# Results V1.0

## Orthogroups in order of gene overlap (>2)

OG000529 has 4 overlaps with 1 copy present in all Ac. And represents flgE flagellar hook protein.

OG0000246 and OG0000968 both with 3 overlaps, representing biofilm formation protein PslI and PslD respectively from Pseudomonas aeruginosa. PslI is present in all but Ac. ferrooxidans, with a second copy present in Ac. prosperus. PslD is present only in Ac. ferrooxidans and Ac. yilgarnensis.

OG0000331 and OG0000562, both with 3 overlaps and single copies present in all Ac. species. These orthogroups represent two enzymes part of the de novo purine biosynthesis pathway.

OG0000622, OG0000902, and OG0000905 with 3 overlaps representing the flagellar motor switch, flagella capping protein, and flagellar m-ring protein respectively are also present with single copies in all Ac. species.

OG0000576 with 3 overlaps is also present in all Ac. species. It represents haloacid dehalogenase.

## Interesting by heatmap

### Significant expansion

OG0000015 with no overlap representing NADP(H)-dependent aldo-keto reductase and present in all Ac. species with 4 in Ac. aeolianus, 2 in Ac. ferrooxidans, 6 in Ac. prosperus, and 5 in Ac. yilgarnensis.

OG0000020 with 2 overlaps representing lipopolysaccharide glucosyltransferase I. Present in all Ac. species with 6 in species aeolianus, ferrooxidans, and yilgarnensis and 3 in Ac. prosperus.

OG0000023 with no overlap representing type IV pilus twitching motility protein. Present in all Ac. species with 7 in Ac. species ferrooxidans, prosperus, and yilgarnensis and 3 in aeolianus.

OG0000051 no overlap, represents MotA/TolQ/ExbB proton channel family protein. In all Ac. species, 4 in aeolianus and prosperus, 3 in ferrooxidans and 5 in yilgarnensis.

### Moderate expansion

OG0000058 no overlap, represents dihydrolipoyl dehydrogenase. In all Ac. species, 4 in aeolianus and ferrooxidans, 2 in prosperus, 3 in yilgarnensis.

OG0000066 no overlap, represents branched-chain amino acid aminotransferase. In all Ac. species, 3 in aeolianus and prosperus, 4 in ferrooxidans, 2 in yilgarnensis.

OG0000079 no overlap, represents amidase. In all Ac. species, 4 in aeolianus, 3 in ferrooxidans and prosperus, 1 in yilgarnensis.

OG0000187 no overlap, represents UDP-glucose 6-dehydrogenase. In all Ac. species, 3 in aeolianus, prosperus and yilgarnensis, and 1 in ferrooxidans.

OG0000407 no overlaps, represents diacylglycerol kinase. In all Ac. species, 2 in all.

OG0000596 no overlaps, represents mercury(II) reductase. In all Ac. species, 2 in all.

### Other potentially interesting things?? (are there other ways of determining interesting things?)

OG0000324 no overlap. TRZ/ATZ family hydrolase. One in all and 4 in ferrooxidans.

OG0000435 1 overlap. glucose-6-phosphate isomerase. One in all, 3 in yilgarnensis.

Note - Ac. species completely lack fimbriae according to E\_coli orthofinder results. However, they do appear to have orthology with WP\_044570658.1 (OG0001130 no overlaps) fimV fimbrial protein in Stenotrophomonas with 1 in all Ac. species.

### Missing

There are 200 orthogroups with relation to the candidate genes where Acidihalobacter has 0 copies.

# Discussion

## Motility

### Flagellar motility

OG000529, OG0000622, OG0000902, and OG0000905

(<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8940932/>) shows that the inactivation of flgE in Pseudomonas aeruginosa results in the promotion of aggregated structures and a decreased expression in biofilm matrix genes. With a lack of flgE, the organism had a reduced ability to adhere to surfaces, although once adhered, demonstrated faster biofilm growth.

The flagellar may act as a switch/sensor, indicating to the organism that it has found a surface. (<https://pubmed.ncbi.nlm.nih.gov/23927648/>) indicates that surface mediated restriction of flagellar rotation may allow the organism to sense/detect the surface, triggering the activation of biofilm related pathways and subsequent restriction of flagellar expression in further proliferation.

### Pili motility

OG0000023

Twitching motility is a flagella independent method of motility on a solid surface and appears to be mostly but not completely restricted to proteobacteria (beta, **gamma,** and delta)(<https://doi.org/10.1146/annurev.micro.56.012302.160938>).

While it is widely accepted that type IV pili mediates adherence, a lack of pili does not necessarily prevent surface adherence, although it does inhibit the formation of a complex/multicell layered biofilm (<https://doi.org/10.1046/j.1365-2958.1998.01062.x>).

## Extracellular Polymeric Substances (EPS)

### Pseudomonas Psl Exopolysaccharide

OG0000246 and OG0000968

Polysaccharide locus (Psl) is one of the three major exopolysaccharides in Pseudomonas biofilm (<https://dx.doi.org/10.1016%2Fj.mib.2007.09.010>).

(the two other pathways for biofilm formation in Pseudo is Pel and alginate, ctrl+F in genes table finds some Pel genes in some Ac. species, but can’t find anything with alginate, worth adding alginate candidate genes to look for?)

Psl is a 15-gene operon that encodes for the Psl exopolysaccharide and is important in the role of adhesion, biofilm maintenance, and cell-cell interaction. *“Although Psl was proposed to promote both cell-cell and cell-surface interactions, little is known about how Psl is anchored on the bacterial cell surface and how it might promote these interactions.”* (<https://doi.org/10.1128/jb.01202-06>).

## ?Metabolic

OG0000331 and OG0000562

De novo purine biosynthesis pathway that has been demonstrate to influence biofilm production. (<https://pubmed.ncbi.nlm.nih.gov/26788425/>)

OG0000576

Haloacid dehalogenase - *“In plant pathogen bacteria, HADs might also be involved in biofilm formation during biotic interactions”*

(<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8490876/>)

OG0000015

NADP(H)-dependent aldo-keto reductase (not really sure what this does, waste/oxidative stress maybe?)

OG0000020

Lipopolysaccharide glucosyltransferase I. Important for attachment (<https://pubmed.ncbi.nlm.nih.gov/10398745/>).

OG0000051

MotA/TolQ/ExbB proton channel family protein

OG0000058

Dihydrolipoyl dehydrogenase.

OG0000066

Branched-chain amino acid aminotransferase.

OG0000079

Amidase

OG0000187

UDP-glucose 6-dehydrogenase

OG0000407

Diacylglycerol kinase, identified as important to the formation of biofilm in Strep (?gram positive) (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3575517/>).

OG0000596

Mercury(II) reductase (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC92343/>).

OG0000324

TRZ/ATZ family hydrolase

OG0000435

Glucose-6-phosphate isomerase (<https://journals.asm.org/doi/10.1128/mSphere.01231-20>).