

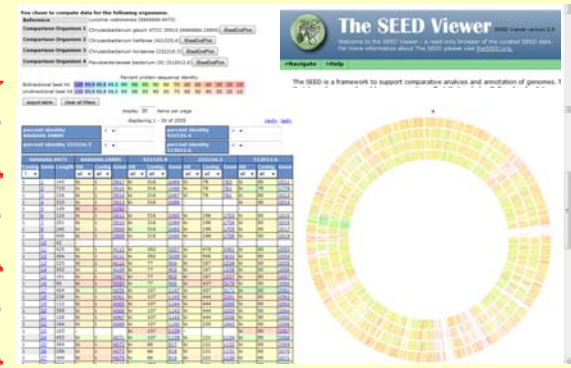
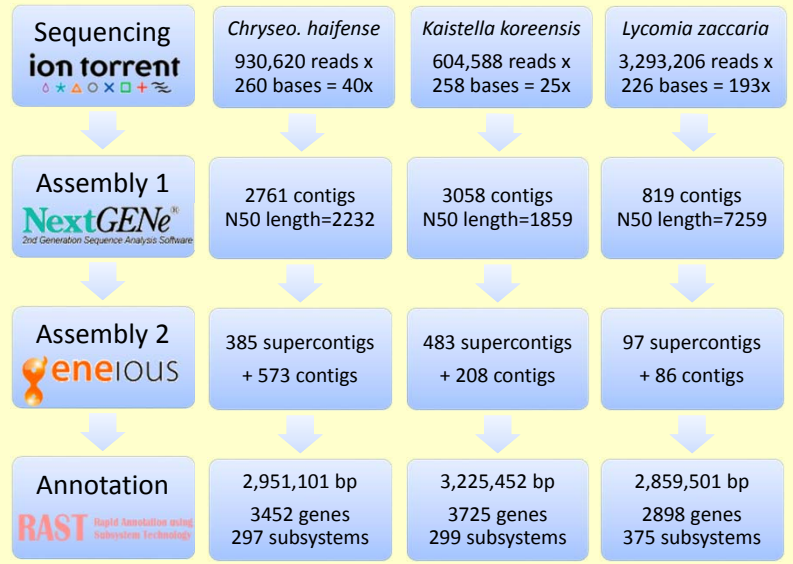
Assembly and Annotation Strategies for Three Microbial Genomes Sequenced Through the Genome Consortium for Active Teaching – NextGen Sequencing Group (GCAT-SEEK). Jeff Newman, Timothy Hosteley, Thomas Sontag, Jordan Krebs. Lycoming College .



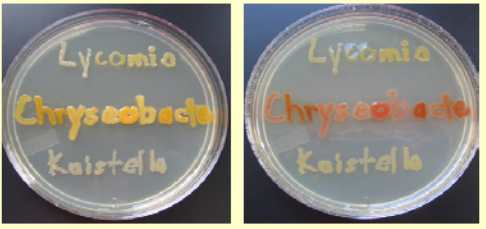
GCAT-SEEKquence is a Research Collaboration Network Initiated by Mike Boyle and Led by Vince Buonaccorsi at Juniata College.

Mission

1. Bring functional genomic methods into undergraduate curriculum primarily through independent and classroom based student research using centralized cores to make NextGen sequence data accessible to undergraduates.
2. Create a Clearing House of Information for educators to use when teaching NextGen sequencing and related topics.
3. Create large database of raw data and analyzed results for pedagogical use by GCAT-SEEK members
4. Develop a Global Network of educators using functional genomics and NextGen sequencing in the undergraduate curriculum



1. Sequence based comparison with related genomes suggests adjacent contigs/ supercontigs.
2. Contigs/supercontigs can be reordered & reverse complemented in Artemis
3. Detailed sequence/ORF analysis can determine if contigs/ supercontigs overlap, or if gap exists to be closed by PCR and sequencing during finishing.
4. Comparison of genes allows identification of species specific genes, genes shared by different groups



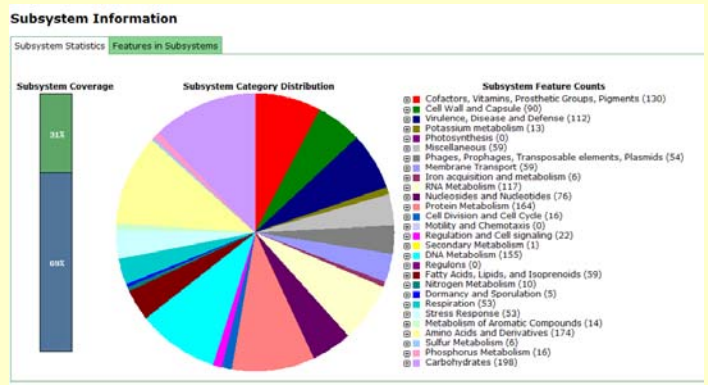
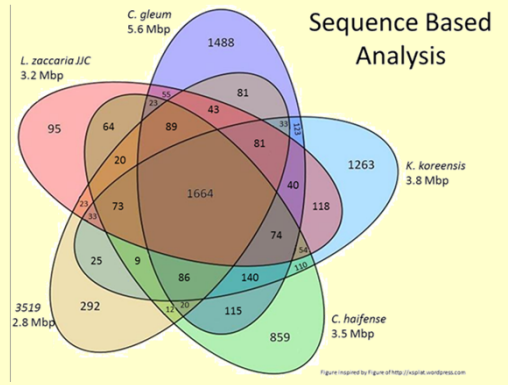
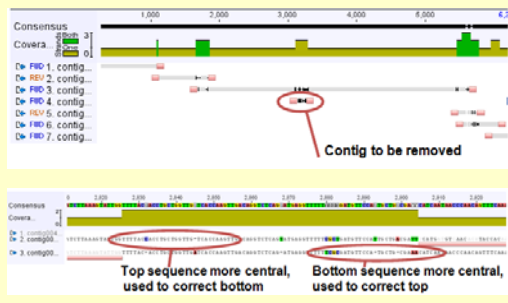
TSA for 24 hrs @ 36°C

- C = *Chryseobacterium angustatum* DSM 21467
- h = *Chryseobacterium angustum* NCIC 21467
- r = *Chryseobacterium proteoferrireducens* UMI 1047
- y = *Chryseobacterium jeikeium* DSM 25899
- o = *Chryseobacterium jeikeium* DSM 18222
- e = *Chryseobacterium jeikeium* DSM 18222
- l = *Chryseobacterium jeikeium* DSM 18222
- b = *Chryseobacterium jeikeium* DSM 24047
- g = *Chryseobacterium jeikeium* DSM 17227
- c = *Chryseobacterium jeikeium* DSM 17227
- t = *Chryseobacterium jeikeium* DSM 18222
- a = *Chryseobacterium jeikeium* DSM 22227

After spraying with 40% KOH solution red color = + for flexirubin

- ly = *Lycomia haifense* DSM 19056
- ca = *Lycomia zaccaria* DSM 20257
- ma = *Lycomia zaccaria* DSM 20257
- ka = *Kaistella kaistella* DSM 20257

1. Trimming low coverage ends of NextGen contigs allows further assembly, into "supercontigs"
2. Requires manual editing of ambiguous bases
3. Supercontigs and unused contigs combined, uploaded to RAST for annotation



- Our Projects** – de novo sequencing and assembly of microbial genomes
- *Flavobacteriaceae* sp. JJC – Isolated from Ice creek, Possible novel genus to be named *Lycomia zaccaria*.
 - *Chryseobacterium haifense* DSM 19056^T - Does not produce Flexirubin pigments, probably incorrectly classified, to be renamed *Lycomia haifensis*
 - *Kaistella koreensis* CCG 49689^T - Moved into genus *Chryseobacterium* in 2009, based on similarity to *C. haifense*

- Reference Strains** – with Genome sequences available
- *Flavobacteriaceae* sp. 3519-10 – Isolated from Ice core over Lake Vostok, not yet officially classified, to be named *Lycomia vostokensis*.
 - *Chryseobacterium gleum* ATCC 35910^T – Sequenced as part of the human microbiome project

References

- Aziz, R., Bartels, D., Best, D., DeJongh, M., Disz, T., Edwards, R., et al. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75.
- Carver, T., Berriman, M., Tivey, A., Patel, C., Böhme, U., Barrell, B., Parkhill, J., and Rajandream, M. (2008). Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24, 2672-6.
- Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Swidan F, Thierer T, Wilson A (2012) Geneious v5.6. Available from <http://www.geneious.com>

Acknowledgements
This work was supported by supported by NSF Award DBI-1061893 to Michael D Boyle. The authors also thank Deb Grove and the staff of the Penn State University Genomics Core Facility

