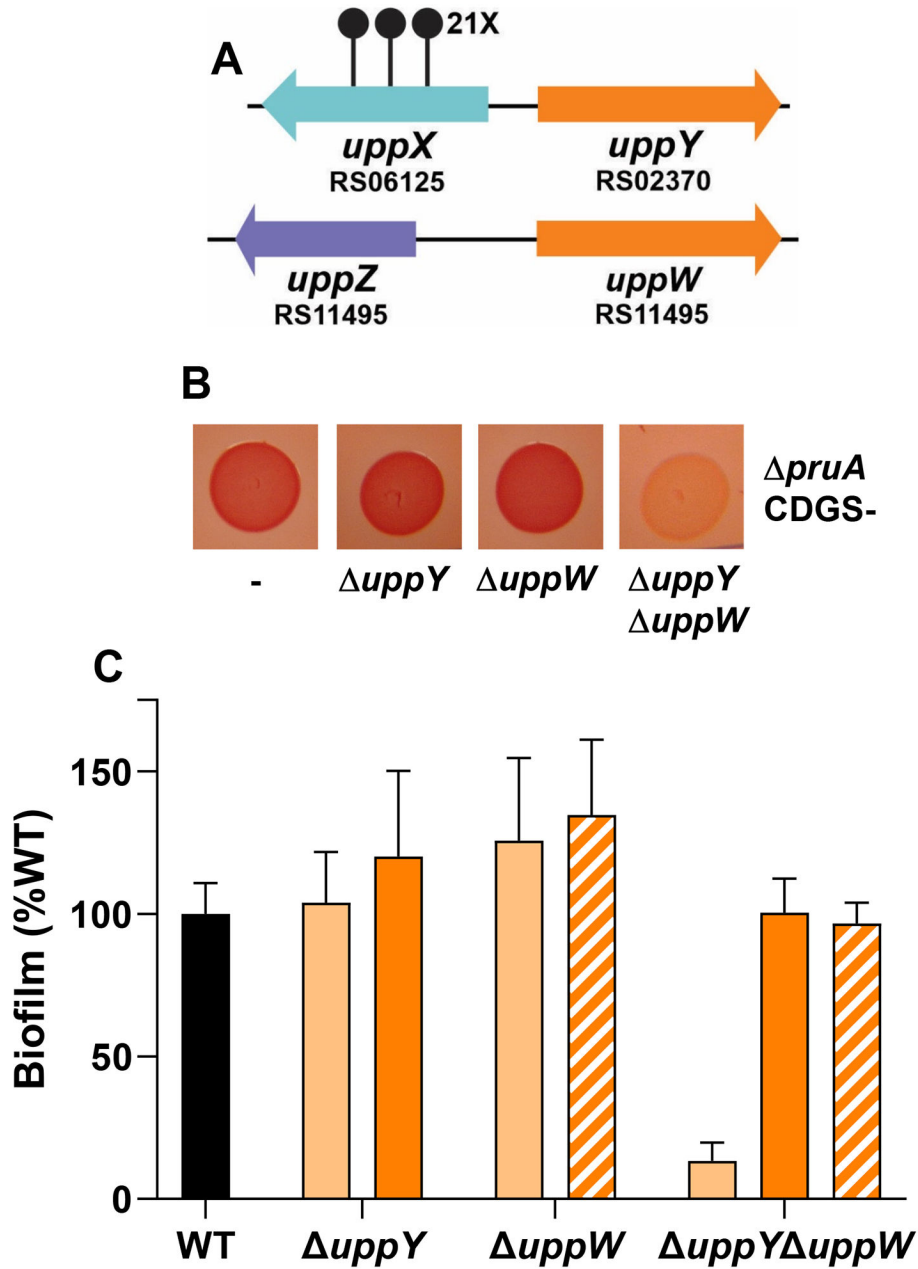


Figure 8. Two Wzy-type polymerase proteins are redundant for attachment and biofilm formation.

(A) Genetic loci of *uppY* and *uppW*. Stalked circles represent independent, sequenced transposon insertion sites and numbers with X indicate 14 or more independent transposon insertions. (B) ATGN CR colony phenotypes of the indicated deletion mutants in the *pruA* CDGS- background. (C) Quantitative biofilm as described in experimental procedures with the ratio of A_{600} of acetic acid-solubilized CV-stained coverslips to planktonic OD_{600} normalized to percent wild type for the indicated in-frame deletion mutants in otherwise wild type C58 after 48 h incubation in ATGN. The indicated strains with solid black and light orange bars have the pSRK-Gm vector control, whereas the strains represented by dark orange bars harbor a P_{lac} -*uppY* plasmid and the strains with cross hatched bars harbor the

P_{lac}-uppW plasmid. Cultures grown in 400 μ M IPTG to induce *P_{lac}*. Error bars are standard deviation of assays performed in triplicate.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

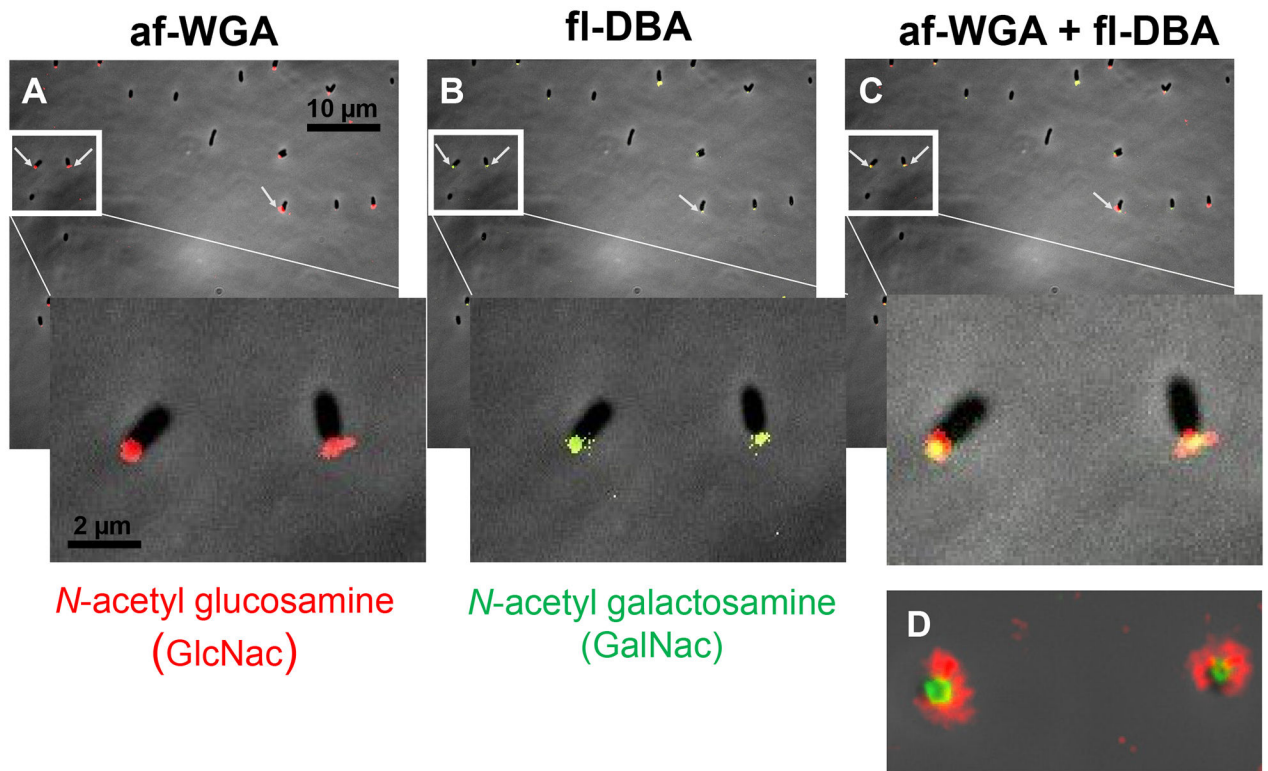


Figure 9. Dual lectin binding of UPP with WGA and DBA.

Short term attachment assay of *A. tumefaciens pruA* (CDGS⁻) on a glass coverslip co-incubated with lectins. (A) UPP stained with af-WGA showing red fluorescent polar foci. (B) UPP stained with fl-DBA showing green fluorescent polar foci. (C) UPP stained with both af-WGA and fl-DBA showing the sectored pattern of dual red and green fluorescence. Arrows indicate clearly labeled UPP. Insets – magnification of boxed area. (D) Image of two attached dual labeled cells at the focal plane of surface contact. Microscopy was performed as described in experimental procedures on a Nikon Eclipse Ti-E microscope with a Hamamatsu ORCAFlash 4.0 camera with the 60X oil immersion objective.

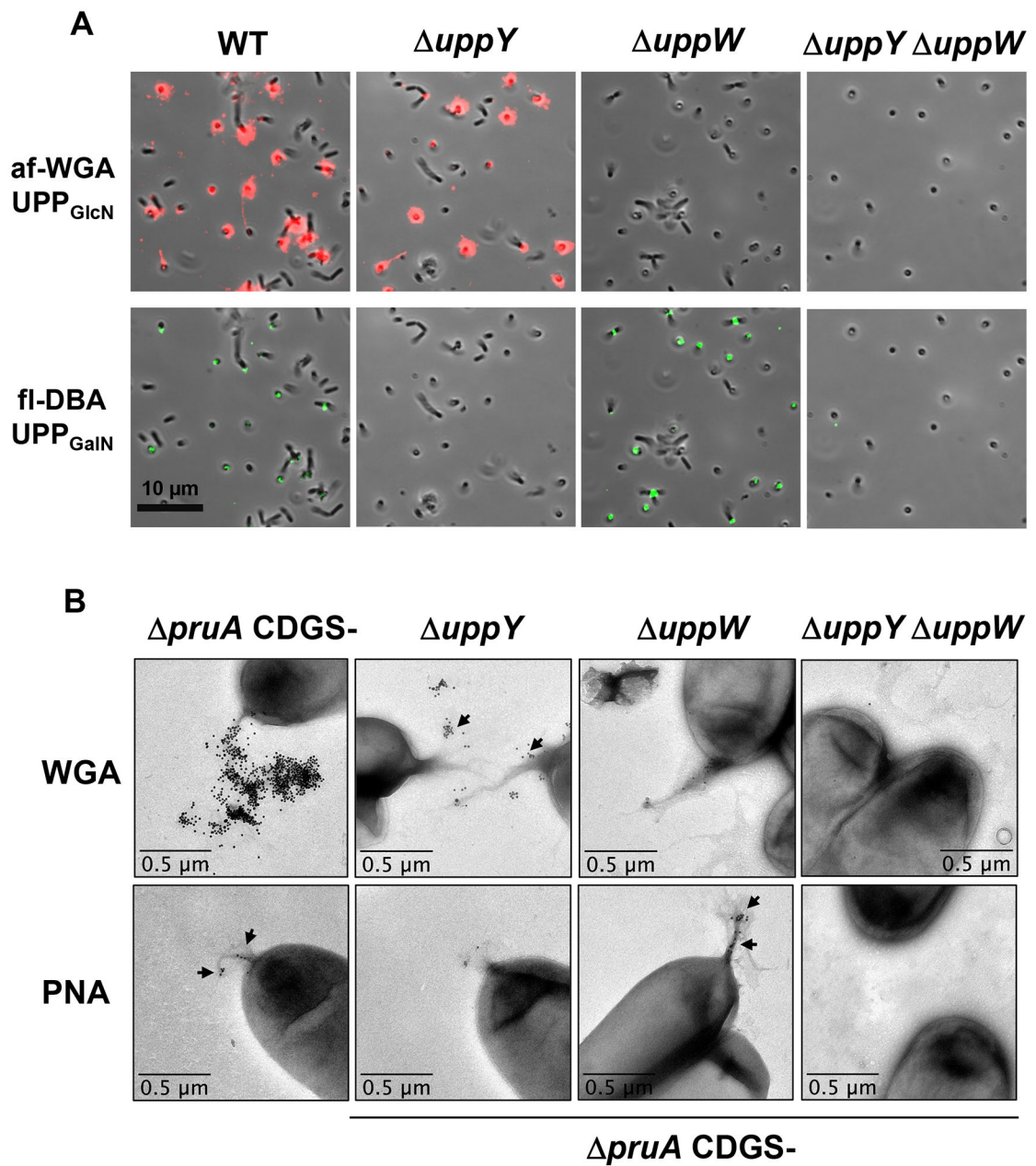


Figure 10. Dual lectin binding in *uppW* and *uppY* mutants shows reciprocal specificity. (A) Short term attachment assay on a glass coverslip co-incubated with WT, *uppY*, *uppW*, and *uppY uppW* mutants co-incubated with af-WGA and fl-DBA showing reciprocal fluorescence of the polymerase mutants. Microscopy was performed as described in experimental procedures on a Nikon Eclipse Ti-E microscope with a Hamamatsu ORCAFlash 4.0 camera with the 60X oil immersion objective. (B) Transmission electron microscopy (TEM) of UPP from polymerase deletion mutants. TEM analysis of *A. tumefaciens pruA* CDGS- mutant derivatives on a JEOL JEM-1010 Transmission Electron Microscope at 80 kV. using WGA lectin conjugated with gold beads (20 nm) and PNA lectin conjugated with gold beads (10 nm).

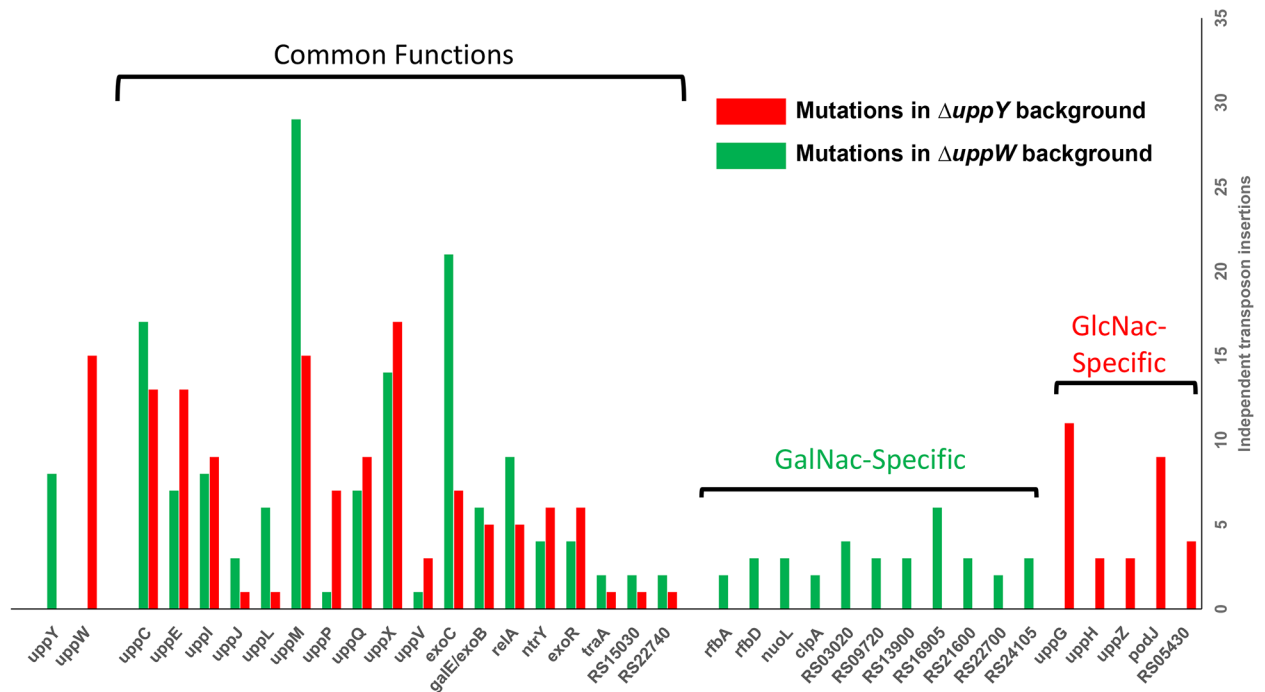


Figure 11. Genes isolated in UPP pathway-specific screens.

On the y-axis are number of independent transposon insertions that resulted in a scorable DCR phenotype for the *pruA* (CDGS⁻) polymerase deletion mutants (*uppY* and *uppX*), and on the x-axis are the genes identified with transposon insertions. Red bars indicate genes disrupted in the *uppY* background and therefore predicted to be important for UPP_{GlcN} synthesis. Green bars indicate genes identified in the *uppW* background and thus predicted to be important in the synthesis of UPP_{GalN}. Red and green bars together indicate genes identified in both mutant screening backgrounds.

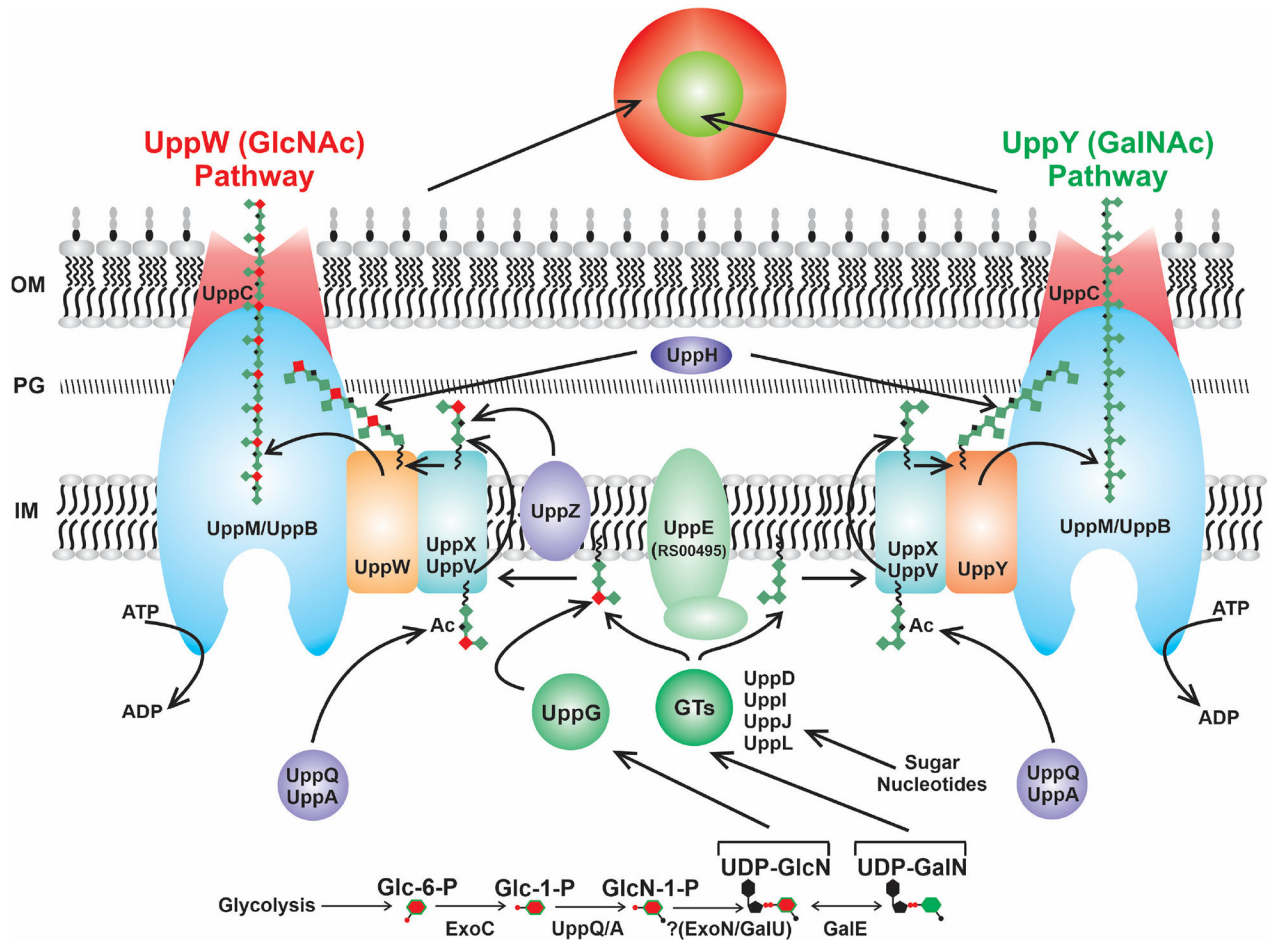


Figure 13. Model of UPP biosynthesis pathway in *A. tumefaciens*. Dual Wzx/Wzy polysaccharide biosynthesis pathway showing proteins that exhibit specificity for UPP_{GlcN} or UPP_{GalN}, and proteins involved in synthesizing both polysaccharides. Color-coding is consistent with the legend in Fig. 2. IM – inner membrane; PG– Peptidoglycan; OM – outer membrane; GT – glycosyl transferase; Ac – acetylation.