

An Extreme Metabolism: Iso-Seq analysis of the Ruby-Throated Hummingbird



Rachael Workman¹, Kenneth Welch², G. William Wong³, Winston Timp¹

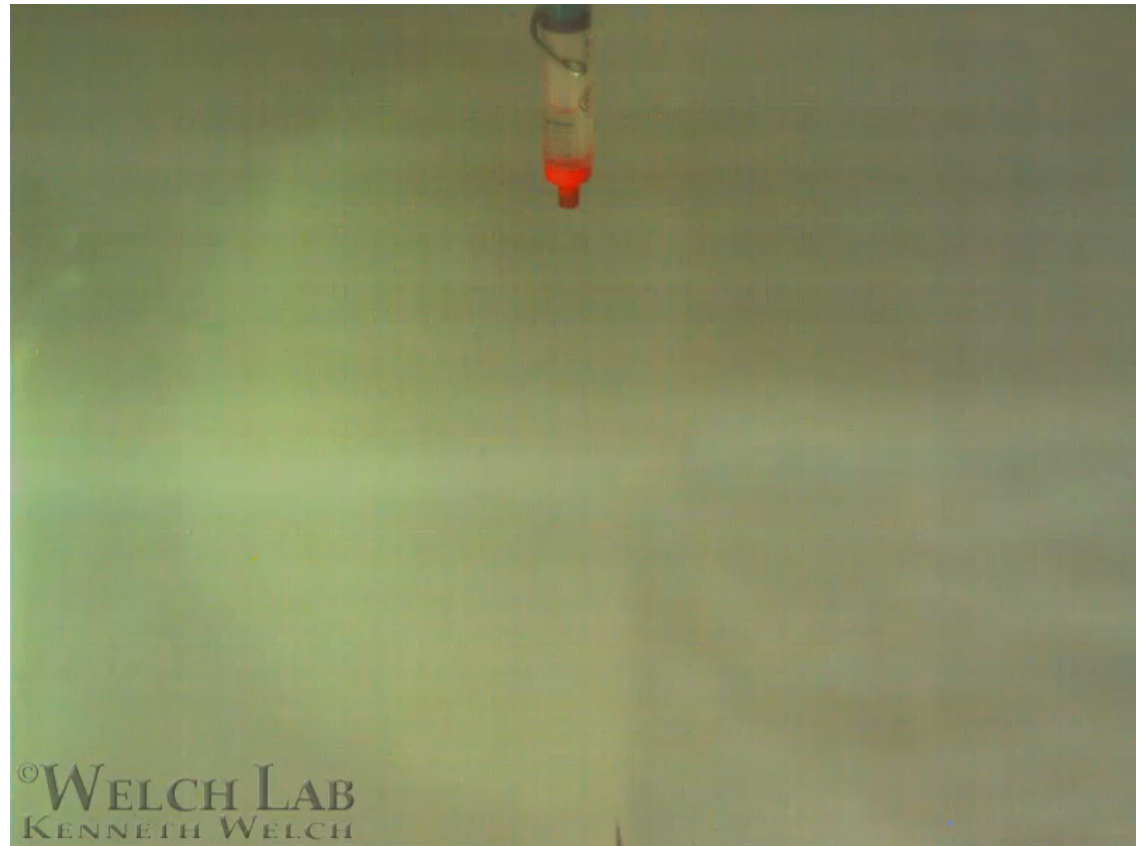
¹Department of Biomedical Engineering, Johns Hopkins University

²Department of Biological Sciences, University of Toronto

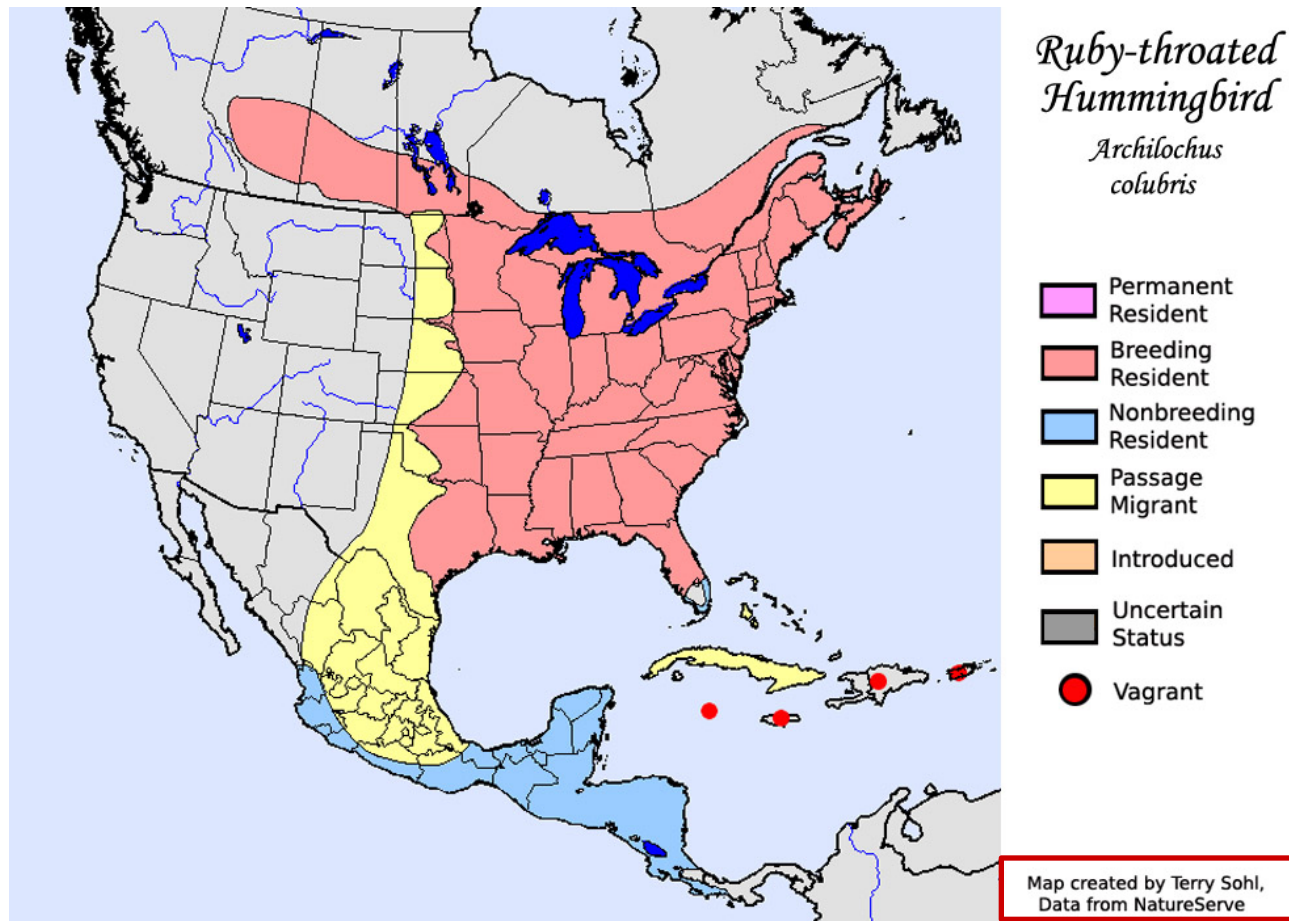
³Department of Physiology, Johns Hopkins University

Hummingbird Amazing Facts

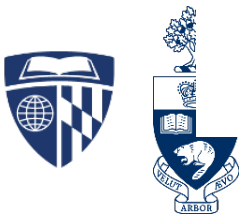
- Can fly at 25 mph
- Heart rate to 1260 bpm when flying
- Hummingbirds can enter torpor
 - Body temperature change up to 20C
 - 95% metabolism reduction
 - 20 min to recover upon waking



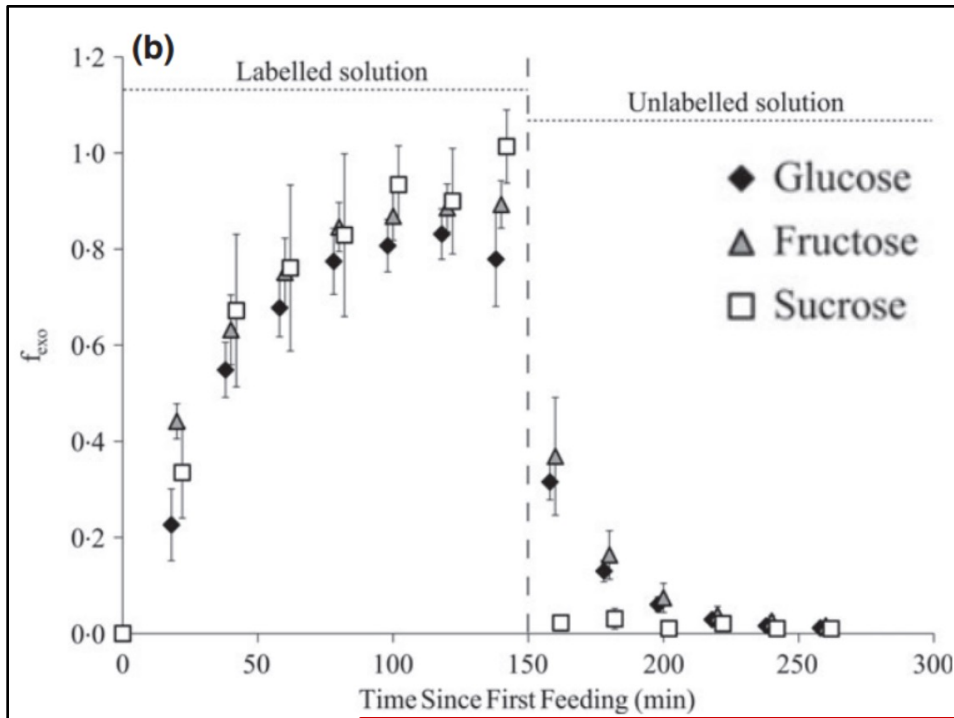
Migration



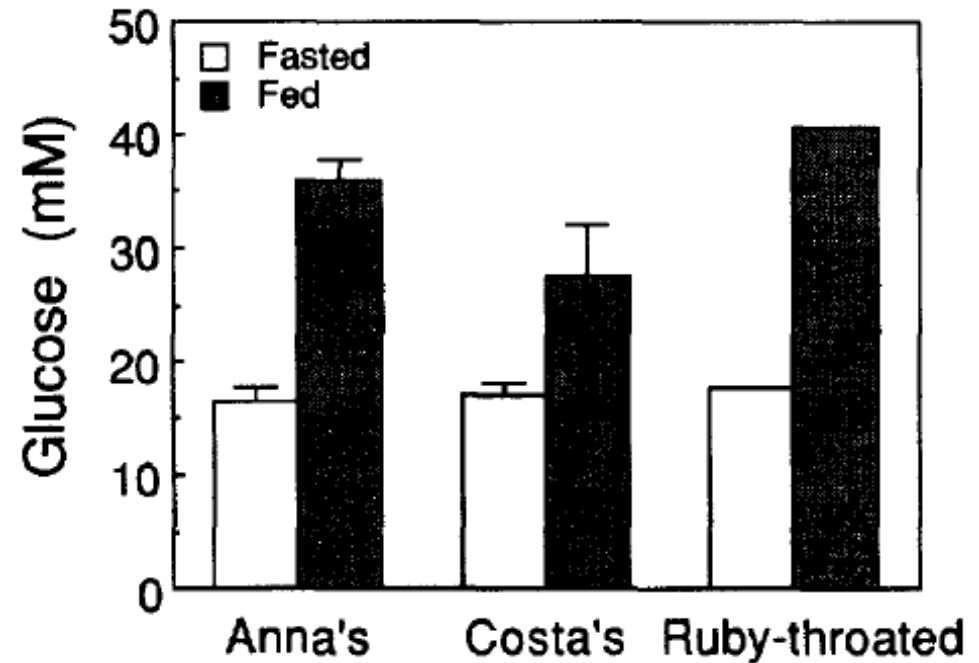
- Despite its high metabolic demands, the hummingbird migrates over 500 mi (800 km); in some cases cutting across the Gulf of Mexico (20 hr flight)
- To fuel this flight, the hummingbird can gain up to 50% of its pre-migration body mass



Hummingbird Fuel

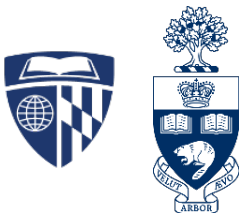


Chen and Welch (2014) Functional Ecology

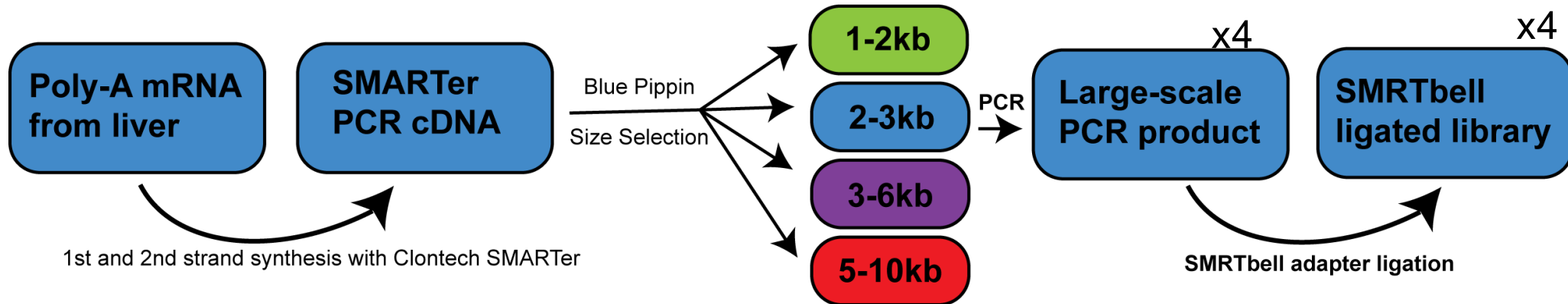


Beuchat and Chong, 1998

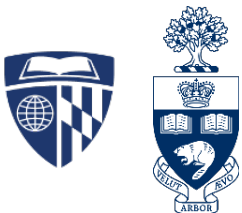
- Fuel (nectar) is typically around 50% sucrose, 25% glucose 25% fructose.
- Sugar is rapidly turned over – average sugar molecule half-life ~20 min. Similar in humans, but we have insulin!
- Any sugar not immediately used is shifted to fat storage
- Human fasted glucose 4-5mM, fed <10mM, compared with ~20 mM fasted with hummingbird. Human DKA at 13-33 mM



IsoSeq Library Prep

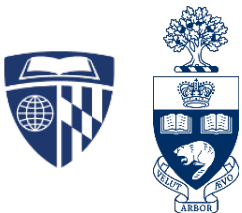
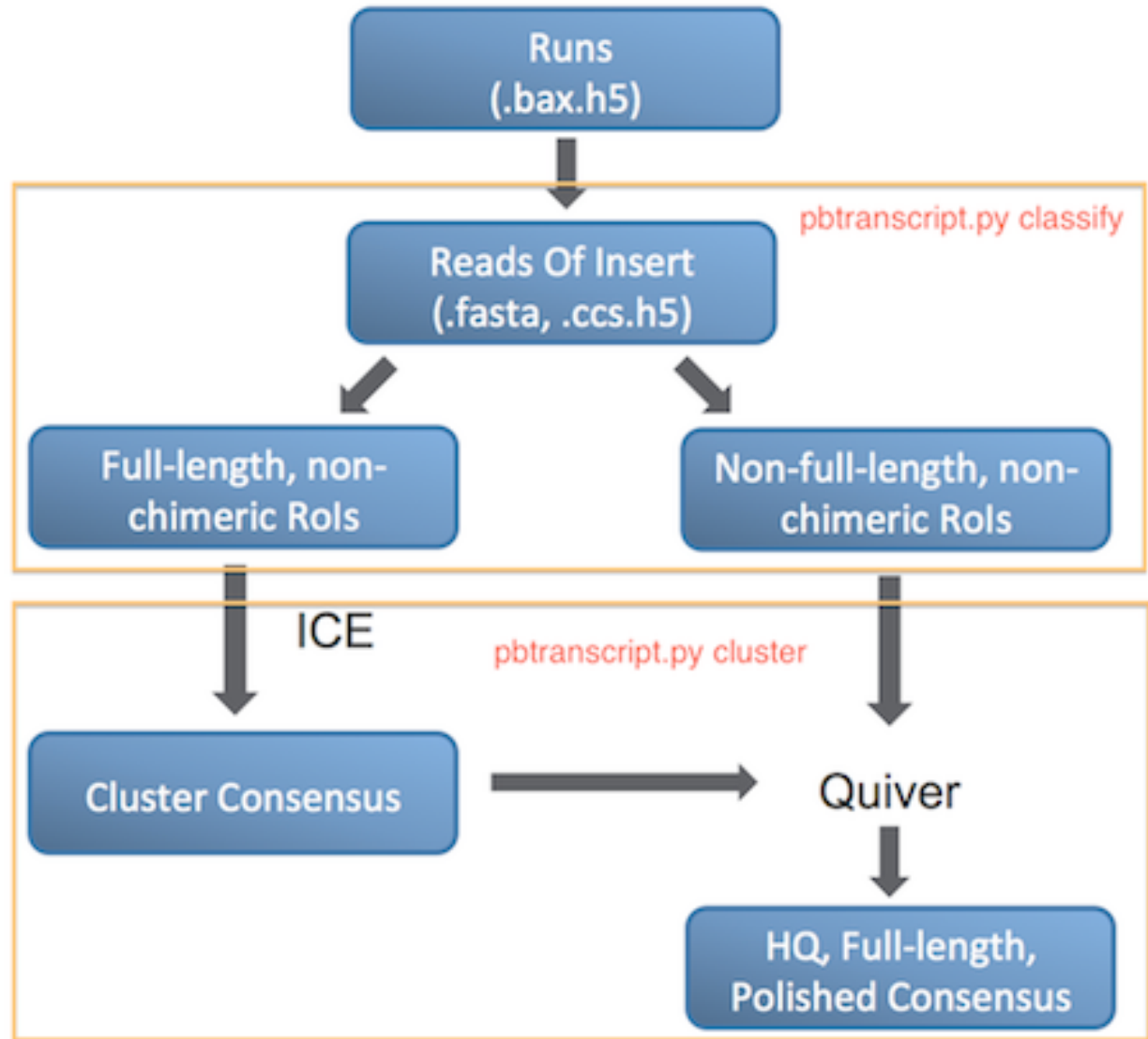


- Total RNA extracted from ruby-throated hummingbird livers (important for controlling metabolism)
- RNA is converted to full-length cDNA using Clontech SMARTer kit – primed against polyA tail
- Fragments are size selected into 4 pools
- PCR amplification, then ligation of PacBio “SMRTbell” adapters
- Each size fraction is run on 10 SMRTcells on a PacBio RSII (JHU), 40 cells total



IsoSeq SMRT analysis

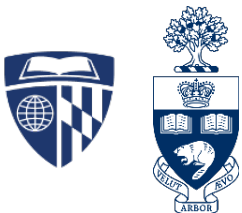
- Circular consensus (CCS) reads generated by the sequencer processing repeatedly through single molecule
- These long reads are then used to generate a read of insert (ROI)
- Full length = poly-A tail + F and R Clontech primers
- Custom AML on AWS for large dataset



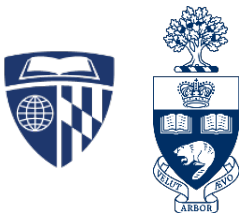
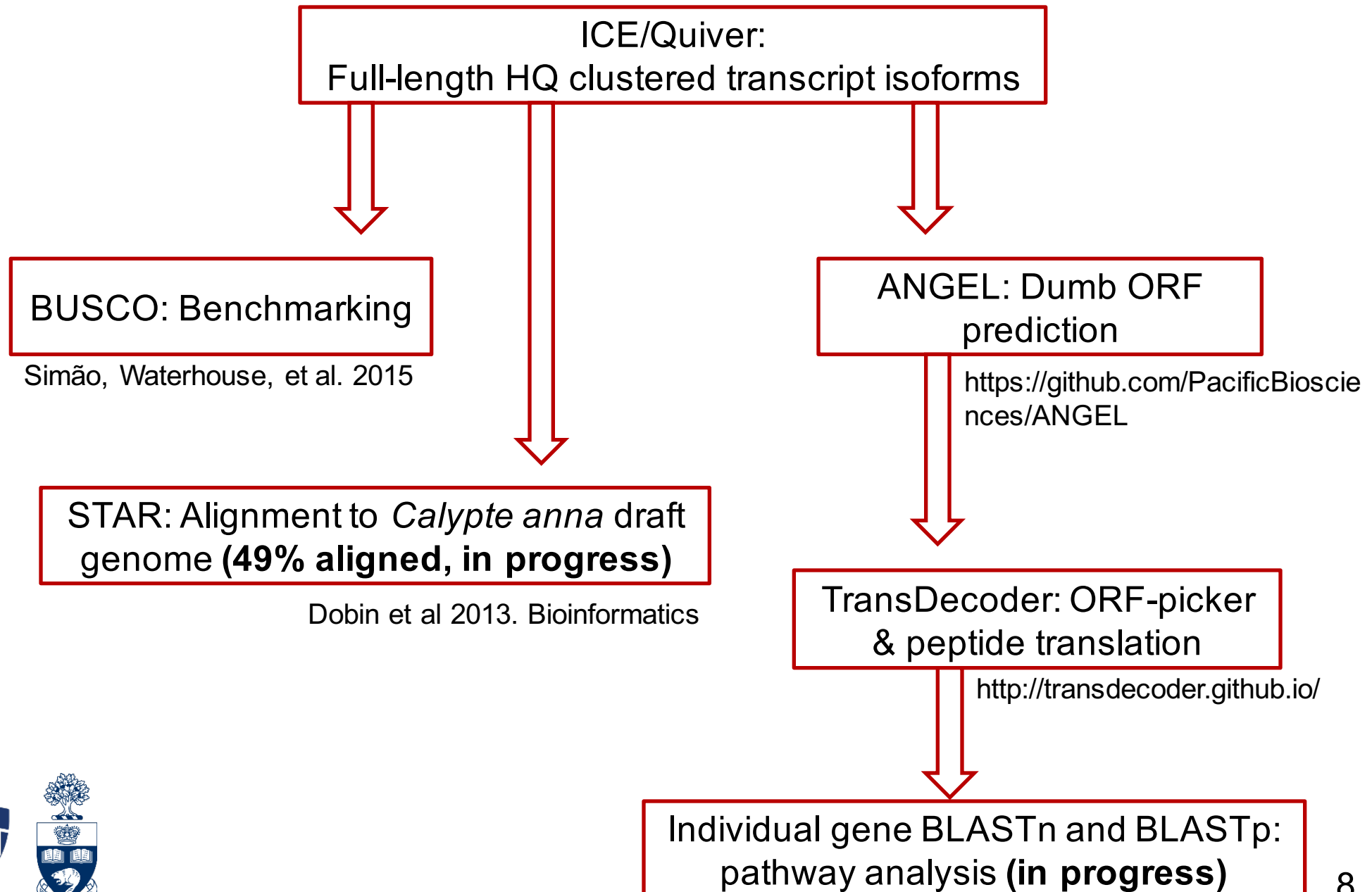
QC Stats

Size Fraction	1-2kb	2-3kb	3-6kb	5-10kb	Total
# of cells	10	10	10	10	40
Reads of Insert (ROI)	688,069	591,050	735,670	625,194	2,639,983
Avg length ROI (bp)	1533	2464	3650	5444	
ROI Yield (Mbp)	1055	1457	2685	3404	8601
Filtered Reads (FLNC)	430,381	306,841	272,781	193,906	1,203,909
# Consensus Isoforms	101,792	101,035	126,911	123,696	453,434
Consensus Yield (Mbp)	143	250	502	806	1701
# HQ consensus isoforms	52,538	48,282	49,390	20,021	170,231
% HQ	51.61%	47.79%	38.92%	16.19%	37.54%
Avg HQ length	1315	2329	3629	5491	
HQ Yield (Mbp)	69	112	179	110	470
# LQ consensus isoforms	49,254	52,753	77,521	103,675	283,203
% LQ	48.39%	52.21%	61.08%	83.81%	62.46%
Avg LQ length	1503	2621	4170	6718	
LQ Yield (Mbp)	74	138	323	696	1231

- ~1M reads after filtering our 40 flowcells, for ~9Gb of data
- Number of “LQ” clusters increases as the clusters get longer, less transits of polymerase around SMRTbell



Initial analysis



BUSCO results

Archilochis colubris

C:62%[D:36%], F:9.6%, M:27%, n:3023

1885	Complete BUSCOs
782	Complete and single-copy BUSCOs
1103	Complete and duplicated BUSCOs
293	Fragmented BUSCOs
845	Missing BUSCOs
3023	Total BUSCO groups searched



- Of 3023 ortholog groups “common” to all vertebrates, 62% were found in our assembly.
- For comparison, we ran BUSCO on chicken (*Gallus gallus*) mRNA from NCBI. Found 68% missing BUSCOs shared.

Gallus gallus

C:41%[D:8.6%], F:8.0%, M:50%, n:3023

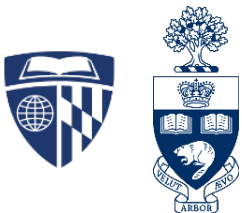
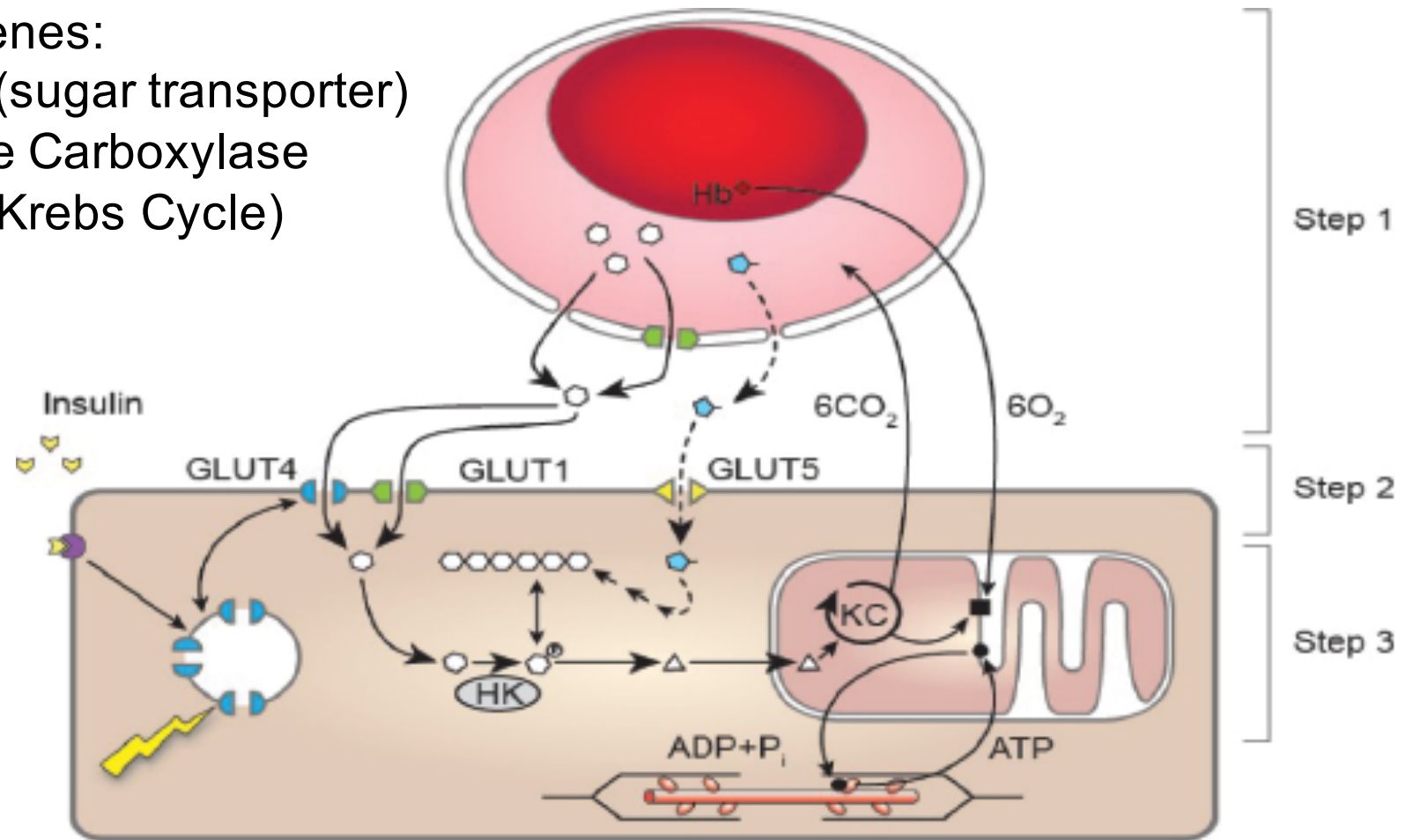


1267	Complete BUSCOs
1007	Complete and single-copy BUSCOs
260	Complete and duplicated BUSCOs
242	Fragmented BUSCOs
1514	Missing BUSCOs
3023	Total BUSCO groups searched

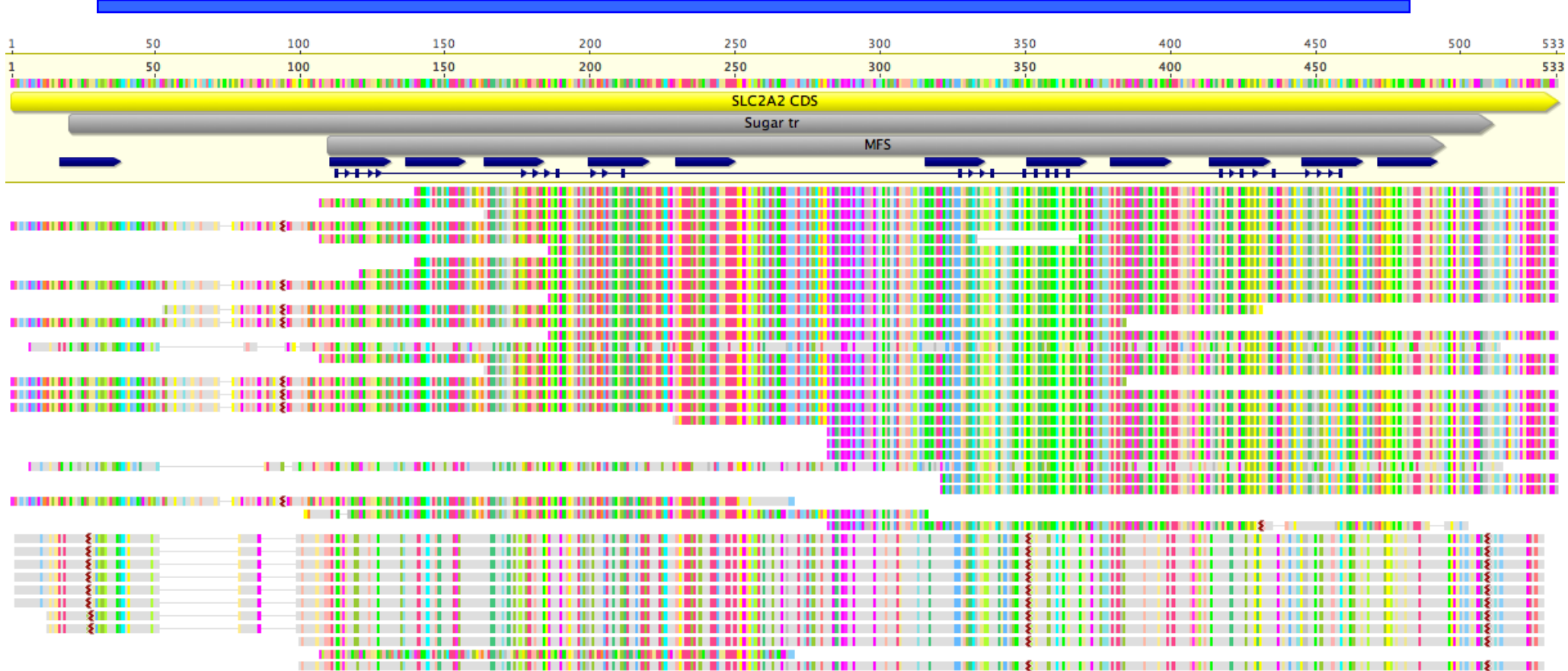


Metabolic Pathway

- Whole-transcriptome analysis in progress
- Some snapshots of key metabolic genes:
 - GLUT2 (sugar transporter)
 - Pyruvate Carboxylase (Part of Krebs Cycle)



GLUT2 (SLC2A2)

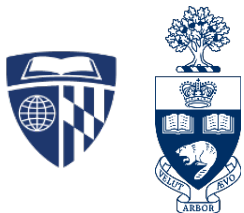
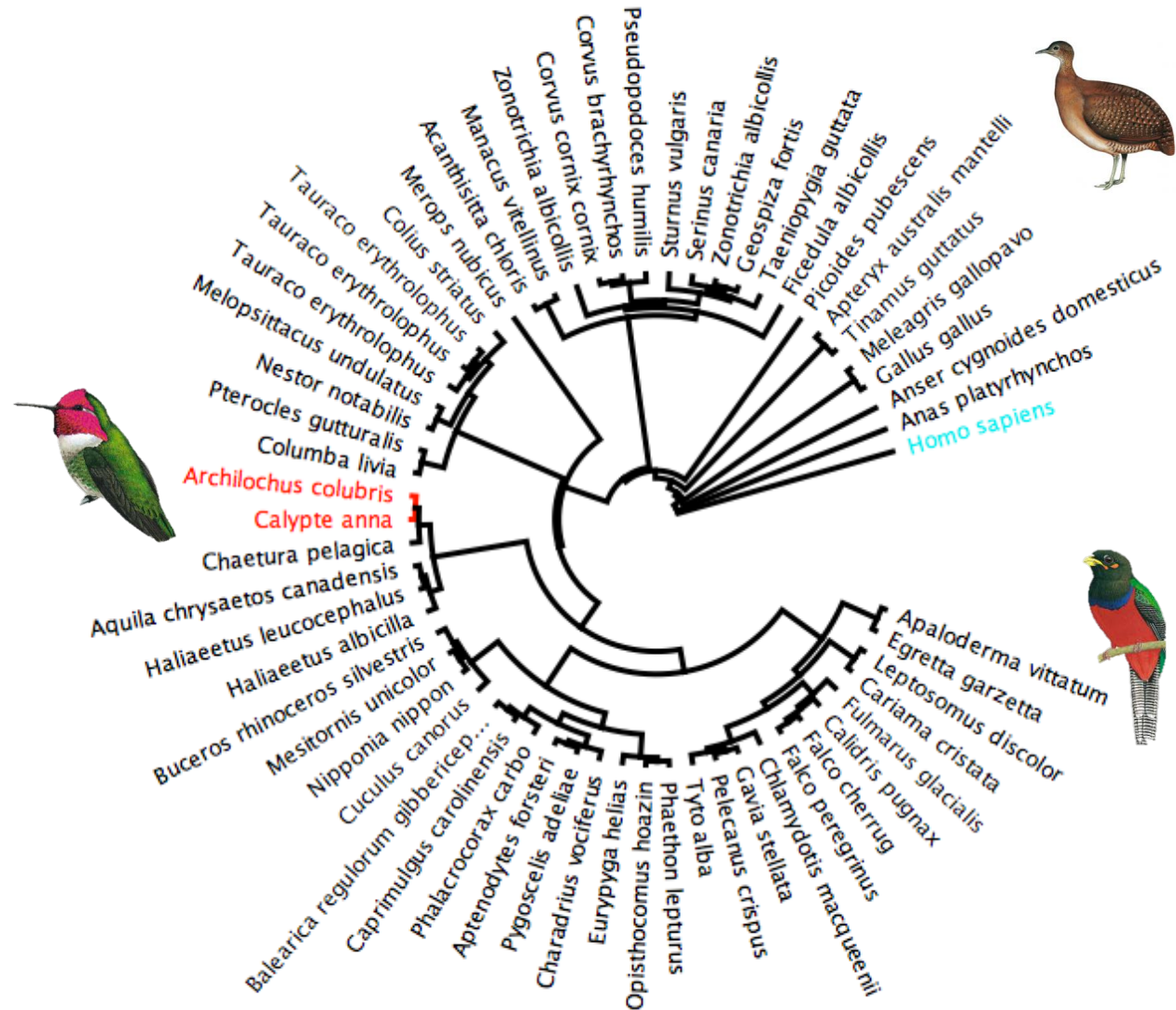


- GLUT2 is a glucose transporter primarily expressed in liver; principle transporter between liver and blood
- We BLASTed the chicken (*Gallus gallus*) form of GLUT2 against translated ORFs of our proteins. Each line= a cluster with X number supporting reads
- Highlighted (colored) parts of translated IsoSeq ORFs match, grey is mismatch

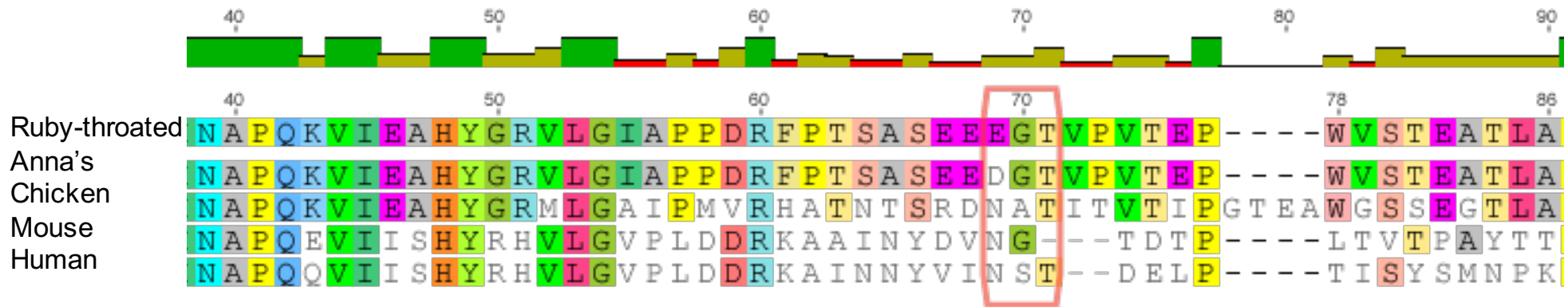


GLUT2

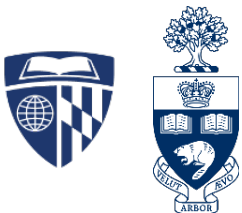
- BLASTing one of the protein hits against full RefSeq protein database had primarily avian hits
- Phylogenetic tree analysis showed closest response with *Calypte anna* as expected
- Identity for avian hits ranges from 83.46% (*Tinamus guttatus*) to 99.62% (*Calypte anna*)



GLUT2 Glycosylation

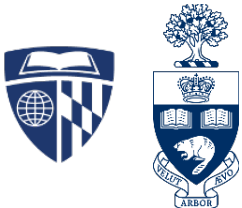


- We noticed in that low identity region a N(S)T motif which is a glycosylation motif, but absent in both hummingbird sequences.
- This may alter the transport of the protein and hence its trafficking.



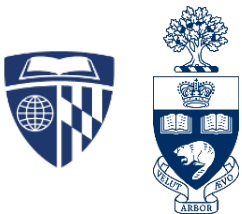
Pyruvate carboxylase (PC)

- Pyruvate carboxylase: catalyzes the creation of oxaloacetate from pyruvate in mitochondria: Krebs cycle
- Chicken PC mRNA does not align in the chicken genome
- *Calypte anna* doesn't have a predicted PC transcript
- PC not generated as a single transcript?
Separate domains seem to be on independent ORFs



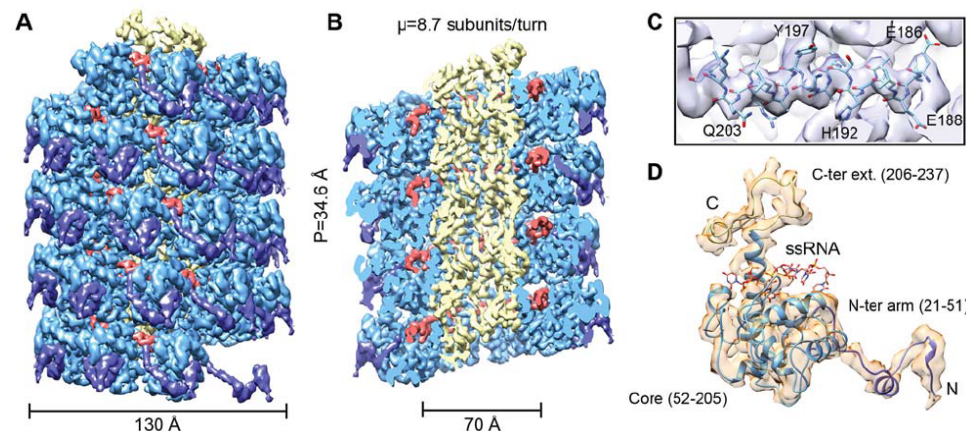
Conclusions

- We were able to generate full-length mRNA transcripts using the Iso-seq protocol
 - Longest ID'd transcript award (so far): ACC1, >7kb nucleotides with 2323 amino acids!
- More careful clustering parameters needed to reduce # of isoforms with one read only
- Cursory results show some essential genes to be quite different from avian relatives, need whole-transcriptome analysis to see pattern universality

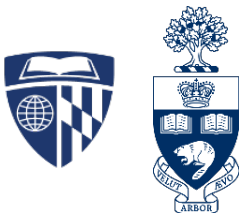


Future work

- Extensive BLAST studies to examine orthologous and non-orthologous genes in key metabolic pathways
- dS/dN to explore gene conservation in coding regions
- As part of Human Frontiers Grant:
 - Iso-seq of different tissue types
 - Cloning of interesting genes and/or splice variants to examine protein product
 - Cryo-EM to probe protein shape and function



M.Valle et al, eLife 2015



Acknowledgements

- University of Toronto-Scarborough
 - **Ken Welch**
 - Alex Myrka
- Johns Hopkins University – Center for Metabolism
 - **Will Wong**
- Johns Hopkins University – Biomedical Engineering
 - **Winston Timp**
 - Rachael Workman
 - Yunfan Fan
- Johns Hopkins University – Deep Sequencing Core
 - Haiping Hao
- Pacific Biosciences
 - Emily Hatas
 - Yan Guo
 - Roberto Lleras

