Annotation of the *Kytococcus sedentarius* Genome From DNA Coordinates 1137484 to 1140893

Erika Wicher

Biomedical Sciences Major

SEPA SCIENCE EDUCATION PARTNERSHIP AWARD
Supported by the National Institutes of Health

Abstract

Kytococcus sedentarius is a free-living, non-motile, Gram-positive bacteria that is strictly aerobic. It has a spherical/coccoid shape that occurs in tetrads. It has the ability to produce oligoketide antibiotics and is an opportunistic pathogen that can cause endocarditis, pitted keratolysis, and hemorrhagic pneumonia (Sims et al., 2009). The whole genome of *K. sedentarius* has been sequenced, but not all genes have been manually confirmed. The goal of this project is to fill this gap in knowledge using the modules in GENI-ACT (https://geni-act.org/), which include looking at sequence based similarity, structure based similarity, cellular localization, enzymatic function, and other modules. All of my proposed annotations coincide with the computer, except for KSED_05440, which I predict to be YraN. I found that three of these genes, KSED_RS05440, KSED_RS05445, and KSED_RS05450, are involved in homologous recombination of exogenous DNA. I am proposing that these three genes are located on the same operon.

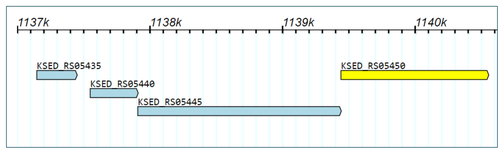

Figure 1 Genes of Interest

Results

KSED_RS05450

- BLAST, COG, TIGRFAM, PFAM, PDB all showed hits of DprA with high scores and low E-values.
- WebLogo and T-Coffee showed well conserved alignments.
- DprA is involved in the uptake of exogenous chromosomal ssDNA in natural transformation (Mortier-Barriere et all, 2007).
- DprA binds to the ssDNA to prevent degradation before it can bind to RecA, which aids in the exchanging of homologous DNA strands (Lesetti et al. 2002).

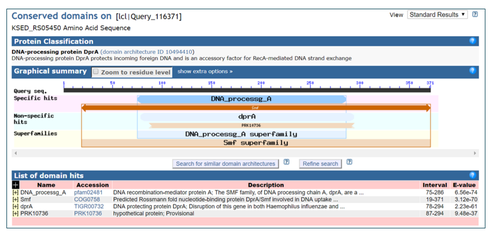

Figure 2 CDD results of KSED_RS05450

Computer Annotation: DNA protecting protein, DprA
Proposed Annotation: Agree with computer, DprA

KSED_RS05440

- BLAST, COG and PFAM all showed hits of YraN or UPF0102 with high scores and low E-values.
- WebLogo and T-Coffee showed well conserved alignments.
- YraN has predicted endonuclease activity and nucleic acid binding in Holliday junction resolution.

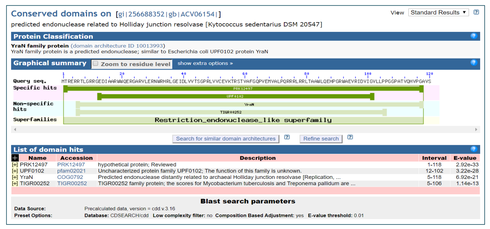

Figure 3 CDD results of KSED_RS05440

- TMHMM and SignalP showed negative results of transmembrane helices and signal peptides.
- PSORT-B predicted this protein to be in the cytoplasm.

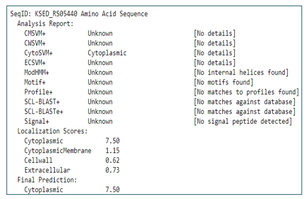

Figure 4 PSORT-B results for KSED_RS05440

Computer Annotation: Hypothetical protein
Proposed Annotation: YraN; predicted endonuclease activity involved in Holliday junction resolution

KSED_RS05445

- COG, PFAM, and TIGRFAM all gave hits of Magnesium chelatase with high scores and low E-values.
- SP BLAST gave a competence protein, which allows the cell to undergo transformation by uptake of exogenous DNA (Chen et al. 2001).
- Has enzymatic function with the EC number of 6.6.1.1, which can be involved in chlorophyll biosynthesis.
- Magnesium chelatase catalyzes the insertion of a magnesium ion into a protoporphyrin, which involves ATP. *K. sedentarius* cannot undergo photosynthesis because it is strictly aerobic.
- It has been researched that magnesium and ATP is needed in the early steps of homologous recombination (Register et al., 1985).
- Well conserved in WebLogo and T-Coffee, with no evidence of horizontal gene transfer.
- PSORT-B predicted the localization to be in the cytoplasm.

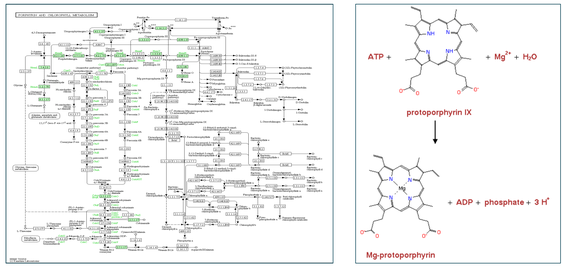

Figure 5 KEGG pathway (left) and MetaCyc reaction of Magnesium Chelatase (right)

Computer Annotation: Magnesium chelatase-related protein
Proposed Annotation: Agree with computer, magnesium chelatase-related

Prediction of Operon

- All functionally-related proteins found in the cytoplasm.
- Found consecutively in the genome.
- Gene neighborhood showed that is was conserved in other organisms.

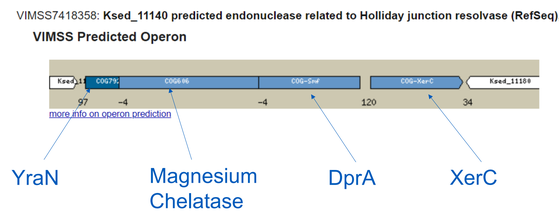

Figure 6 Operon Prediction from MicrobesOnline

- All proteins are involved in the process of transformation, which is a type of horizontal gene transfer.
- DprA protecting the incoming DNA from degradation.
- Magnesium chelatase needed in the early steps of homologous recombination.
- YraN cutting the DNA to which the exogenous DNA will be integrated into.
- XerC is tyrosine recombinase, which is involved in strand exchange in homologous recombination (Blakely et al. 1994).

KSED_RS05435

- Significant hits for domain of unknown function family in BLAST and PFAM.
- No significant hits for COG and TIGRFAM.
- Very well conserved in WebLogo.

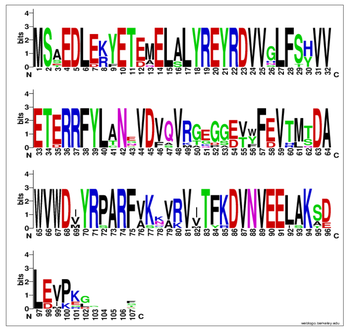

Figure 7 WebLogo for KSED_RS05435

Computer Annotation: Hypothetical protein
Proposed Annotation: Agree with computer

Conclusion

Locus Tag	Computer Prediction	Agree with Computer?
KSED_RS05435	Hypothetical protein	Yes
KSED_RS05440	Hypothetical protein	No, YraN protein with predicted endonuclease activity
KSED_RS05445	Magnesium chelatase-related protein	Yes
KSED_RS05450	DNA protecting protein, DprA	Yes

- Prediction of KSED_RS05440, KSED_RS05445, and KSED_RS05450 to be on the same operon, along with XerC (KSED_RS05455). All involving homologous recombination.

References

1. Sims D, Brettin T, Detter, J, Han C, Lapidus A, Copeland A, Del Rio T, Nolan M, Chen F, Lucas S, et al.2009. Standard Genomic Sci. 1(1):12-20.
2. Mortier-Barriere I, Velton M, Dupaigne P, Le Cam E, Polard P, Claverys J. 2007. Cell. 130(5):824-836.
3. Lesetti S, Cox M. 2002. Annual Review of Biochemistry. 71:71-100.
4. Chen I, Gotschlich EC. 2001. J Bacteriol. 183 (10):3160-3168.
5. Li X, Heyer W. 2008. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. 18 (1):99-113.
6. Register JC, Griffith J. 1985. Mol Gen Genet. 199(3):415-20.
7. Blakely GW, Sherratt DJ. 1994. Nucleic Acid Res. 22(25): 5613-5620.

University at Buffalo The State University of New York

Jacobs School of Medicine and Biomedical Sciences
buffalo.edu

Acknowledgments

Dr. Rama Dey-Rao, Dr. Stephen Koury
Supported by an NIH Science Education Partnership (SEPA) Award - R25ODO10536