



Article

# Bioinformatic Analysis of Sulfotransferases from an Unexplored Gut Microbe, *Sutterella wadsworthensis* 3\_1\_45B: Possible Roles towards Detoxification via Sulfonation by Members of the Human Gut Microbiome

Lauryn Langford <sup>1,2</sup> and Dhara D. Shah <sup>1,2,\*</sup>

<sup>1</sup> Biodesign Center for Fundamental and Applied Microbiomics, Arizona State University, Tempe, AZ 85281, USA; llangfo2@asu.edu

<sup>2</sup> School of Mathematical and Natural Sciences, Arizona State University, Glendale, AZ 85306, USA

\* Correspondence: dhara.shah1@asu.edu

**Abstract:** Sulfonation, primarily facilitated by sulfotransferases, plays a crucial role in the detoxification pathways of endogenous substances and xenobiotics, promoting metabolism and elimination. Traditionally, this bioconversion has been attributed to a family of human cytosolic sulfotransferases (hSULTs) known for their high sequence similarity and dependence on 3'-phosphoadenosine 5'-phosphosulfate (PAPS) as a sulfo donor. However, recent studies have revealed the presence of PAPS-dependent sulfotransferases within gut commensals, indicating that the gut microbiome may harbor a diverse array of sulfotransferase enzymes and contribute to detoxification processes via sulfation. In this study, we investigated the prevalence of sulfotransferases in members of the human gut microbiome. Interestingly, we stumbled upon PAPS-independent sulfotransferases, known as aryl-sulfate sulfotransferases (ASSTs). Our bioinformatics analyses revealed that members of the gut microbial genus *Sutterella* harbor multiple *asst* genes, possibly encoding multiple ASST enzymes within its members. Fluctuations in the microbes of the genus *Sutterella* have been associated with various health conditions. For this reason, we characterized 17 different ASSTs from *Sutterella wadsworthensis* 3\_1\_45B. Our findings reveal that *SwASSTs* share similarities with *E. coli* ASST but also exhibit significant structural variations and sequence diversity. These differences might drive potential functional diversification and likely reflect an evolutionary divergence from their PAPS-dependent counterparts.

**Keywords:** sulfotransferases; *Sutterella*; human gut microbiome; sulfation; detoxification; PAPS; aryl-sulfate sulfotransferases; human sulfotransferases; gut microbial enzymes



**Citation:** Langford, L.; Shah, D.D. Bioinformatic Analysis of Sulfotransferases from an Unexplored Gut Microbe, *Sutterella wadsworthensis* 3\_1\_45B: Possible Roles towards Detoxification via Sulfonation by Members of the Human Gut Microbiome. *Int. J. Mol. Sci.* **2024**, *25*, 2983. <https://doi.org/10.3390/ijms25052983>

Academic Editor: Andreas Burkovski

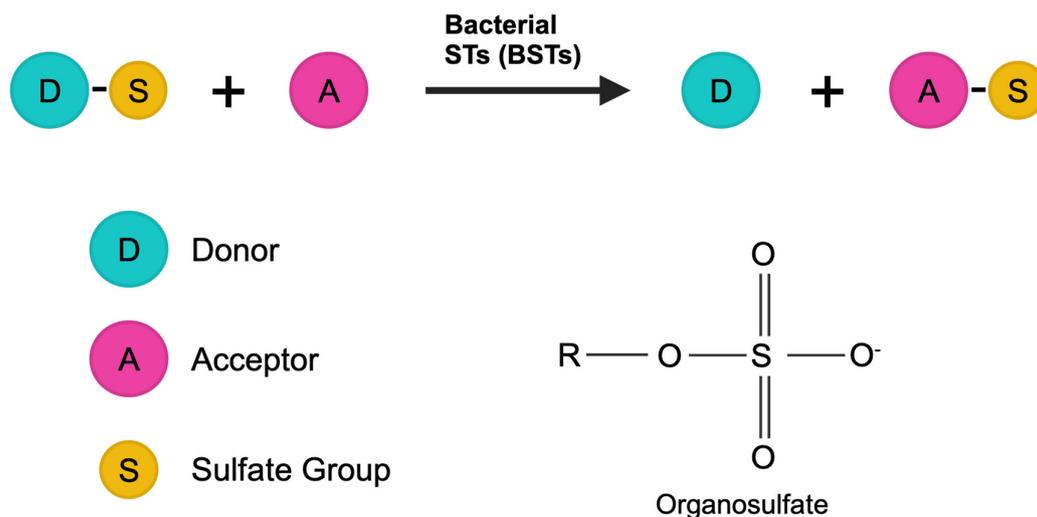
Received: 19 January 2024  
Revised: 27 February 2024  
Accepted: 28 February 2024  
Published: 4 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Sulfonation is a key detoxification pathway, where endogenous compounds and foreign substances undergo a transformation to increase their water solubility, aiding in metabolism and elimination [1]. This process has traditionally been attributed to a diverse array of human cytosolic sulfotransferases [2–20]. However, recent discoveries have identified enzymes capable of cholesterol sulfonation within the *Bacteroides* genus, a prominent member of the human gut microbiota [21,22]. This suggests the potential presence of additional sulfotransferases in the gut microbiome. While microbial sulfotransferases have been previously studied, they lack the extensive characterization seen with human enzymes [23–26]. Sulfotransferases generally facilitate a biochemical reaction in which a sulfo group is transferred from a donor to an acceptor molecule (Figure 1).



**Figure 1.** Depiction of a sulfotransferase-catalyzed reaction. This scheme illustrates the general chemical reaction catalyzed by sulfotransferases where a sulfo group is transferred from a donor molecule to an acceptor molecule.

Extensive research has been conducted on human sulfotransferases. To date, 14 cytosolic human sulfotransferases (hSULTs) have been identified and characterized (Table S1) [2–20]. These enzymes exhibit a high degree of sequence similarity, and all utilize 3′-phosphoadenosine 5′-phosphosulfate (PAPS) as the common sulfo donor. However, their sulfo acceptor preferences differ, giving each hSULT a unique function [27]. This variability is mostly present in the substrate-binding regions of the enzymes [28]. Specifically, the substrate binding loops of hSULTs demonstrate significant diversity [27]. The 14 distinct isoforms of human sulfotransferases (hSULTs) have tissue-specific expression and are responsible for the sulfonation of numerous small molecules, both endogenous and exogenous [28]. These enzymes are typically characterized by their dependency on PAPS for the sulfotransferase activity and are recognized by the presence of PAPS binding motifs [29]. In addition, hSULTs are categorized under the large class of aryl sulfotransferases (EC 2.8.2.1) [26]. Human cytosolic sulfotransferases facilitate sulfo transfer reactions, utilizing either a nonsequential (random Bi Bi) mechanism, where the substrates bind independently without intermediate formation, or a sequential (ordered Bi Bi) mechanism, where the binding of one substrate facilitates the attachment of the other [20,30,31]. PAPS-dependent aryl sulfotransferases are not exclusive to humans but are also found across various eukaryotic and prokaryotic species [7,9,11,17,21,22]. The widespread and prevalent nature of these enzymes is not fully understood. One hypothesis is that these enzymes might be involved in detoxification of molecules that are present at varying levels in the environments of these diverse life forms [30].

In the late 1980s, a series of studies demonstrated that the intestinal flora exhibited sulfation activity on small phenolic compounds [23,24,26]. These investigations revealed a new class of previously unknown microbial sulfotransferases, distinctive in its utilization of sulfo donors other than the mammalian default, 3′-phosphoadenosine 5′-phosphosulfate (PAPS). Interestingly, these microbial enzymes were capable of utilizing a variety of non-physiological phenolic sulfate esters as donors, with p-nitrophenol sulfate (pNPS) being predominantly used.

Due to the differences in donor specificities, reaction mechanisms, and kinetic profiles, these enzymes were designated as aryl-sulfate sulfotransferases (ASSTs), classified under EC 2.8.2.22 [32]. Beak et al. conducted studies on the distribution of ASSTs, showing their presence across both gram-positive and gram-negative bacterial species [33]. Furthermore, ASSTs were observed to exist both with and without signal peptides [33]. Cumulatively, these studies show that there are many variations in microbial ASSTs. The widespread prevalence of ASSTs and the current gap in understanding their functional importance demonstrate the need for further study.

The integral role of gut microbiota in metabolizing a range of xenobiotics and endogenous substances is already well-recognized [34]. As mentioned earlier, a common route for the metabolism of these compounds is via sulfonation. Motivated by these findings, we have delved into the genomic landscape of gut microbes, specifically searching for genes annotated as aryl-sulfate sulfotransferases. Interestingly, we discovered that *Sutterella*, a genus within the human gut microbiome, contains various members that harbor multiple genes annotated for these sulfotransferases. It is very unusual to find more than two annotated metabolic genes predicted to produce the same type of protein.

*Sutterella* species have been isolated from human fecal samples [35–41] and implicated in various human health conditions, including ulcerative colitis [42], autism spectrum disorders (ASD) [43,44], bacteremia [45], multiple sclerosis [46], and autoimmune based thyroid disease [47]. Additionally, *Sutterella* is known to degrade immunoglobulin A (IgA), suggesting a pro-inflammatory role [48]. Interestingly, in the context of ulcerative colitis, a decline in *Sutterella wadsworthensis* has been associated with drug-free remission [42]. Given its widespread presence and links to several human health concerns, we decided to investigate and understand the presence of annotated *asst* genes and their products (ASST proteins) in the genome of *S. wadsworthensis*. Our bioinformatics study has revealed a predominant presence of multiple aryl-sulfate sulfotransferase (ASST) enzymes harboring members within the genus *Sutterella*. These enzymes exhibit considerable sequence and structural homogeneity, albeit with enough variability that can have functional consequences. Additionally, while sharing similarities with prokaryotic ASSTs, such as those in *E. coli*, *Sutterella*'s sulfotransferases appear to have diverged evolutionarily from the PAPS-dependent class of sulfotransferases.

## 2. Results and Discussion

### 2.1. Distribution of Aryl-Sulfate Sulfotransferase (*asst*) Genes in Human Gut Microbiota

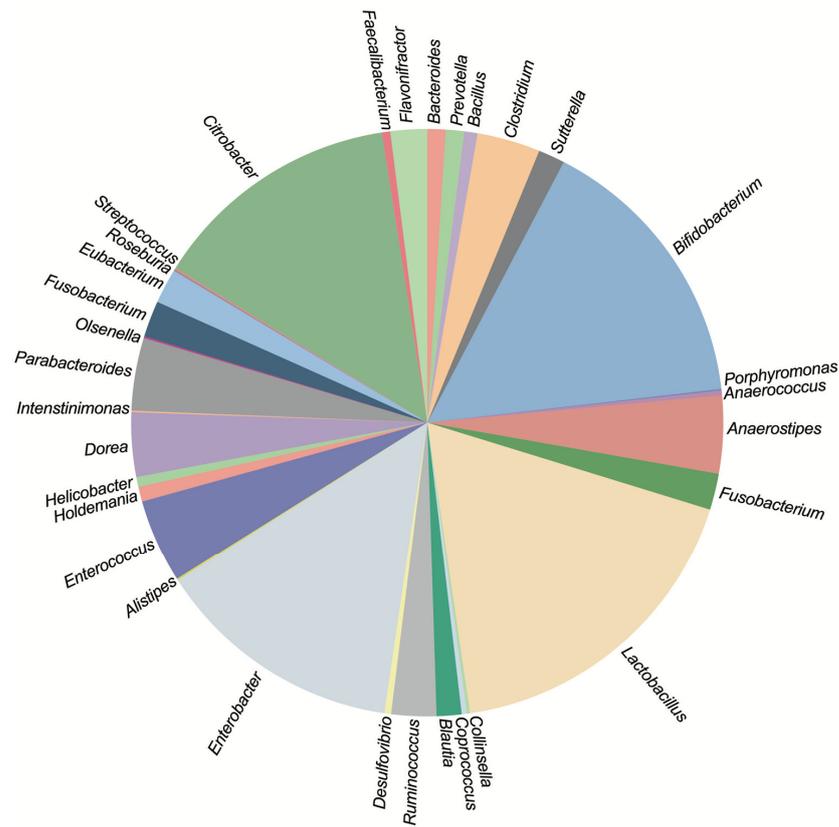
Aryl-sulfate sulfotransferases (ASSTs) are enzymes found extensively across various genera of gut microbiota. Figure 2 illustrates the prevalence of annotated *asst* genes within diverse gut microbial genera. The data depict the presence of these genes across different species and strains within each genus. The analysis omits numerous *E. coli* strains due to their high abundance and to avoid skewing the data representation.

Our research indicates a broad distribution of *asst* genes among gut microbes. Interestingly, a majority of these microbes possess only one or two annotated *asst* genes per organism. Genera such as *Lactobacillus*, *Bifidobacterium*, *Enterobacter*, and *Citrobacter* are predominant, collectively harboring 60% of all *asst* gene annotations identified in gut microbes. Despite the apparent ubiquity of these genes, the underlying reasons for their prevalence in human gut microbiota remain elusive.

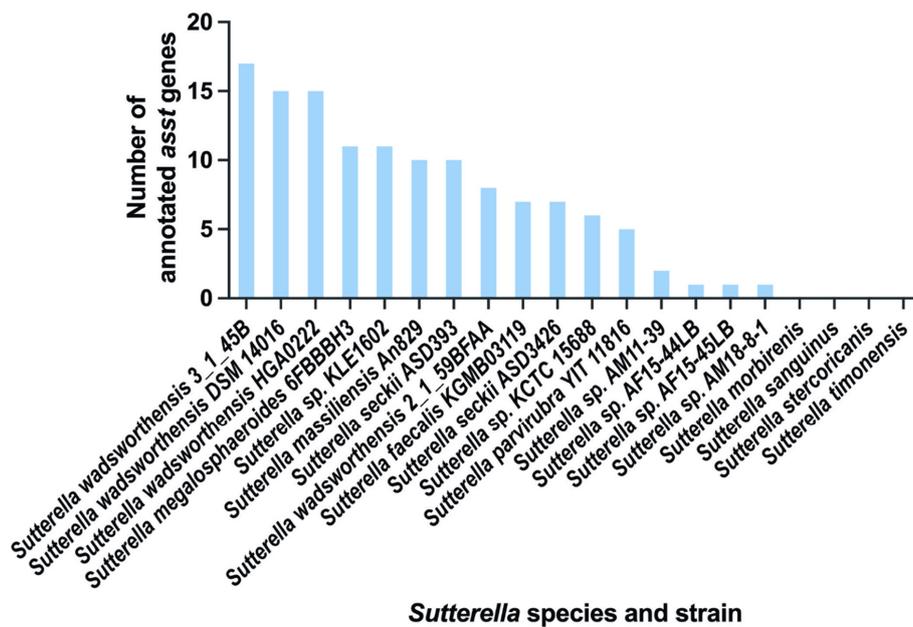
### 2.2. Variability of Annotated Aryl-Sulfate Sulfotransferase (*asst*) Genes in the Genus *Sutterella*

Our search through IMG/M database has uncovered multiple annotated aryl-sulfate sulfotransferases within the genus *Sutterella*. The bar graph in Figure 3 and Table S2 details the distribution of these genes among different species and strains of *Sutterella*. Notably, *Sutterella wadsworthensis* exhibits the largest number of annotated *asst* genes, with individual strains showing variability ranging from 8 to 17 genes.

Furthermore, *Sutterella megalosphaeroides* and *Sutterella* sp. KLE1602 each harbor a total of 11 annotated *asst* genes. Close behind, *Sutterella massiliensis* has 10 such genes. Distinct strains of *Sutterella seckii*, ASD393 and ASD3426, comprise 10 and 7 annotated *asst* genes, respectively. Other species, including *Sutterella faecalis*, *Sutterella* sp. KCTC 15688, and *Sutterella parvirubra*, display between 5 and 7 genes. Contrastingly, several strains, such as AM11-39, AF15-44LB, AF15-45LB, and AM18-8-1 of *Sutterella* sp., contain only 1 to 2 annotated *asst* genes. However, species like *Sutterella morbirenis*, *Sutterella sanguinus*, *Sutterella stercoricanis*, and *Sutterella timonensis* have no detectable annotated *asst* genes.



**Figure 2.** Distribution of annotated *asst* genes in the human gut microbes. The pie chart represents the distribution of annotated aryl-sulfate sulfotransferase (*asst*) genes across various known gut microbial genera. Each segment’s proportion reflects the relative count of *asst* genes within that particular genus.



**Figure 3.** Annotated *asst* genes within the genus *Sutterella*. The bar graph represents the quantity of annotated *asst* genes across various species and strains of *Sutterella*, highlighting the substantial variability within the genus.

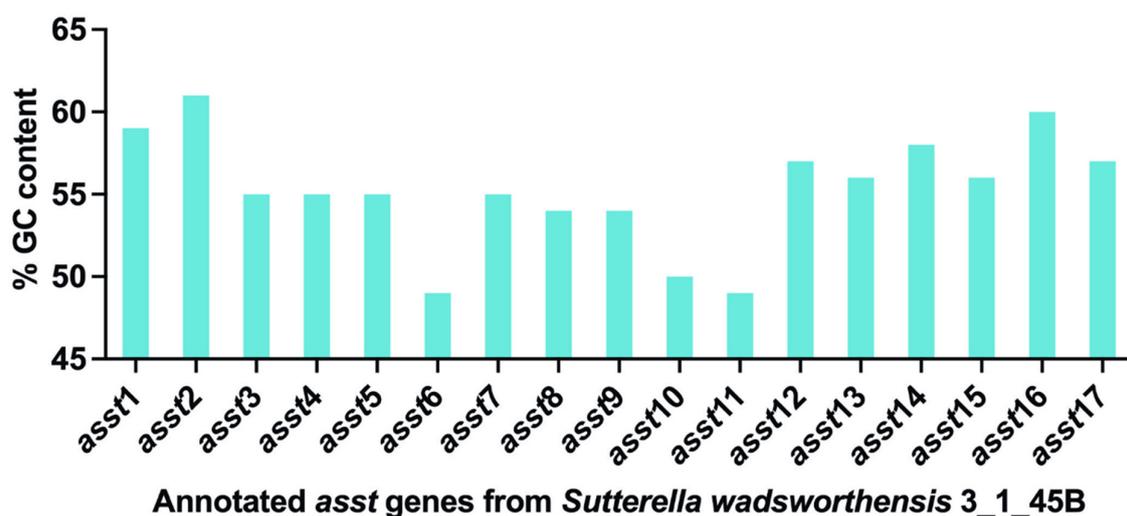
This variation within the same genus raises intriguing questions: Why do certain *Sutterella* species have a multitude of *asst* genes while others have none? Addressing these questions will be possible only after the development of advanced genetic manipulation tools specifically tailored for the *Sutterella* genus.

### 2.3. Bioinformatic Analysis of Annotated *asst* Genes in *S. wadsworthensis* 3\_1\_45B

Given that *S. wadsworthensis* has the highest count of annotated genes encoding aryl-sulfate sulfotransferases (*assts*), we focused on a detailed characterization of these genes and their corresponding enzyme products from the strain *S. wadsworthensis* 3\_1\_45B, which contains 17 identified *asst* genes. The comprehensive list and characteristics of these 17 genes and enzymes can be found in Table S3. For *asst* genes, we examined the percent GC content and the gene neighborhoods.

### 2.4. GC Content Variation in *asst* Genes

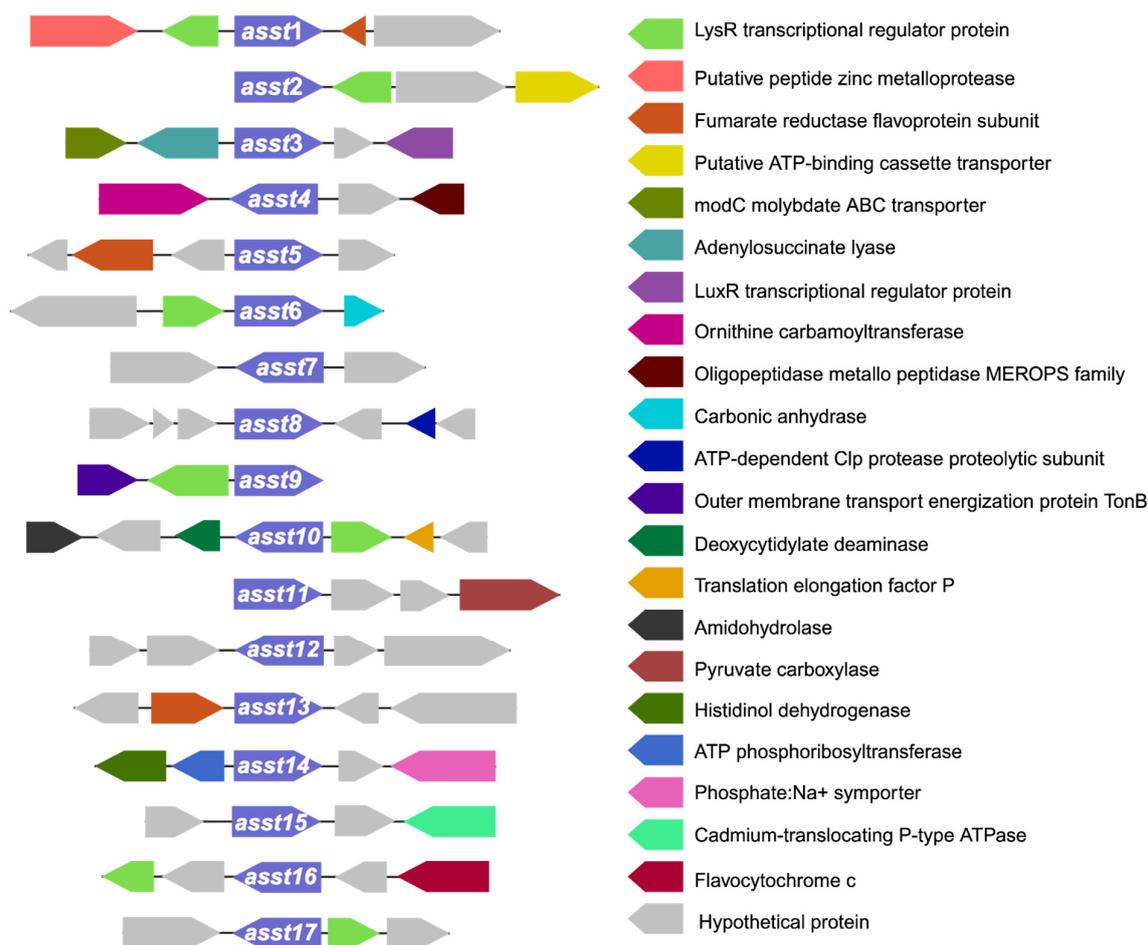
We first analyzed the percent GC content of the *asst* genes from *S. wadsworthensis* 3\_1\_45B, with data sourced from the IMG/M database. By plotting the GC content, we aimed to determine any variations or similarities that might exist within *asst* genes. The GC content of various *S. wadsworthensis* strains is generally around 55% [45]. The mean GC content of the *asst* genes was found to be 55.3%, with a median of 55%, aligning closely with the previously reported values for a GC content of *S. wadsworthensis* [45]. However, the GC content of the 17 annotated *asst* genes did display some diversity, ranging from a minimum of 49% to a maximum of 61% (Figure 4). It has been seen that the overall base composition of highly expressed genes can affect bacterial fitness. Additionally, mRNAs with higher GC content display higher stability [49]. GC content can also affect gene expression and regulation. This could also help organism to adapt to various environmental conditions. Thus, GC content variation in the annotated *asst* genes might have physiological relevance.



**Figure 4.** Variation in the %GC content of *asst* genes. This figure shows a bar graph representation of %GC content for all annotated *asst* genes from *S. wadsworthensis* 3\_1\_45B.

### 2.5. Analysis of Gene Neighborhoods for *asst* Genes in *S. wadsworthensis* 3\_1\_45B

Investigating the gene neighborhood is critical for understanding the physiological roles of specific genes and identifying potential operon structures, which may elucidate their functions. Interestingly, in *E. coli*, the *asst* gene is part of an operon associated with the formation of disulfide bonds [29]. Accordingly, such analyses are important in shedding light on the functions of the annotated sulfotransferases. Figure 5 illustrates the comprehensive analysis of gene neighborhoods of *asst* genes in *S. wadsworthensis* 3\_1\_45B.



**Figure 5.** Gene neighborhood maps of annotated *asst* genes. Each gene is depicted as an arrow and is represented by a unique color. All *asst* genes (from 1 to 17) are represented in purple color. Size of each arrow is approximately proportional to the gene size.

An analysis of the seventeen annotated *asst* genes has revealed that seven of these have the *lysR* gene in close vicinity (Figure 5 and Table 1). The *lysR* gene encodes LysR family transcription regulator. LysR-type transcriptional regulators (LTTRs) play dual roles as activators or repressors of gene expression [50]. In *E. coli*, for instance, the LysR-type regulator CysB governs the expression of genes crucial for sulfate assimilation and organic sulfur utilization [51]. This regulatory control extends to sulfate starvation responses in *P. putida* as well [52]. Furthermore, recent discoveries have identified a novel LysR type transcriptional regulator (LTTR) from *A. baumannii*, which is involved in controlling the expression of genes involved in the uptake and reduction of various sulfur compounds [53]. LTTRs are prolific within prokaryotes and are known to regulate a multitude of bacterial functions, such as stress response, antibiotic resistance, motility, quorum sensing, degradation of aromatic compounds, and biosynthesis of amino acids [54].

**Table 1.** List of genes found in the neighborhood of annotated *asst* genes that encode for *SwASST* proteins.

Annotated Genes in the Neighborhood of <i>asst</i> Genes	Number of Occurrences around <i>asst</i> Genes	Found in the Neighborhood of this <i>asst</i> Gene
LysR transcriptional regulator protein	7	1, 2, 6, 9, 10, 16, 17
Fumarate reductase flavoprotein subunit	3	1, 5, 13
Putative peptide zinc metalloprotease	1	1
Putative ATP-binding cassette transporter	1	2
LuxR transcriptional regulator protein	1	3
modC molybdate ABC transporter	1	3
Adenylosuccinate lyase	1	3
Ornithine carbamoyltransferase	1	4
Oligopeptidase metallo peptidase MEROPS family	1	4
Carbonic anhydrase	1	6
ATP-dependent Clp protease proteolytic subunit	1	8
Outer membrane transport energization protein TonB	1	9
Translation elongation factor P	1	10
Deoxycytidylate deaminase	1	10
Amidohydrolase	1	10
Pyruvate carboxylase	1	11
ATP phosphoribosyltransferase	1	14

Information from Figure 5 shown as a list.

Given that the ASST proteins are predicated to catalyze sulfo transfer reactions, it is plausible that LysR may influence the expression of ASSTs. Our findings indicate that *lysR* is a predominant gene in the vicinity of genes encoding *SwASST* proteins. Other genes located near *assts* are listed in Table 1, providing a comprehensive view of the gene neighborhood.

In the vicinity of *asst* genes, particularly those encoding *SwASSTs* 1, 5, and 13, we frequently identified the gene for the fumarate reductase flavoprotein subunit. This protein is integral to the process of anaerobic respiration [55], where sulfate can serve as one of the terminal electron acceptors. In addition, certain gut microbes are known to use naturally occurring organosulfonates, such as taurine, as terminal electron acceptors [56]. Given that aryl-sulfate sulfotransferases (ASSTs) are involved in the production of organosulfonates through the transfer of sulfates between organic molecules [23,24,26], it raises the question of whether fumarate reductase could facilitate the transfer of electrons to these organosulfonates. Could there be a coordinated expression between ASSTs and fumarate reductase, suggesting a metabolic linkage? This presents an interesting hypothesis, which requires experimental validation. Beyond the two genes commonly found near various *asst* genes, our analysis has identified additional genes situated in close proximity to single and specific *asst* genes encoding *SwASSTs*.

For instance, in the gene neighborhood of *asst1*, there is a gene that encodes a putative zinc metalloprotease. These enzymes are ubiquitous and versatile, often associated with pathogenicity and virulence in pathogenic microbes, though they are also present in nonpathogenic species. The presence of zinc metalloprotease is speculated to confer an evolutionary advantage to the microorganisms that retain it [57]. However, the connection between an ASST protein and zinc metalloprotease is unclear.

In the vicinity of the *asst2* gene, there is a gene that encodes a putative ATP-binding cassette transporter (ABC transporter). ABC transporters, present in both eukaryotes and prokaryotes, perform a wide range of functions. Among their roles in microbial organisms

is the transportation of sulfate, sulfonates, and sulfate esters [58,59]. Further research may reveal whether this ABC transporter is responsible for the translocation of the products generated by the activity of *SwASST2*.

Similar to the LysR family, the LuxR family transcriptional regulators can also function as either activators or repressors of gene expression, predominantly managing genes associated with quorum sensing. However, their regulatory influence extends beyond quorum sensing to other critical microbial biological functions [60]. One such case is a novel LuxR-type regulator involved in catechol metabolism within the prominent gut microbe, *E. lenta* [61]. A gene for LuxR family transcriptional regulator is also found adjacent to the *asst3* gene, which encodes *SwASST3*. Whether it serves as a regulator for the expression of the *asst* gene, potentially influencing the metabolism of sulfated molecules in *Sutterella* spp., is a question that remains open for investigation.

Adjacent to the *asst3* encoding *SwASST3* are two other genes in addition to the gene that produces LuxR regulator (Table 1). The first is an annotated gene that encodes molybdate ABC transporter. Previous studies have demonstrated the competitive interaction between sulfate ions and molybdate for uptake via similar transport systems [62,63], an observation that has been made in the small intestine of sheep [64]. This raises the question: might there be a similar competitive mechanism affecting the transport of the sulfated products created by *SwASST3* and molybdate?

The other gene in proximity to *asst3* encodes adenylosuccinate lyase, an enzyme implicated in purine biosynthesis via the de novo pathway. Its enzymatic activity is essential for the synthesis of AMP [65]. Any possible connection between an ASST protein and an adenylosuccinate lyase has not been explored before.

Surrounding the gene encoding *SwASST4*, two notable genes are present (Table 1). The gene for ornithine carbamoyltransferase, which modulates ornithine levels, is one such gene. Microbes capable of breaking down ornithine demonstrate a competitive advantage [66]. This raises the question: could this enzyme also confer a survival benefit to *Sutterella wadsworthensis* in a healthy gut, and might *SwASST4* play a role in this dynamic? Further investigation is needed to explore these possibilities. An additional gene adjacent to *asst4* is a gene for oligopeptidase A, part of the M03A family of metallopeptidases, known for their ability to cleave small peptides starting with alanine or glycine [67]. There are no reports of any interaction between an oligopeptidase A and an ASST.

In the vicinity of the gene (*asst6*) encoding *SwASST6*, lies a gene for a LysR family transcription regulator, previously discussed (Table 1). Another neighboring gene encodes a carbonic anhydrase, pivotal for maintaining  $\text{CO}_2$ ,  $\text{HCO}_3^-$ , and  $\text{H}^+$  balance within microbes and facilitating crucial exchange of these molecules with various metabolic pathways. Gut microbial carbonic anhydrases significantly differ from those in the host [68], suggesting that their activity levels can impact the survival of gut microbiota [69].

The gene neighborhood of *asst8* includes an ATP-dependent Clp protease gene (Table 1). Proteases like these are essential for the removal of damaged or misfolded proteins. Because of their critical role in protein degradation pathways, these proteins are imperative for microbial physiology [70].

The gene encoding *SwASST9* is associated with two distinct genes within its local genomic landscape (Table 1). In addition to the previously noted LysR family transcription regulator, this neighborhood also includes a gene for the outer membrane transport protein, TonB. TonB-dependent transporters are known to leverage the proton motive force to enable the translocation of substances across the outer membrane. Specifically, TonB plays a critical role in importing nutrients, particularly large polysaccharides derived from the diet, into gut microbes. The considerable size of these molecules necessitates an energy-dependent mechanism for membrane passage, a process facilitated by TonB-dependent transporters [71,72].

In addition to the above mentioned *lysR* family transcription regulator gene, three other annotated genes surrounding *asst10* gene that encodes *SwASST10* (Table 1). These genes are for translation elongation factor P, deoxycytidylate deaminase, and an amidohydrolase. The

translation elongation factor P is instrumental in microbial protein synthesis, particularly for polypeptides comprising polyproline sequences [73]. The deoxycytidylate deaminase plays a crucial role in the synthesis of the thymidine nucleotide, a building block of DNA [74]. The amidohydrolase, a versatile enzyme, is capable of cleaving a variety of bonds, such as carbon–oxygen, phosphorous–oxygen, phosphorous–sulfur, carbon–nitrogen, carbon–sulfur, and carbon–chlorine, with known involvement in the metabolism of xenobiotics [75].

Adjacent to the gene encoding *SwASST11*, pyruvate carboxylase has been identified, as listed in Table 1. This enzyme plays a pivotal role in anaplerosis, the process of replenishing intermediates of the tricarboxylic acid (TCA) cycle that are consumed by various biosynthetic pathways. Pyruvate carboxylase specifically catalyzes the formation of oxaloacetate, a key intermediate in the TCA cycle. Moreover, carboxylases are integral to carbon fixation reactions, facilitating the incorporation of inorganic carbon into organic molecules [76].

Surrounding the gene encoding *SwASST14*, three significant genes have been annotated: histidinol dehydrogenase, ATP phosphoribosyltransferase, and a phosphate:Na<sup>+</sup> symporter, as noted in Table 1 and Figure 5. The biosynthesis of histidine is a well-conserved pathway found across various life forms, with the exception of mammals [77]. ATP phosphoribosyltransferase and histidinol dehydrogenase are key enzymes in this pathway, representing the first and the final steps, respectively [78]. ATP phosphoribosyltransferase initiates the pathway by condensing 1-(5-phospho-D-ribosyl)-ATP with a diphosphate molecule. The phosphate:Na<sup>+</sup> symporter may facilitate the uptake of phosphate, the substrate required for this first reaction in histidine biosynthesis. The potential involvement of *SwASST14* in histidine biosynthesis or its regulation poses an intriguing question for future research endeavors.

Located in the genomic vicinity of the gene encoding *SwASST15* is a cadmium-translocating P-type ATPase (Table 1). P-type ATPases play a critical role in cellular homeostasis by regulating the intracellular concentrations of both essential and toxic transition metals. They achieve this by actively expelling toxic metals from the cells, thereby mitigating metal toxicity [79].

Lastly, there are two annotated genes adjacent to the gene encoding *SwASST16* (Table 1). One is the previously mentioned LysR family transcription regulator, and the other encodes flavocytochrome c. Flavocytochromes c are typically produced in large quantities under anaerobic conditions [80]. They are prominent in sulfate-reducing microbes, where soluble flavocytochrome c proteins transfer electrons to various acceptors, including sulfate, thiosulfate, sulfur, nitrate, and fumarate [81]. The involvement of flavocytochrome c proteins in the reduction of organosulfates has not been established in the literature. For this reason, the likelihood of *SwASST16* reaction products, which are organosulfates, acting as direct terminal electron acceptors for flavocytochrome c seems minimal. Nonetheless, it would be intriguing to explore whether these organosulfates participate indirectly in a broader metabolic network that includes both ASST and flavocytochrome c proteins, thereby contributing to the electron transfer to organosulfates through more complex pathways.

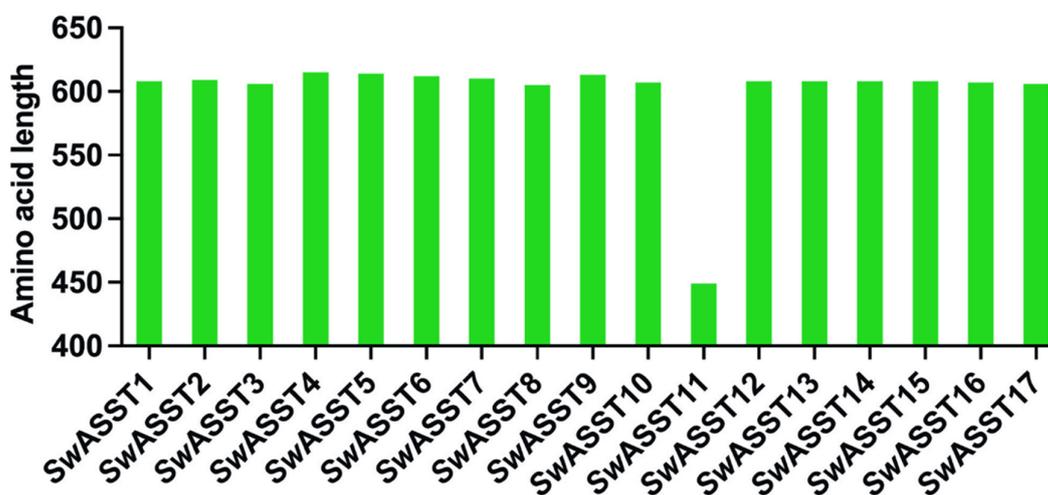
Our gene neighborhood analysis has indicated that *asst* genes from *S. wadsworthensis* 3\_1\_45B typically do not integrate into recognizable operons. Instead, they appear to exist independently within the genome, as depicted in Figure 5. This finding leads us to hypothesize that the *asst* genes generally function as isolated units rather than as part of operonic clusters. Nevertheless, to fully understand the biological relevance of this arrangement, experimental investigations into the interactions between *SwASSTs* and the products of the neighboring genes are warranted.

## 2.6. Bioinformatic Analysis of Predicted ASST Proteins from *S. wadsworthensis* 3\_1\_45B

For the predicted *SwASST* proteins, we assessed protein length, presence of signal peptides, probable cellular localization, presence of transmembrane regions, amino acid sequence conservation, and structural variability.

Protein length variation among ASST enzymes of *S. wadsworthensis* 3\_1\_45B. The data presented in Figure 6 and Table S3 detail the amino acid lengths of the 17 identified

ASST enzymes from *S. wadsworthensis* 3\_1\_45B. The majority of these enzymes consist of approximately 600 amino acids. Specifically, the longest ASST comprises 615 amino acids, while the shortest contains 449 amino acids. 16 of the 17 ASST proteins have lengths within a narrow range of 605 to 615 amino acids, exhibiting consistency in size among these enzymes.



**Predicted ASST from *Sutterella wadsworthensis* 3\_1\_45B**

**Figure 6.** Variation in the length of ASST enzymes of *S. wadsworthensis* 3\_1\_45B. The plot illustrates the range of amino acid counts as comparison of protein lengths across the 17 predicted ASST enzymes.

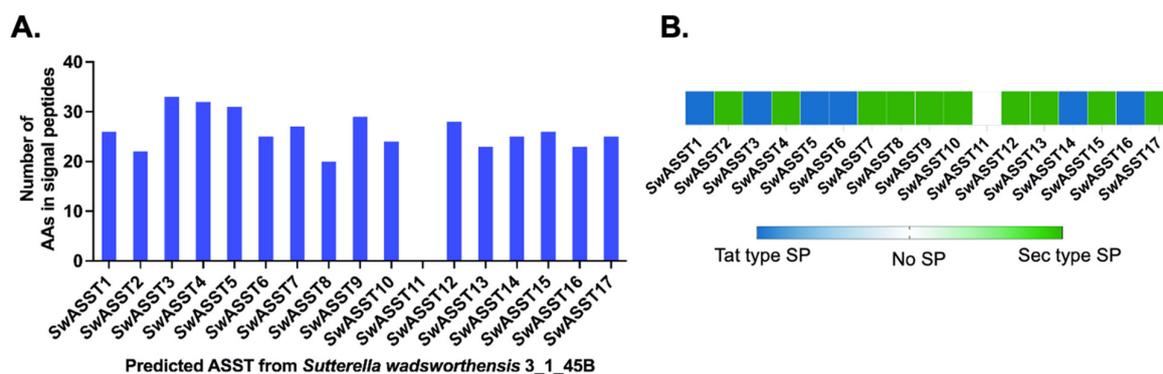
### 2.7. Analysis of Signal Peptides in SwASST Enzymes

Given that a well-characterized ASST from *E. coli* has a signal peptide [29], we conducted a sequence analysis to determine the prevalence and types of signal peptides in SwASSTs from *S. wadsworthensis* 3\_1\_45B. Our findings reveal that 16 of the 17 annotated SwASSTs are predicted to have signal peptides, suggesting their translocation across the cytoplasmic membrane (Figure 7A and Table S3). The analysis of these signal peptides was carried out using SignalP 6.0, confirming the annotations provided by the IMG/M database. Upon examining the lengths of these signal peptides, the average length was found to be 26 amino acids, with a median of 25.5 amino acids. Among the 16 SwASSTs with signal peptides, 10 (SwASSTs 2, 4, 7, 8, 9, 10, 12, 13, 15, 17) harbor the Sec-type signal peptides associated with the general secretion pathway, while 6 (SwASSTs 1, 3, 5, 6, 14, 16) contain the Tat-type signal peptides linked to the twin-arginine translocation pathway (Figure 7B and Table 2). In addition, there is significant variability in the amino acid sequences of both Sec and Tat type signal peptides across the different annotated SwASSTs (Figure 8A,B).

Amino acid sequences of signal peptides from 17 annotated SwASSTs with the description of the type of secretory pathways (Sec or Tat). Underlined motifs are for Tat type signal peptides harboring conserved twin-arginine residues.

Sec and Tat translocases are common mechanisms for the transfer of proteins across the cytoplasmic membrane. The Sec system mediates protein transport across the membrane in an unfolded state, whereas the Tat system transports proteins in a folded state [82]. These proteins have distinct N-terminal signal peptides, with the Sec pathway typically featuring a positively charged N-terminus, a hydrophobic core, and polar C-terminal residues (Figure 8A and Table 2). The Tat pathway signal peptides are more conserved, particularly the twin-arginine residues (Figure 8B). However, SwASST14 and SwASST16 diverge from this pattern, presenting a lysine (K) in place of the first arginine (R) in the conserved region—a rare variant with potential implications for translocation efficiency, as reported in literature [83,84]. This observation suggests that SwASST14 and 16 may have a lower translocation efficiency than other SwASSTs with canonical twin-arginine motifs. Ad-

ditionally, both Sec and Tat signal peptides frequently contain the conserved A-X-A motif at the C-terminus, a feature commonly found in N-terminal signal peptides [85].

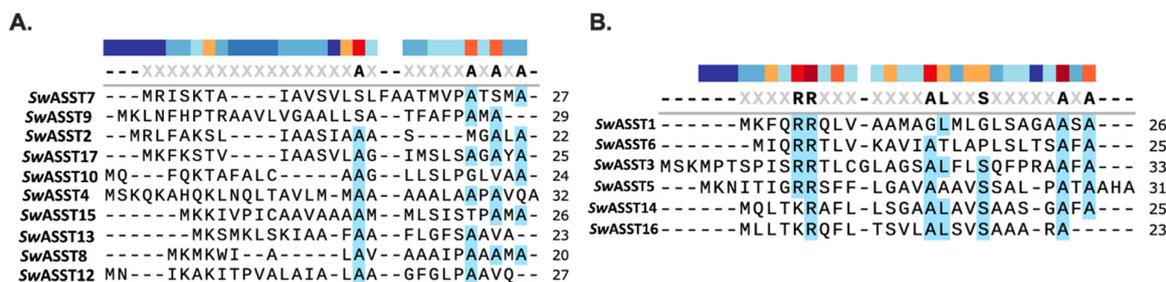


**Figure 7.** Signal peptide characteristics in predicted *SwASSTs*. (A) Depicts the length of signal peptides in amino acid (AA) residues. (B) Categorizes the signal peptides by type (Sec or Tat) within the *SwASSTs*, where Sec type is represented as green and Tat type as blue.

**Table 2.** Signal peptide sequences and types of *SwASSTs*.

Predicted Protein	Type of SP	Signal Peptide Sequence
<i>SwASST1</i>	Tat type	MKFQRRQLVAAMAGMLGLSAGAASA
<i>SwASST2</i>	Sec type	MRLFAKSLIAASIAAASMGALA
<i>SwASST3</i>	Tat type	MSKMPTSPISRRLCGLAGSALFLSQFPRAAFA
<i>SwASST4</i>	Sec type	MSKQKAHQKLNQLTAVLMMAAAAALAAPAVQA
<i>SwASST5</i>	Tat type	MKNITIGRRSFFLGAVAAAVSSALPATAAHA
<i>SwASST6</i>	Tat type	MIQRRTL <sup>U</sup> VKAVIATLAPLSLTS <sup>A</sup> FA
<i>SwASST7</i>	Sec type	MRISKTAIAVSVLSLFAATMVPAT <sup>S</sup> MA
<i>SwASST8</i>	Sec type	MKMKWIALAVAAAIPAAAMA
<i>SwASST9</i>	Sec type	MKLN <sup>F</sup> HPTRA <sup>A</sup> VLVGAALLSAT <sup>F</sup> FAFPAMA
<i>SwASST10</i>	Sec type	MQFQKTAFALCAAGLLSLPGLVAA
<i>SwASST12</i>	Sec type	MNIKAKITPVALAIALAAGFGLPA <sup>A</sup> VQA
<i>SwASST13</i>	Sec type	MKSMKLSKIAAFAAFLGFSAAVA
<i>SwASST14</i>	Tat type	MQLTKRAFLLSGAALAVSAASGAFA
<i>SwASST15</i>	Sec type	MKKIVPICA <sup>A</sup> VAAAAMMLSISTPAMA
<i>SwASST16</i>	Tat type	MLLTKRQFLTSVLALS <sup>V</sup> SAAARA
<i>SwASST17</i>	Sec type	MKFKSTVIAASVLGIMSL <sup>S</sup> SAGAYA

Proteins secreted through the Sec or Tat pathways may localize to the periplasm, integrate into the cytoplasmic membrane, or be exported outside the cell via other secretion systems, especially in gram-negative bacteria [86]. Periplasmic and extracellular proteins typically follow a SecB-mediated route, with extracellular proteins requiring an additional translocation step across the outer membrane via Type II or Type V secretion systems (T2SS or T5SS) [86]. Genomic analysis of *S. wadsworthensis* 3\_1\_45B revealed the presence of T2SS components, suggesting that some *SwASSTs* might be T2SS substrates. Table 3 combines our findings regarding *SwASSTs* harboring the Sec or Tat signal peptides and proposes their potential cellular destinations.



**Figure 8.** Sequence alignments for signal peptides of *SwASSTs*. (A) Sequence alignment for Sec type signal peptides of *SwASSTs*. (B) Sequence alignment for Tat type signal peptides of *SwASSTs*. Color bar above the sequences shows the amino acid conservation at that position, where dark red is highly conserved and dark blue indicates the least conserved position.

**Table 3.** Potential cellular destinations of *SwASSTs*.

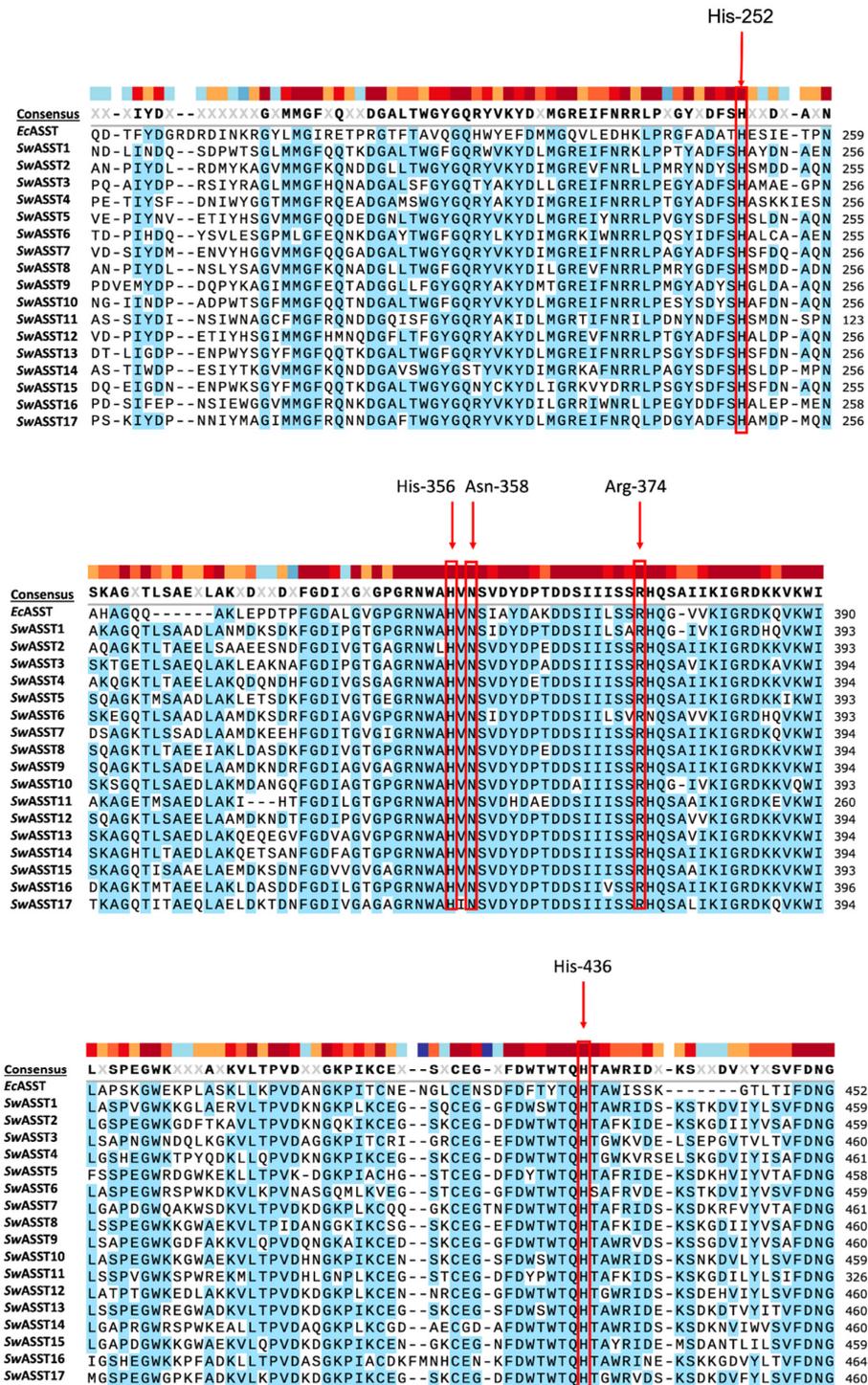
Predicted Protein	Signal Peptide Type	Sec Guided (Unfolded State) Periplasmic/Extracellular Proteins	Tat Guided (Folded State) Periplasmic/Extracellular Proteins
<i>SwASST1</i>	Tat		
<i>SwASST2</i>	Sec	+	
<i>SwASST3</i>	Tat		+
<i>SwASST4</i>	Sec		
<i>SwASST5</i>	Tat		+
<i>SwASST6</i>	Tat		+
<i>SwASST7</i>	Sec		
<i>SwASST8</i>	Sec		
<i>SwASST9</i>	Sec		
<i>SwASST10</i>	Sec	+	
<i>SwASST11</i>	-		
<i>SwASST12</i>	Sec		
<i>SwASST13</i>	Sec	+	
<i>SwASST14</i>	Tat		
<i>SwASST15</i>	Sec		
<i>SwASST16</i>	Tat		+
<i>SwASST17</i>	Sec		

The table indicates the type of signal peptides from *SwASSTs* and possible cellular destinations.

### 2.8. Sequences Similarity Analysis of *SwASSTs*

For the multiple sequence alignments, *E. coli* ASST (*EcASST*) served as a reference sequence because it has been studied well in the context of structure–function relationships [29]. Utilizing the Clustal Omega tool from UniProt for multiple sequence alignment, we observed that all the key active site residues of *EcASST*, crucial for catalysis are all well conserved across all 17 *SwASSTs*, especially residues corresponding to His-252 (H252), His-356 (H356), Asn-358 (N358), Arg-374 (R374), and His-436 (H436) [29] (Figure 9). His-436 is the most important catalytic residue that becomes transiently sulfated during the catalysis by *E. coli* sulfotransferase. This happens due to two half-reactions that occur during the ASST catalysis. In the first half-reaction, a donor adds a sulfate group to the active site histidine residue. In the second half reaction, this sulfate group from a histidine residue is transferred to the acceptor molecule, which completes the catalytic cycle. The multiple sequence alignment of

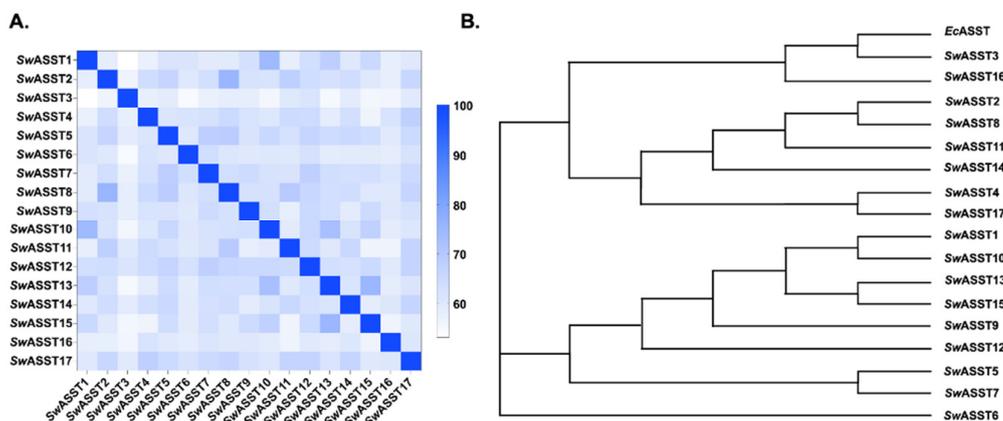
all 17 *SwASST* enzymes confirms that this catalytic histidine residue is perfectly conserved, suggesting a similar catalytic mechanism may be at play within these enzymes (Figure 9).



**Figure 9.** Multiple sequence alignment of *SwASSTs*. Alignment of *SwASSTs* with *EcASST* highlighting catalytically crucial residues in red. Color bar above the consensus line shows the amino acid conservation at that position. Dark red is highly conserved, while dark blue has little to no conservation.

This alignment revealed an average sequence similarity of 62% among all 17 *SwASSTs*. The range of similarity spanned from a minimum of 53% to a maximum of 75.5%, with the median and mode being 62% and 61.25%, respectively (Figure 10A). The degree of similarity

indicates that while the *Sw*ASST enzymes share a common structural framework, there is sufficient variation to suggest they may have specialized functions. Interestingly, *Sw*ASST2 and *Sw*ASST8 exhibited the greatest sequence similarity at 75.5%, whereas *Sw*ASST1 and *Sw*ASST3 shared the least similarity at 53%.



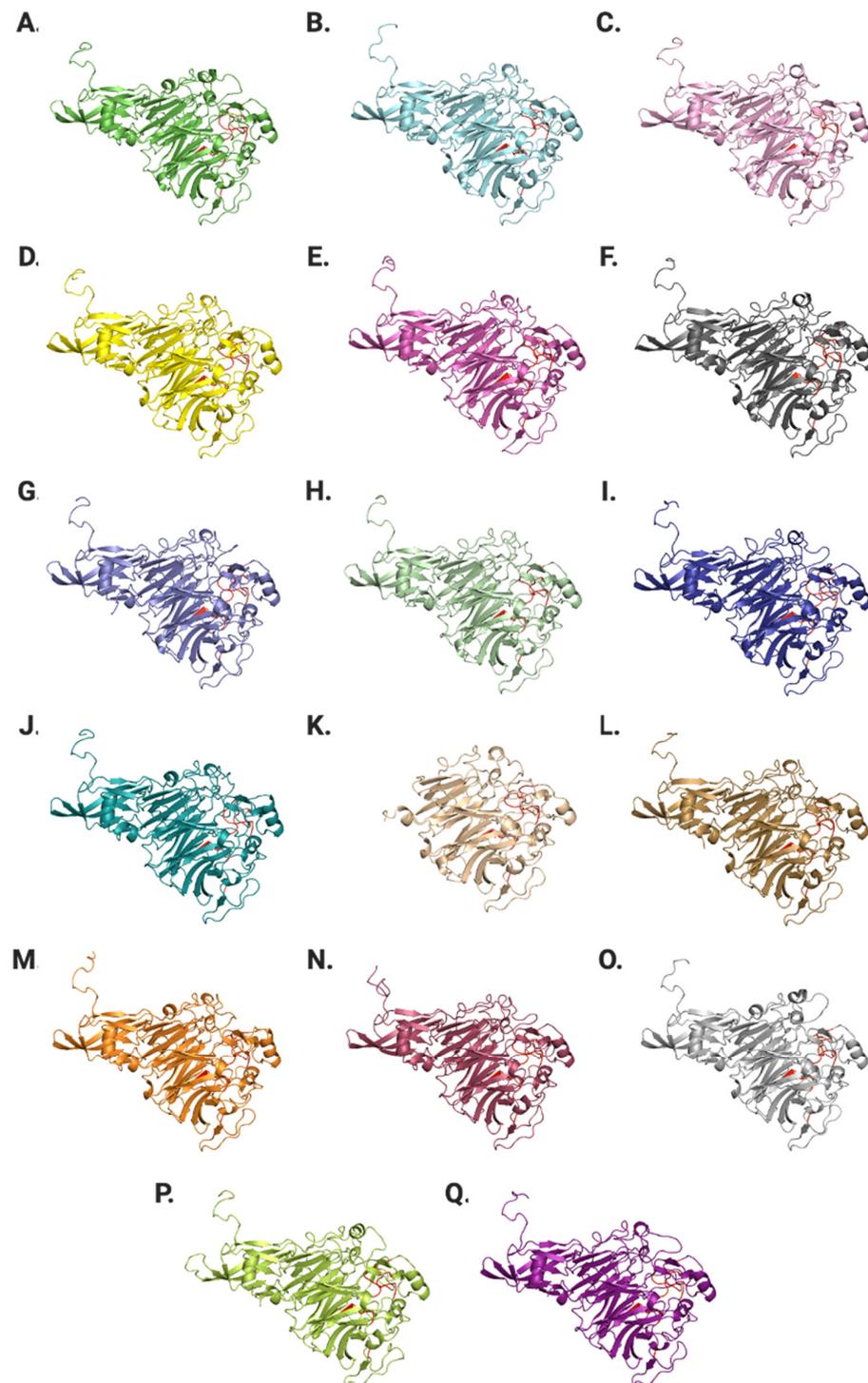
**Figure 10.** Sequence similarity analysis of *Sw*ASSTs from *S. wadsworthensis* 3\_1\_45B. (A) Percent identity matrix (B) Phylogenetic tree of the 17 predicted *Sw*ASSTs and *Ec*ASST (*Ec*ASST) from *E. coli* strain CFT073.

Additionally, our sequence alignments show that *Sw*ASSTs harbor more sequence similarities with each other than with *Ec*ASST. There are residues that are absolutely conserved in *Sw*ASSTs but are completely different in *Ec*ASST (Figure S1). For example, there are two conserved asparagine residues in all *Sw*ASSTs (1st one around 160–163 in and 2nd one around 161–164 in all *Sw*ASSTs except in the shorter *Sw*ASST11 where these are located at positions 28 and 29), which are not conserved in *Ec*ASST. *Ec*ASST has histidine and glycine residues in this place. Similarly, a threonine (T166) is replaced by an absolutely conserved glycine (164–167) residue in all *Sw*ASSTs. There are more such examples that can be seen in extended sequence alignment of these proteins. Based on sequence alignment, it is apparent that *Sw*ASST3 and 16 are more closely related to *Ec*ASST protein (Figure 10B, phylogenetic tree).

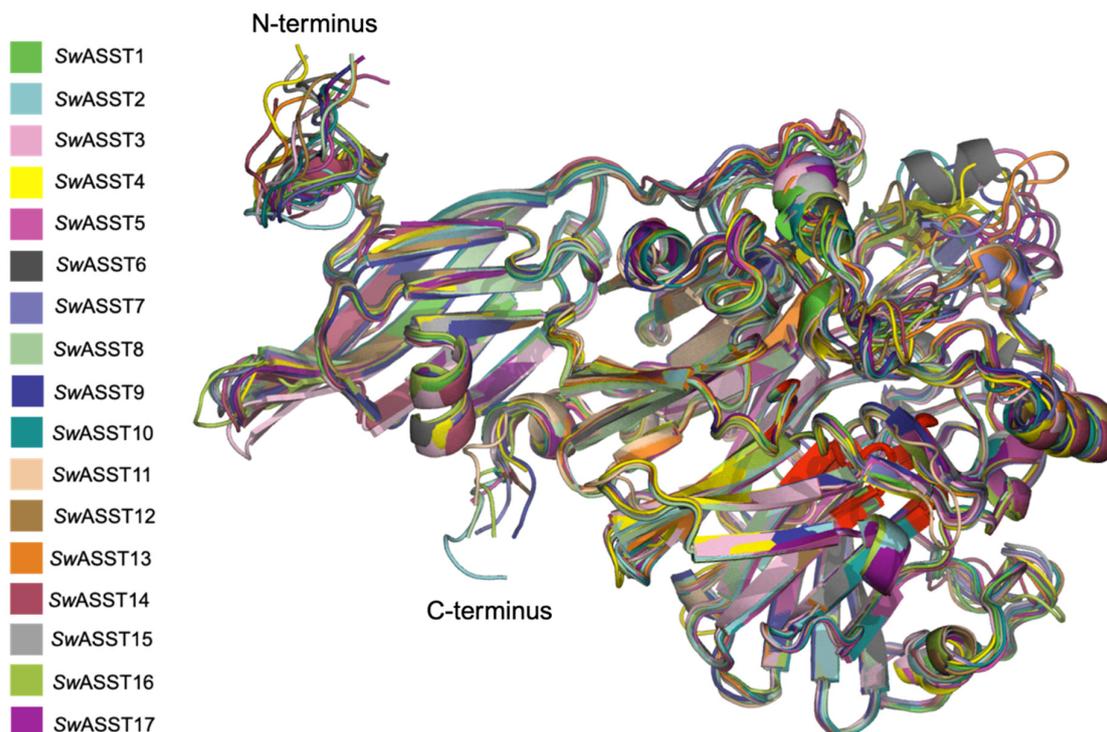
### 2.9. Structural Variations in *Sw*ASST Proteins

To understand differences at the structure level, structures of all 17 *Sw*ASSTs were generated using AlphaFold2. The resulting confidence levels of each model, measured by the local distance difference test (IDDT) score, are plotted against their respective amino acid positions, which were generated by AlphaFold2 (Figure S2) [87]. Generally, the IDDT confidence plots show decreased prediction reliability around residue positions 200, 350, and 550 for most *Sw*ASSTs. However, *Sw*ASST3, *Sw*ASST12, and *Sw*ASST16 exhibit less pronounced dips in confidence around residue 550. *Sw*ASST11, being smaller, displays a distinct confidence profile with no notable decrease in prediction reliability around the residues corresponding to position 550 in the other *Sw*ASSTs, such as in *Sw*ASSTs 3, 12, and 16. These observations highlight subtle structural differences among the *Sw*ASST proteins, possibly reflecting variations in their functional attributes. Structural alignment of *Sw*ASSTs, conducted using PyMol, revealed a shared common fold among these enzymes (Figure 11). The ability to superimpose these structures, shown in Figure 12, further supports their structural convergence. Additionally, an average root-mean-square deviation (RMSD) of 0.52 Å across all 17 aligned structures suggests a high degree of similarity [88]. While the overall secondary structure elements, such as alpha helices and beta sheets, are consistent across the majority of *Sw*ASSTs, *Sw*ASST11 deviates slightly, displaying two fewer beta sheets at the N-terminus, likely due to its shorter length. Additionally, a beta

sheet spanning residues 318–323 in *SwASST1* is oriented in the reverse direction compared to its counterparts in other *SwASST* enzymes.



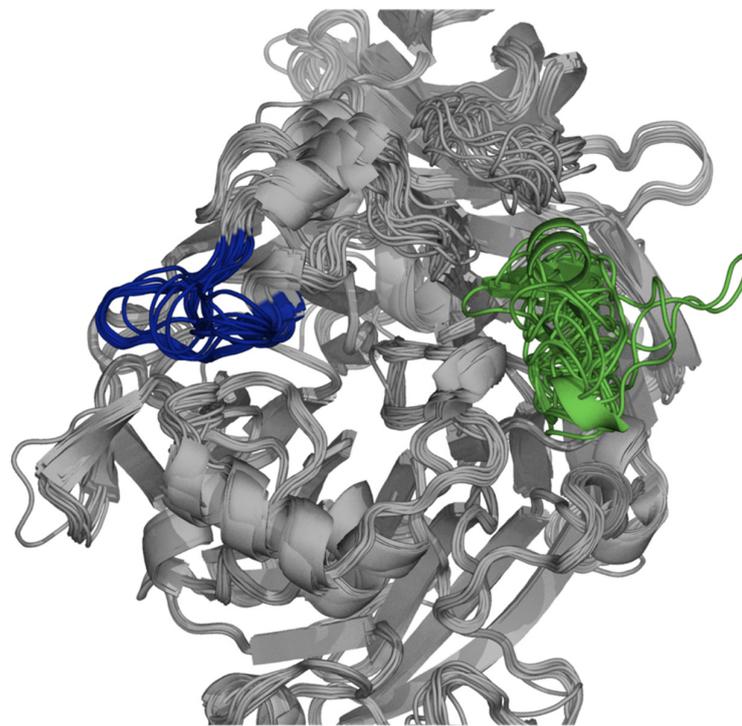
**Figure 11.** Structures of *SwASST* proteins generated by AlphaFold2. From (A–Q) are predicted structures of *SwASST1* to *SwASST17*. Conserved regions harboring active site residues are highlighted in red.



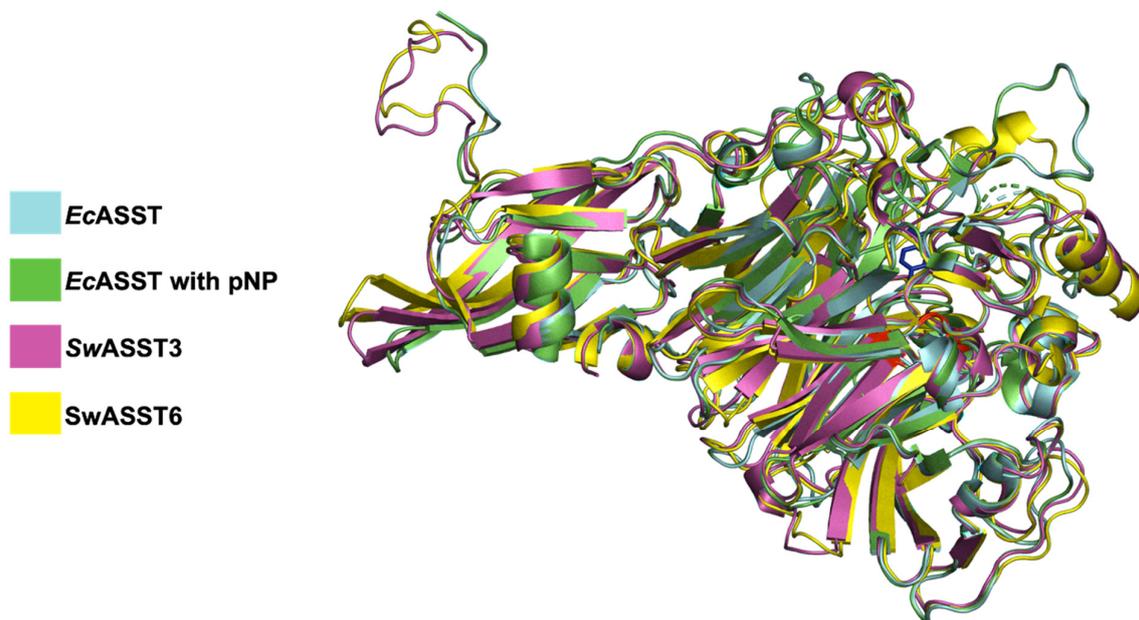
**Figure 12.** Structural alignment of *SwASST* proteins. The figure depicts aligned AlphaFold2-derived structures of *SwASSTs* with conserved regions highlighted in red.

As can be seen in Figure 12, most areas of *SwASSTs* are completely superimposable and very well aligned. However, there are some regions where we see a lower degree of structural alignment (Figure 13, blue and green areas). It is interesting to note that the areas with variable structural alignments can have either high or low amino acid sequence conservation. One such region is at the N-terminus with the conserved motif V/s/t-W-N-N-P-X-G-G-A-L/m/v-E-W (Figure 13, blue cluster, Figure S3), displays considerable variation in the loop positions among the aligned *SwASST* structures. Contrastingly, at the C-terminus, which is the other variable region (Figure 13, green cluster, Figure S3), the level of structural alignment diminishes relative to the rest of the *SwASST* structures. This region is characterized by low sequence conservation and exhibits the most significant variability in secondary structural elements across the *SwASST* family.

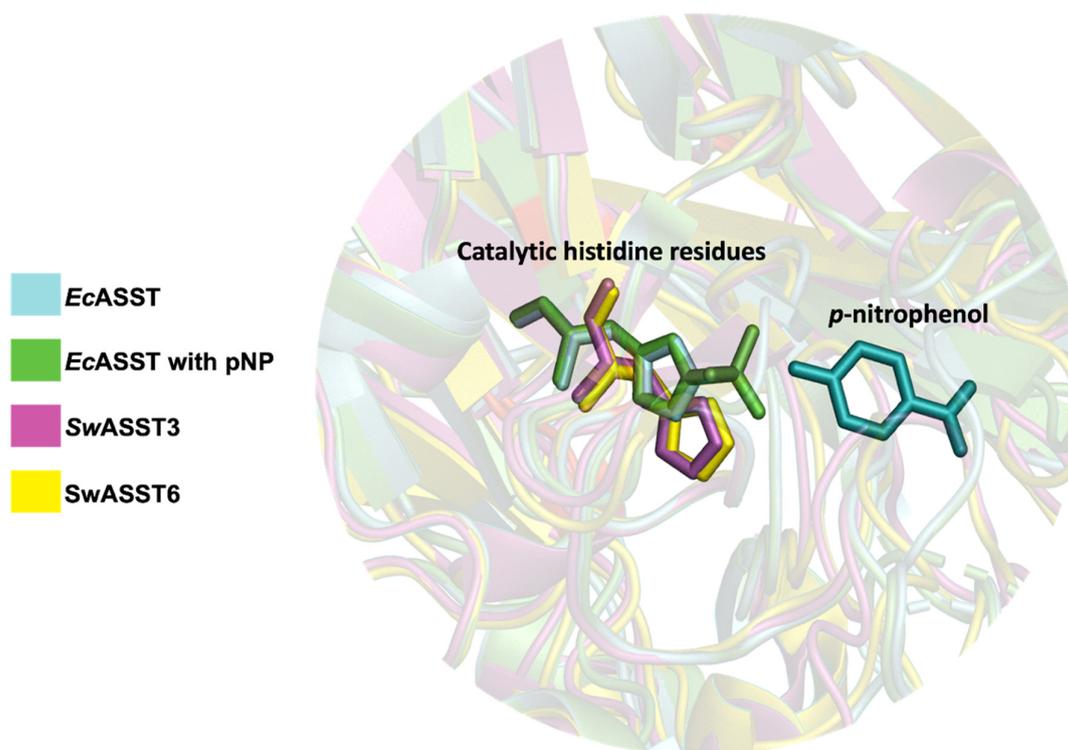
Sequence similarity studies and phylogenetic analyses indicate that *SwASST3* is evolutionarily closest to *E. coli* ASST (*EcASST*), while *SwASST6* is the most divergent (Figure 10B). Structural comparisons of *EcASST* [29] with *SwASST3* and *SwASST6* exhibit a substantial degree of alignment across the three structures (Figure 14). However, a loop region spanning residues 151–164 present in *EcASST* is absent in *SwASST3* and *SwASST6*. Additionally, *SwASST6* uniquely features an alpha helix and a loop between residues 555 and 568. We also included a ligand-bound structure of *EcASST* in our alignment (Figure 14). In the ligand-bound structure of *EcASST* (green), the active site binds a molecule of para-nitrophenol (pNP), a reaction product of *EcASST* with the sulfate donor para-nitrophenyl sulfate (pNPS). Figure 15 presents a detailed view of the active site, showing the ligand bound within *EcASST* and the catalytic histidine residues from all three structures. While the histidine residues from the two *E. coli* structures align perfectly with each other, the histidine residues from the *SwASSTs* align amongst themselves but show slight positional differences when compared to the catalytic histidine residue of *EcASST*. This observation suggests subtle variations in the active site residue positioning in *SwASSTs* relative to *EcASST*.



**Figure 13.** Unique clusters in *Sw*ASSTs alignments. Dark blue region represents the N-terminus cluster, and the C-terminus cluster is depicted in green. These regions show a lower degree of structural alignments across the *Sw*ASSTs.



**Figure 14.** Alignment *Sw*ASST3 and *Sw*ASST6 with *Ec*ASST. The figure shows aligned structures of *Sw*ASST3 (pink) and *Sw*ASST6 (yellow) with free (cyan, PDB ID: 3ELQ) and ligand bound (green, PDB ID: 3ETT) structures of *Ec*ASST. There is a molecule of para-nitrophenol (shown as sticks in blue color) bound in the active site of *Ec*ASST (green, PDB ID: 3ETT). Regions in red show catalytic residues from *Ec*ASST (His-252, His-356, Asn-358, Arg-374, and His-436).



**Figure 15.** Active site of *SwASSTs* (3 and 6) and *EcASST* highlighting the catalytic histidine residues. The figure displays a zoomed-in active site of *SwASST3* (pink), *SwASST6* (yellow), *EcASST* (cyan, free), and *EcASST* (green, pNP bound). Catalytic histidine residues are presented as sticks and highlighted in the same colors for each of the corresponding structures. pNP is shown as sticks (in cyan).

#### 2.10. Different Classes of Sulfotransferases from the Members of the Human Gut Microbiome and Their Comparison to Human Sulfotransferases

Sulfotransferases have historically been categorized based on their dependency on PAPS as either PAPS-dependent, using PAPS as a sulfo donor, or PAPS-independent [29]. Human sulfotransferases, typically around 300 amino acids in length, exclusively use PAPS, whereas microbial aryl-sulfate sulfotransferases (ASSTs), also known as PAPS-independent sulfotransferases, are generally larger. Recent findings, however, have identified PAPS-dependent sulfotransferases in gut commensals; these microbial enzymes are longer than their human counterparts, averaging around 370 amino acids, yet smaller than ASSTs [21,22]. To understand the prevalence of sulfotransferase classes, we explored the genomes of dominant human gut microbes, using protein blast searches with either *SwASST1* (a PAPS-independent aryl-sulfate sulfotransferase) or *BT\_0416* (a PAPS-dependent cholesterol sulfotransferase from *Bacteroides thetaiotaomicron*) as references. The results demonstrate a wide distribution of both sulfotransferase types across gut microbiome members (Figure S4). Specifically, the *Bacteroides* genus predominantly carries PAPS-dependent sulfotransferases, with *Parabacteroides* following closely (Figure S5), whereas PAPS-independent sulfotransferases (aryl-sulfate sulfotransferases) are abundant in *Sutterella* (Figure S6). Sequence (Figure S7) and structural (Figure S8) alignments were performed to elucidate the differences between these enzymes. The sequence similarity between *SwASST1* and other PAPS-dependent sulfotransferases (*BT\_0416*, hSULT1A1, and hSULT2B1) is very low. Within the PAPS-dependent class, however, there is higher sequence conservation. Structurally, there is a significant divergence between PAPS-independent (*SwASST1*) and PAPS-dependent sulfotransferases (Figure S8A). Human PAPS-dependent sulfotransferases align closely with each other and are completely superimposable (Figure S8B). When aligning all three PAPS-dependent sulfotransferases, including two from humans (hSULT1A1, and hSULT2B1) and one from gut microbes (*BT\_0416*), some overlapping regions are apparent, but the structures are not completely superimposable (Figure S8C), particularly a large alpha-helix

in *BT\_0416* that does not align with the others. A striking contrast is observed in the secondary structure composition between these two types of sulfotransferases. PAPS-dependent sulfotransferases are rich in alpha helices (Figure S8B,C), whereas PAPS-independent *Sw*ASSTs feature beta sheets as their dominant structural elements (Figure 12). This contrast suggests a divergent evolutionary path for these two classes of sulfotransferases.

### 3. Materials and Methods

#### 3.1. Members of the Human Gut Microbiome Harboring Annotated *asst* Genes

IMG/M database (Department of Energy, Berkeley, California, USA) was used to search for the prevalent human gut microbes [89,90] that contain annotated genes for *asst* that are predicted to produce the protein aryl-sulfate sulfotransferases (ASSTs). A search of IMG/M database with either enzyme name, aryl-sulfate sulfotransferases or the enzyme ID, EC 2.8.2.22, produced lists of genomes containing predicted *asst* genes. By selecting genome name under the filter column, lists of predicted ASSTs for each specific genus were collected. The total number of sulfotransferase genes present in a genus was counted by adding annotated genes from all the species under that genus.

#### 3.2. Annotated *asst* Genes from the Genus *Sutterella*

By selecting genus *Sutterella* in the IMG/M database, we were able to collect different species and strains harboring annotated *asst* genes. From this, *Sutterella wadsworthensis* 3\_1\_45B was selected due to the number of annotated sulfotransferases found in its genome. Locus tags, gene IDs, GC content, and gene neighborhoods were retrieved for all annotated *asst* genes via IMG/M database.

#### 3.3. Properties of Predicted ASST Proteins from *Sutterella wadsworthensis* 3\_1\_45B

Amino acid sequences for all 17 predicted ASST proteins from *S. wadsworthensis* 3\_1\_45B were obtained from the IMG/M database and UniProt and were aligned to confirm that amino acid sequences retrieved from both sources are exactly the same.

#### 3.4. Analysis of Signal Peptides

Amino acid sequences for all predicted sulfotransferases were analyzed with SignalP 6.0 using preset parameters to gain insights into the presence and variability of signal peptides. These collected datasets were also verified with IMG/M database. In addition to providing information about the presence or absence of signal peptides, SignalP 6.0 also allowed for identification of the possible signal peptide type, along with the length of the signal.

#### 3.5. Multiple Sequence Alignment

Multiple sequence alignment for all 17 ASST amino acid sequences was performed with the help of UniProt alignment tool, which utilizes the Clustal Omega program 2.1. This analysis provided information about sequence conservation in predicted ASST proteins. Additionally, percent identity matrix for these 17 ASST sequences was obtained from the alignments. Multiple alignments were performed with and without the signal peptides. Alignments with signal peptides were generated to understand differences in these regions. Alignments without signal peptides were generated to understand the variability in ASST sequences.

#### 3.6. ASST Structure Predictions with AlphaFold2

Amino acid sequences of ASSTs without signal peptides were uploaded to Collaboratory AlphaFold2 (ColabFold v1.5.3: AlphaFold2 using MMseqs2, DeepMind Technologies Limited) and default settings were used for all runs to predict protein structures [91,92]. The three conserved regions found from the multiple sequence alignment of *Sw*ASSTs with *Ec*ASST were highlighted in red in all ASST structures.

### 3.7. ASST Structural Alignments with PyMOL

The PyMOL (version 2.5.4, Schrödinger) alignment tool with the default parameters and align command with five iteration cycles and a cutoff of 2 Å was utilized to align all the structures except for the alignment 'c' described below. These alignments are as follows: (a) For the alignment among all 17 ASSTs, all structures were aligned to *SwASST1*. (b) The alignment of *E. coli* ASST (*EcASST*) with *SwASST3* and *SwASST6*, where two separate PDB structures of *EcASST* were utilized. Structure with PDB ID 3ETT has para-nitrophenol (pNP) bound to the active site, while 3ELQ has no bound ligands in the active site. (c) The alignment of *SwASST1* with *Bacteroides thetaiotaomicron* VPI-5482 sulfotransferase *BT\_0416* (structure generated via AlphaFold2), human sulfotransferase 1A1 (hSULT1A1, PDB ID 1LS6), and human sulfotransferase 2B1b (hSULT2B1b, PDB ID 1Q1Z). hSULT2B1b, hSULT1A1, *BT\_0416*, and *SwASST1* were aligned using the cealign command in PyMOL's alignment tool due to low sequence similarity.

### 3.8. Search for PAPS-Dependent and PAPS-Independent Sulfotransferases in the Members of the Human Gut Microbiome

Using NCBI protein blast (National Library of Medicine, Bethesda, MD, USA), two separate blasts, one with a sulfotransferase from *Bacteroides thetaiotaomicron* sulfotransferase, *BT\_0416* and another one with an aryl sulfate sulfotransferase from *Sutterella wadsworthensis* 3\_1\_45B, *SwASST1* were performed against members of the human gut microbiome. *BT\_0416* is PAPS-dependent sulfotransferase and *SwASST1* is a PAPS-independent sulfotransferase. *Bacteroides* and *Sutterella* were each excluded from their respective blast searches. Each genus of gut microbes was entered into the filter bar to determine if that genus contained PAPS-dependent and/or PAPS-independent sulfotransferases. From this search a binary dataset was created where presence was marked by 1 and absence was marked by 0. This dataset was utilized to create a heatmap. Predicted sulfotransferase proteins from different species and strains of each genus were added together to calculate the total number of predicted sulfotransferases (PAPS-dependent or independent) per genus.

## 4. Conclusions

Sulfotransferases from *Sutterella wadsworthensis* 3\_1\_45B exhibit a mixture of shared and unique structural characteristics. While these enzymes generally align with a common structural fold, certain regions display variability in both sequence conservation and structural configuration. The enzymes evaluated in this study are recognized for their activity on phenolic compounds, as established by prior research involving *E. coli* ASSTs and ASSTs from other gut microbes. These enzymes facilitate the transfer of sulfo groups, which modulates the levels of phenolic molecules. Uniquely, within the *Sutterella* genus, certain members possess a diversity of ASSTs within a single organism, with significant sequence variation—no two ASSTs examined here exhibit more than 75% sequence similarity. This is notable since proteins from the same genus and species, but different strains, that catalyze identical reactions usually display around 90% or more sequence homology.

In contrast, the ASSTs in this study, despite catalyzing the same reaction, exhibit lower homology, suggesting potential structural and functional variation. This is paralleled in human sulfotransferases (hSULTs). Humans have several hSULTs, which, although catalyzing identical chemical reactions, have divergent substrate specificities and functional diversifications. This diversity is partly due to variations in the sequences and conformations of loops near the active sites of these enzymes, which influence substrate specificity. Our predictions of functional divergence are based on sequence and structural alignments of *SwASSTs* and the understanding of the differences between ASSTs and human SULTs. Despite having structural distinctions and limited sequence similarity between these two classes of enzymes, the type of chemical reactions catalyzed by both are the same but with different donor and acceptor specificities. Therefore, it appears there has been an evolutionary branching in these two categories of sulfotransferases.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25052983/s1>.

**Author Contributions:** Conceptualization: D.D.S.; methodology, L.L. and D.D.S.; formal analysis, L.L. and D.D.S.; investigation, L.L. and D.D.S.; resources, D.D.S.; data curation, L.L.; writing—original draft preparation, L.L. and D.D.S.; writing—review and editing, D.D.S.; visualization, L.L.; supervision, D.D.S.; project administration, D.D.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Coughtrie, M.W.; Bamforth, K.J.; Sharp, S.; Jones, A.L.; Borthwick, E.B.; Barker, E.V.; Roberts, R.C.; Hume, R.; Burchell, A. Sulfation of endogenous compounds and xenobiotics—Interactions and function in health and disease. *Chem. Biol. Interact.* **1994**, *92*, 247–256. [[CrossRef](#)]
2. Idris, M.; Mitchell, D.J.; Gordon, R.; Sidharthan, N.P.; Butcher, N.J.; Minchin, R.F. Interaction of the Brain-Selective Sulfotransferase SULT4A1 with Other Cytosolic Sulfotransferases: Effects on Protein Expression and Function. *Drug Metab. Dispos.* **2020**, *48*, 337–344. [[CrossRef](#)]
3. Bairam, A.F.; Rasool, M.I.; Alherz, F.A.; Abunnaja, M.S.; El Daibani, A.A.; Gohal, S.A.; Alatwi, E.S.; Kurogi, K.; Liu, M. Impact of SULT1A3/SULT1A4 genetic polymorphisms on the sulfation of phenylephrine and salbutamol by human SULT1A3 allozymes. *Pharmacogenet. Genom.* **2019**, *29*, 99–105. [[CrossRef](#)]
4. Bairam, A.F.; Rasool, M.I.; Alherz, F.A.; Abunnaja, M.S.; El Daibani, A.A.; Kurogi, K.; Liu, M. Effects of human SULT1A3/SULT1A4 genetic polymorphisms on the sulfation of acetaminophen and opioid drugs by the cytosolic sulfotransferase SULT1A3. *Arch. Biochem. Biophys.* **2018**, *648*, 44–52. [[CrossRef](#)] [[PubMed](#)]
5. Kurogi, K.; Shimohira, T.; Kouriki-Nagatomo, H.; Zhang, G.; Miller, E.R.; Sakakibara, Y.; Suiko, M.; Liu, M.-C. Human Cytosolic Sulphotransferase SULT1C3: Genomic analysis and functional characterization of splice variant SULT1C3a and SULT1C3d. *J. Biochem.* **2017**, *162*, 403–414. [[CrossRef](#)]
6. Guidry, A.L.; Tibbs, Z.E.; Runge-Morris, M.; Falany, C.N. Expression, purification and characterization of human cytosolic sulfotransferase (SULT) 1C4. *Horm. Mol. Biol. Clin. Investig.* **2017**, *29*, 27–36. [[CrossRef](#)]
7. Lu, J.; Li, H.; Zhang, J.; Li, M.; An, X.; Liu, M.-C.; Chang, W. Crystal structures of SULT1A2 and SULT1A1 \*3: Insights into the substrate inhibition and the role of Tyr149 in SULT1A2. *Biochem. Biophys. Res. Commun.* **2010**, *396*, 429–434. [[CrossRef](#)]
8. Cook, I.T.; Duniec-Dmuchowski, Z.; Kocarek, T.A.; Runge-Morris, M.; Falany, C.N. 24-hydroxycholesterol sulfation by human cytosolic sulfotransferases: Formation of monosulfates and disulfates, molecular modeling, sulfatase sensitivity, and inhibition of liver x receptor activation. *Drug Metab. Dispos.* **2009**, *37*, 2069–2078. [[CrossRef](#)]
9. Lu, L.-Y.; Hsieh, Y.-C.; Liu, M.-Y.; Lin, Y.-H.; Chen, C.-J.; Yang, Y.-S. Identification and characterization of two amino acids critical for the substrate inhibition of human dehydroepiandrosterone sulfotransferase (SULT2A1). *Mol. Pharmacol.* **2008**, *73*, 660–668. [[CrossRef](#)]
10. Allali-Hassani, A.; Pan, P.W.; Dombrowski, L.; Najmanovich, R.; Tempel, W.; Dong, A.; Loppnau, P.; Martin, F.; Thonton, J.; Edwards, A.M.; et al. Structural and chemical profiling of the human cytosolic sulfotransferases. *PLoS Biol.* **2007**, *5*, e97.
11. Gamage, N.U.; Tsvetanov, S.; Duggleby, R.G.; McManus, M.E.; Martin, J.L. The structure of human SULT1A1 crystallized with estradiol. An insight into active site plasticity and substrate inhibition with multi-ring substrates. *J. Biol. Chem.* **2005**, *280*, 41482–41486. [[CrossRef](#)]
12. Pedersen, L.C.; Petrotchenko, E.; Shevtsov, S.; Negishi, M. Crystal structure of the human estrogen sulfotransferase-PAPS complex: Evidence for catalytic role of Ser137 in the sulfuryl transfer reaction. *J. Biol. Chem.* **2002**, *277*, 17928–17932. [[CrossRef](#)]
13. Fuda, H.; Lee, Y.C.; Shimizu, C.; Javitt, N.B.; Strott, C.A. Mutational analysis of human hydroxysteroid sulfotransferase SULT2B1 isoforms reveals that exon 1B of the SULT2B1 gene produces cholesterol sulfotransferase, whereas exon 1A yields pregnenolone sulfotransferase. *J. Biol. Chem.* **2002**, *277*, 36161–36166. [[CrossRef](#)]
14. Freimuth, R.R.; Raftogianis, R.B.; Wood, T.C.; Moon, E.; Kim, U.-J.; Xu, J.; Siciliano, M.J.; Weinshilboum, R.M. Human sulfotransferases SULT1C1 and SULT1C2: cDNA characterization, gene cloning, and chromosomal localization. *Genomics* **2000**, *65*, 157–165. [[CrossRef](#)]
15. Wang, J.; Falany, J.L.; Falany, C.N. Expression and characterization of a novel thyroid hormone-sulfating form of cytosolic sulfotransferase from human liver. *Mol. Pharmacol.* **1998**, *53*, 274–282. [[CrossRef](#)]

16. Sakakibara, Y.; Yanagisawa, K.; Katafuchi, J.; Ringer, D.P.; Takami, Y.; Nakayama, T.; Suiko, M.; Liu, M. Molecular cloning, expression, and characterization of novel human SULT1C sulfotransferases that catalyze the sulfonation of N-hydroxy-2-acetylaminofluorene. *J. Biol. Chem.* **1998**, *273*, 33929–33935. [[CrossRef](#)]
17. Fujita, K.-I.; Nagata, K.; Ozawa, S.; Sasano, H.; Yamazoe, Y. Molecular cloning and characterization of rat ST1B1 and human ST1B2 cDNAs, encoding thyroid hormone sulfotransferases. *J. Biochem.* **1997**, *122*, 1052–1061. [[CrossRef](#)]
18. Falany, C.N.; Krasnykh, V.; Falany, J.L. Bacterial expression and characterization of a cDNA for human liver estrogen sulfotransferase. *J. Steroid Biochem. Mol. Biol.* **1995**, *52*, 529–539. [[CrossRef](#)]
19. Radomska, A.; Comer, K.; Zimniak, P.; Falany, J.; Iscan, M.; Falany, C.N. Human liver steroid sulphotransferase sulphates bile acids. *Biochem. J.* **1990**, *272*, 597–604. [[CrossRef](#)]
20. Whittemore, R.M.; Pearce, L.B.; Roth, J.A. Purification and kinetic characterization of a phenol-sulfating form of phenol sulfotransferase from human brain. *Arch. Biochem. Biophys.* **1986**, *249*, 464–471. [[CrossRef](#)]
21. Le, H.H.; Lee, M.-T.; Besler, K.R.; Comrie, J.M.C.; Johnson, E.L. Characterization of interactions of dietary cholesterol with the murine and human gut microbiome. *Nat. Microbiol.* **2022**, *7*, 1390–1403. [[CrossRef](#)]
22. Yao, L.; D'agostino, G.D.; Park, J.; Hang, S.; Adhikari, A.A.; Zhang, Y.; Li, W.; Avila-Pacheco, J.; Bae, S.; Clish, C.B.; et al. A biosynthetic pathway for the selective sulfonation of steroidal metabolites by human gut bacteria. *Nat. Microbiol.* **2022**, *7*, 1404–1418. [[CrossRef](#)]
23. Kim, D.-H.; Kobashi, K. The role of intestinal flora in metabolism of phenolic sulfate esters. *Biochem. Pharmacol.* **1986**, *35*, 3507–3510. [[CrossRef](#)]
24. Kim, D.-H.; Konishi, L.; Kobashi, K. Purification, characterization and reaction mechanism of novel arylsulfotransferase obtained from an anaerobic bacterium of human intestine. *Biochim. Biophys. Acta* **1986**, *872*, 33–41. [[CrossRef](#)]
25. Kim, D.-H.; Yoon, H.-K.; Koizumi, M.; Kobashi, K. Sulfonation of phenolic antibiotics by sulfotransferase obtained from a human intestinal bacterium. *Chem. Pharm. Bull.* **1992**, *40*, 1056–1057. [[CrossRef](#)]
26. Kobashi, K.; Fukaya, Y.; Kim, D.-H.; Akao, T.; Takebe, S. A novel type of aryl sulfotransferase obtained from an anaerobic bacterium of human intestine. *Arch. Biochem. Biophys.* **1986**, *245*, 537–539. [[CrossRef](#)]
27. Suiko, M.; Kurogi, K.; Hashiguchi, T.; Sakakibara, Y.; Liu, M.-C. Updated perspectives on the cytosolic sulfotransferases (SULTs) and SULT-mediated sulfation. *Biosci. Biotechnol. Biochem.* **2017**, *81*, 63–72. [[CrossRef](#)]
28. Tibbs, Z.E.; Rohn-Glowacki, K.J.; Crittenden, F.; Guidry, A.L.; Falany, C.N. Structural plasticity in the human cytosolic sulfotransferase dimer and its role in substrate selectivity and catalysis. *Drug Metab. Pharmacokinet.* **2015**, *30*, 3–20. [[CrossRef](#)]
29. Malojčić, G.; Owen, R.L.; Grimshaw, J.P.A.; Brozzo, M.S.; Dreher-Teo, H.; Glockshuber, R. A structural and biochemical basis for PAPS-independent sulfuryl transfer by aryl sulfotransferase from uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19217–19222. [[CrossRef](#)]
30. Gamage, N.; Barnett, A.; Hempel, N.; Duggleby, R.G.; Windmill, K.F.; Martin, J.L.; McManus, M.E. Human sulfotransferases and their role in chemical metabolism. *Toxicol. Sci.* **2006**, *90*, 5–22. [[CrossRef](#)]
31. Zhang, H.; Varlamova, O.; Vargas, F.M.; Falany, C.N.; Leyh, T.S. Sulfuryl transfer: The catalytic mechanism of human estrogen sulfotransferase. *J. Biol. Chem.* **1998**, *273*, 10888–10892. [[CrossRef](#)]
32. Konishi-Imamura, L.; Dohi, K.; Sato, M.; Kobashi, K. Improved purification of arylsulfate sulfotransferase from human intestinal bacterium by using polyclonal antibody. *J. Biochem.* **1994**, *115*, 1097–1100. [[CrossRef](#)]
33. Baek, M.-C.; Kwon, A.-R.; Chung, Y.-J.; Kim, B.-K.; Choi, E.-C. Distribution of bacteria with the arylsulfate sulfotransferase activity. *Arch. Pharm. Res.* **1998**, *21*, 475–477. [[CrossRef](#)]
34. Koppel, N.; Rekdal, V.M.; Balskus, E.P. Chemical transformation of xenobiotics by the human gut microbiota. *Science* **2017**, *356*, 6344. [[CrossRef](#)]
35. Wexler, H.M.; Reeves, D.; Summanen, P.H.; Molitoris, E.; McTEAGUE, M.; Duncan, J.; Wilson, K.H.; Finegold, S.M. *Sutterella wadsworthensis* gen. nov., sp. nov., bile-resistant microaerophilic *Campylobacter gracilis*-like clinical isolates. *Int. J. Syst. Bacteriol.* **1996**, *46*, 252–258. [[CrossRef](#)]
36. Engberg, J.; On, S.L.W.; Harrington, C.S.; Gerner-Smidt, P. Prevalence of *Campylobacter*, *Arcobacter*, *Helicobacter*, and *Sutterella* spp. in human fecal samples as estimated by a reevaluation of isolation methods for *Campylobacters*. *J. Clin. Microbiol.* **2000**, *38*, 286–291. [[CrossRef](#)]
37. Sakon, H.; Nagai, F.; Morotomi, M.; Tanaka, R. *Sutterella parvirubra* sp. nov. and *Megamonas funiformis* sp. nov., isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **2008**, *58 Pt 4*, 970–975. [[CrossRef](#)]
38. Ndongo, S.; Cadoret, F.; Dubourg, G.; Delerce, J.; Fournier, P.-E.; Raoult, D.; Lagier, J.-C. '*Collinsella phocaensis*' sp. nov., '*Clostridium merdae*' sp. nov., '*Sutterella massiliensis*' sp. nov., '*Suttarella timonensis*' sp. nov., '*Enorma phocaensis*' sp. nov., '*Mailhella massiliensis*' gen. nov., sp. nov., '*Mordavella massiliensis*' gen. nov., sp. nov. and '*Massiliprevotella massiliensis*' gen. nov., sp. nov., 9 new species isolated from fresh stool samples of healthy French patients. *New Microbes New Infect.* **2017**, *17*, 89–95.
39. Dione, N.; Ngom, I.I.; Valles, C.; Cadoret, F.; Fournier, P.E.; Raoult, D.; Lagier, J.C. '*Collinsella provencensis*' sp. nov., '*Parabacteroides bouchesdurhonensis*' sp. nov. and '*Sutterella seckii*' sp. nov., three new bacterial species identified from human gut microbiota. *New Microbes New Infect.* **2018**, *23*, 44–47. [[CrossRef](#)]
40. Sakamoto, M.; Ikeyama, N.; Kunihiro, T.; Iino, T.; Yuki, M.; Ohkuma, M. *Mesosutterella multiformis* gen. nov., sp. nov., a member of the family Sutterellaceae and *Sutterella megalosphaeroides* sp. nov., isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 3942–3950. [[CrossRef](#)] [[PubMed](#)]

41. Oh, B.S.; Kim, J.S.; Yu, S.Y.; Ryu, S.W.; Park, S.H.; Kang, S.W.; Park, J.-E.; Choi, S.-H.; Han, K.-I.; Lee, K.C.; et al. *Sutterella faecalis* sp. nov., isolated from human faeces. *J. Microbiol.* **2020**, *58*, 99–104. [[CrossRef](#)]
42. Kaakoush, N.O. *Sutterella* Species, IgA-degrading Bacteria in Ulcerative Colitis. *Trends Microbiol.* **2020**, *28*, 519–522. [[CrossRef](#)]
43. Williams, B.L.; Hornig, M.; Parekh, T.; Lipkin, W.I. Application of novel PCR-based methods for detection, quantitation, and phylogenetic characterization of *Sutterella* species in intestinal biopsy samples from children with autism and gastrointestinal disturbances. *mBio* **2012**, *3*, e00261-11. [[CrossRef](#)]
44. Wang, L.; Christophersen, C.T.; Sorich, M.J.; Gerber, J.P.; Angley, M.T.; Conlon, M.A. Increased abundance of *Sutterella* spp. and *Ruminococcus torques* in feces of children with autism spectrum disorder. *Mol. Autism.* **2013**, *4*, 42. [[CrossRef](#)]
45. Kirk, K.F.; Andersen, K.L.; Tarpgaard, I.H.; Nielsen, H.L. Three cases of *Sutterella wadsworthensis* bacteremia secondary to abdominal infections. *Anaerobe* **2021**, *72*, 102460. [[CrossRef](#)]
46. Giri, S.; Mangalam, A. Chapter 34—The Gut Microbiome and Metabolome in Multiple Sclerosis. In *Microbiome and Metabolome in Diagnosis, Therapy, and other Strategic Applications*; Faintuch, J., Faintuch, S., Eds.; Academic Press: Cambridge, MA, USA, 2019; pp. 333–340.
47. Cornejo-Pareja, I.; Ruiz-Limón, P.; Gómez-Pérez, A.M.; Molina-Vega, M.; Moreno-Indias, I.; Tinahones, F.J. Differential Microbial Pattern Description in Subjects with Autoimmune-Based Thyroid Diseases: A Pilot Study. *J. Pers. Med.* **2020**, *10*, 192. [[CrossRef](#)]
48. Hiippala, K.; Kainulainen, V.; Kalliomäki, M.; Arkkila, P.; Satokari, R. Mucosal Prevalence and Interactions with the Epithelium Indicate Commensalism of *Sutterella* spp. *Front. Microbiol.* **2016**, *7*, 1706. [[CrossRef](#)]
49. Raghavan, R.; Kelkar, Y.D.; Ochman, H. A selective force favoring increased G+C content in bacterial genes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14504–14507. [[CrossRef](#)]
50. Schell, M.A. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **1993**, *47*, 597–626. [[CrossRef](#)]
51. van der Ploeg, J.R.; Eichhorn, E.; Leisinger, T. Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Arch. Microbiol.* **2001**, *176*, 1–8. [[CrossRef](#)]
52. Kouzuma, A.; Endoh, T.; Omori, T.; Nojiri, H.; Yamane, H.; Habe, H. Transcription factors CysB and SfnR constitute the hierarchical regulatory system for the sulfate starvation response in *Pseudomonas putida*. *J. Bacteriol.* **2008**, *190*, 4521–4531. [[CrossRef](#)]
53. Pokhrel, A.; Dinh, H.; Li, L.; Hassan, K.A.; Cain, A.K.; Paulsen, I.T. Identification of a novel LysR family transcriptional regulator controlling acquisition of sulfur sources in *Acinetobacter baumannii*. *Microb. Physiol.* **2023**, *33*, 27–35. [[CrossRef](#)]
54. Wang, J.P.; Zhang, W.M.; Chao, H.J.; Zhou, N.Y. PnpM, a LysR-Type Transcriptional Regulator Activates the Hydroquinone Pathway in para-Nitrophenol Degradation in *Pseudomonas* sp. Strain WBC-3. *Front. Microbiol.* **2017**, *8*, 1714. [[CrossRef](#)]
55. Lu, Z.; Imlay, J.A. The Fumarate Reductase of *Bacteroides thetaiotaomicron*, unlike That of *Escherichia coli*, Is Configured so that It Does Not Generate Reactive Oxygen Species. *mBio* **2017**, *8*, e01873-16. [[CrossRef](#)]
56. Laue, H.; Denger, K.; Cook, A.M. Taurine reduction in anaerobic respiration of *Bilophila wadsworthia* RZATAU. *Appl. Environ. Microbiol.* **1997**, *63*, 2016–2021. [[CrossRef](#)]
57. Häse, C.C.; A Finkelstein, R. Bacterial extracellular zinc-containing metalloproteases. *Microbiol. Rev.* **1993**, *57*, 823–837. [[CrossRef](#)]
58. Garmory, H.S.; Titball, R.W. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infect. Immun.* **2004**, *72*, 6757–6763. [[CrossRef](#)] [[PubMed](#)]
59. Kertesz, M.A. Bacterial transporters for sulfate and organosulfur compounds. *Res. Microbiol.* **2001**, *152*, 279–290. [[CrossRef](#)]
60. Chen, J.; Xie, J. Role and regulation of bacterial LuxR-like regulators. *J. Cell Biochem.* **2011**, *112*, 2694–2702. [[CrossRef](#)] [[PubMed](#)]
61. Dong, X.; Guthrie, B.G.H.; Alexander, M.; Noecker, C.; Ramirez, L.; Glasser, N.R.; Turnbaugh, P.J.; Balskus, E.P. Genetic manipulation of the human gut bacterium *eggerthella lenta* reveals a widespread family of transcriptional regulators. *Nat. Commun.* **2022**, *13*, 7624. [[CrossRef](#)]
62. Ryan, J.; McKillen, M.; Mason, J. Sulphate/molybdate interactions: In vivo and in vitro studies on the group VI oxyanion transport system in ovine renal tubule epithelial cells. *Ann. Rech. Vet.* **1987**, *18*, 47–55.
63. Self, W.T.; Grunden, A.M.; Hasona, A.; Shanmugam, K.T. Molybdate transport. *Res. Microbiol.* **2001**, *152*, 311–321. [[CrossRef](#)]
64. Mason, J.; Cardin, C. The competition of molybdate and sulphate ions for a transport system in the ovine small intestine. *Res. Vet. Sci.* **1977**, *22*, 313–315. [[CrossRef](#)]
65. Toth, E.A.; Yeates, T.O. The structure of adenylosuccinate lyase, an enzyme with dual activity in the de novo purine biosynthetic pathway. *Structure* **2000**, *8*, 163–174. [[CrossRef](#)]
66. Pruss, K.M.; Enam, F.; Battaglioli, E.; DeFeo, M.; Diaz, O.R.; Higginbottom, S.K.; Fischer, C.R.; Hryckowian, A.J.; Van Treuren, W.; Dodd, D.; et al. Oxidative ornithine metabolism supports non-inflammatory *C. difficile* colonization. *Nat. Metab.* **2022**, *4*, 19–28. [[CrossRef](#)]
67. Rawlings, N.D.; Barrett, A.J.; Thomas, P.D.; Huang, X.; Bateman, A.; Finn, R.D. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **2018**, *46*, D624–D632. [[CrossRef](#)]
68. Amedei, A.; Capasso, C.; Nannini, G.; Supuran, C.T. Microbiota, Bacterial Carbonic Anhydrases, and Modulators of Their Activity: Links to Human Diseases? *Mediators Inflamm.* **2021**, *2021*, 6926082. [[CrossRef](#)]
69. Grönlund, M.-M.; Lehtonen, O.-P.; Eerola, E.; Kero, P. Fecal microflora in healthy infants born by different methods of delivery: Permanent changes in intestinal flora after cesarean delivery. *J. Pediatr. Gastroenterol. Nutr.* **1999**, *28*, 19–25.

70. Hall, B.M.; Breidenstein, E.B.M.; de la Fuente-Núñez, C.; Reffuveille, F.; Mawla, G.D.; Hancock, R.E.W.; Baker, T.A. Two Isoforms of Clp Peptidase in *Pseudomonas aeruginosa* Control Distinct Aspects of Cellular Physiology. *J. Bacteriol.* **2017**, *199*, e00568-16. [[CrossRef](#)]
71. Bolam, D.N.; van den Berg, B. TonB-dependent transport by the gut microbiota: Novel aspects of an old problem. *Curr. Opin. Struct. Biol.* **2018**, *51*, 35–43. [[CrossRef](#)]
72. Pollet, R.M.; Martin, L.M.; Koropatkin, N.M. TonB-dependent transporters in the Bacteroidetes: Unique domain structures and potential functions. *Mol. Microbiol.* **2021**, *115*, 490–501. [[CrossRef](#)]
73. Doerfel, L.K.; Wohlgemuth, I.; Kothe, C.; Peske, F.; Urlaub, H.; Rodnina, M.V. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* **2013**, *339*, 85–88. [[CrossRef](#)]
74. Kang, T.M.; Yuan, J.; Zhou, A.; Beppler, C.; Miller, J.H. Deoxycytidine deaminase-deficient *Escherichia coli* strains display acute sensitivity to cytidine, adenosine, and guanosine and increased sensitivity to a range of antibiotics, including vancomycin. *J. Bacteriol.* **2014**, *196*, 1950–1957. [[CrossRef](#)]
75. Sugrue, E.; Fraser, N.J.; Hopkins, D.H.; Carr, P.D.; Khurana, J.L.; Oakeshott, J.G.; Scott, C.; Jackson, C.J. Evolutionary expansion of the amidohydrolase superfamily in bacteria in response to the synthetic compounds molinate and diuron. *Appl. Environ. Microbiol.* **2015**, *81*, 2612–2624. [[CrossRef](#)]
76. Erb, T.J. Carboxylases in natural and synthetic microbial pathways. *Appl. Environ. Microbiol.* **2011**, *77*, 8466–8477. [[CrossRef](#)]
77. Monti, S.M.; De Simone, G.; D’Ambrosio, K. L-Histidinol Dehydrogenase as a New Target for Old Diseases. *Curr. Top. Med. Chem.* **2016**, *16*, 2369–2378. [[CrossRef](#)] [[PubMed](#)]
78. Chioccioli, S.; Del Duca, S.; Vassallo, A.; Castronovo, L.M.; Fani, R. Exploring the role of the histidine biosynthetic hisF gene in cellular metabolism and in the evolution of (ancestral) genes: From LUCA to the extant (micro)organisms. *Microbiol. Res.* **2020**, *240*, 126555. [[CrossRef](#)] [[PubMed](#)]
79. Lewinson, O.; Lee, A.T.; Rees, D.C. A P-type ATPase importer that discriminates between essential and toxic transition metals. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4677–4682. [[CrossRef](#)]
80. Pealing, S.L.; Black, A.C.; Manson, F.D.C.; Ward, F.B.; Chapman, S.K.; Reid, G.A. Sequence of the gene encoding flavocytochrome c from *Shewanella putrefaciens*: A tetraheme flavoenzyme that is a soluble fumarate reductase related to the membrane-bound enzymes from other bacteria. *Biochemistry* **1992**, *31*, 12132–12140. [[CrossRef](#)]
81. He, S.H.; DerVartanian, D.V.; LeGall, J. Isolation of fumarate reductase from *Desulfovibrio multispicans*, a sulfate reducing bacterium. *Biochem. Biophys. Res. Commun.* **1986**, *135*, 1000–1007. [[CrossRef](#)]
82. Natale, P.; Brüser, T.; Driessen, A.J. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—Distinct translocases and mechanisms. *Biochim. Biophys. Acta* **2008**, *1778*, 1735–1756. [[CrossRef](#)]
83. Hinsley, A.P.; Stanley, N.R.; Palmer, T.; Berks, B.C. A naturally occurring bacterial Tat signal peptide lacking one of the ‘invariant’ arginine residues of the consensus targeting motif. *FEBS Lett.* **2001**, *497*, 45–49. [[CrossRef](#)]
84. Molik, S.; Karnauchov, I.; Weidlich, C.; Herrmann, R.G.; Klosgen, R.B. The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts: Missing link in the evolution of protein transport pathways in chloroplasts? *J. Biol. Chem.* **2001**, *276*, 42761–42766. [[CrossRef](#)]
85. Frain, K.M.; Robinson, C.; van Dijl, J.M. Transport of Folded Proteins by the Tat System. *Protein J.* **2019**, *38*, 377–388. [[CrossRef](#)]
86. Green, E.R.; Meccas, J. Bacterial Secretion Systems: An Overview. *Microbiol. Spectr.* **2016**, *4*, 1. [[CrossRef](#)]
87. Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. IDDT: A local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **2013**, *29*, 2722–2728. [[CrossRef](#)]
88. Carugo, O.; Pongor, S. A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci.* **2001**, *10*, 1470–1473. [[CrossRef](#)]
89. King, C.H.; Desai, H.; Sylvestsky, A.C.; LoTempio, J.; Ayanyan, S.; Carrie, J.; Crandall, K.A.; Fochtman, B.C.; Gasparyan, L.; Gulzar, N.; et al. Baseline human gut microbiota profile in healthy people and standard reporting template. *PLoS ONE* **2019**, *14*, e0206484. [[CrossRef](#)]
90. Saito, Y.; Sato, T.; Nomoto, K.; Tsuji, H. Identification of phenol- and p-cresol-producing intestinal bacteria by using media supplemented with tyrosine and its metabolites. *FEMS Microbiol. Ecol.* **2018**, *94*, fiy125. [[CrossRef](#)] [[PubMed](#)]
91. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
92. Mirdita, M.; Schütze, K.; Moriawaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making protein folding accessible to all. *Nat. Methods* **2022**, *19*, 679–682. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.