

ISGC Update Sheep Genomics for Discovery & Application: Driving Innovation in Research and Breeding

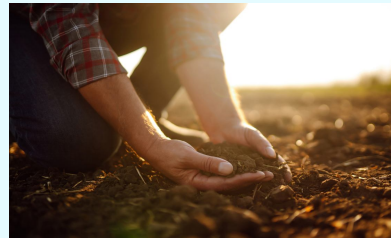
Shannon Clarke and Rudiger Brauning



New Institute for Bioeconomy Science (Bioeconomy Science Institute)



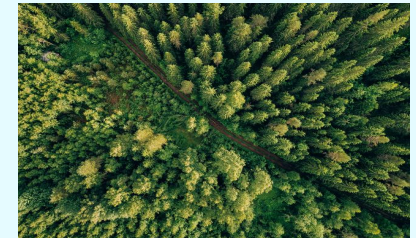
**Pastoral, agrifood
and agritech**



**Terrestrial
biodiversity and
land resources**



**Horticulture, arable,
seafood and food
and beverage**



**Forestry,
biomaterials and
wood-derived
products**

Background

Continued advances in sheep genomics have transformed both fundamental research and practical breeding strategies, enabling efficient approaches to genetic improvement.

- Underpinning resources through to industry applicable tools
- Complement aligned initiatives (e.g. FAANG, PanGenome, FarmGTex....)
- Genotyping tools to enable the genetic variation in flocks/populations/breeding schemes to be captured for your research and/or drive genetic gain in industry
 - developed a suite of SNP array based genotyping tools and lowpass sequencing
 - compliment and enable compatibility between platforms.

ISGC Genomic tools

- 1995-2001
 - Microsatellite-based linkage maps produced through genotyping of the International Mapping Flock
- 2002
 - CHORI-243 BAC library and BAC end sequencing
- 2006
 - Virtual Sheep Genome released
- 2006-2009
 - INRA 1200-rad RH Panel and genome maps
 - USDA 5000-rad RH Panel and genome maps
- 2007
 - Sanger re-sequencing of 2700 loci, 6021 SNPs
 - 1536 SNP chip developed
- 2007-2008
 - Illumina GA sequence (0.8 x) and Roche 454 sequence (3 x) derived from six breeds
 - Illumina Ovine SNP50 BeadChip
- 2009
 - Ovine HapMap based on 50k SNP chip
- 2010-2013
 - SNP loci for Parentage
 - Resequencing of sheep breed diversity panel and HD SNP chip development (606K SNPs)
 - Sheep Genome Assembly OARv3.1 (published in Science 2014)
- 2015
 - Sheep Genome Assembly OARv4.0
 - Low density (15K) SNP chip + imputation
 - SheepGenomes Database (1000 genomes) project initiated
- 2017
 - Ovine FAANG project initiated
 - New genome assembly Oar_rambouillet_v1.0 (Baylor)



ISGC Genomic tools

2019

- Run 2 "1000" genome project completed

2020

- Improvement of Ramb_v2.0 (University of Idaho); ARS-UI_Ramb_v2.0

2021

- Ovine PanGenome project initiated
- Run 3 "1000" genome project initiated (Ramb_v2.0); ~1900 individuals
- Illumina AgResearch Sheep Genomics 60K plus array
~15USD (DNA extraction to results)

2023

- Y chr added to Ramb_v2.0 (Ramb_v3.0)
- Run 3 "1000" genome project re-initiated (Ramb_v3.0); ~3000 individuals
- Ruminant T2T

2024

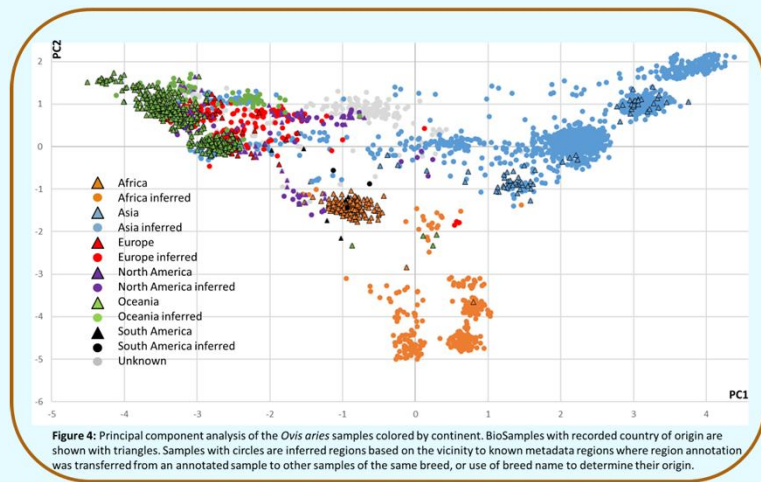
- Low-pass WGS as an alternative to Ovine HD SNP array established
- Illumina AgResearch Sheep Genomics 80K plus array (includes putative CNVs)
~15USD (DNA extraction to results)

2025

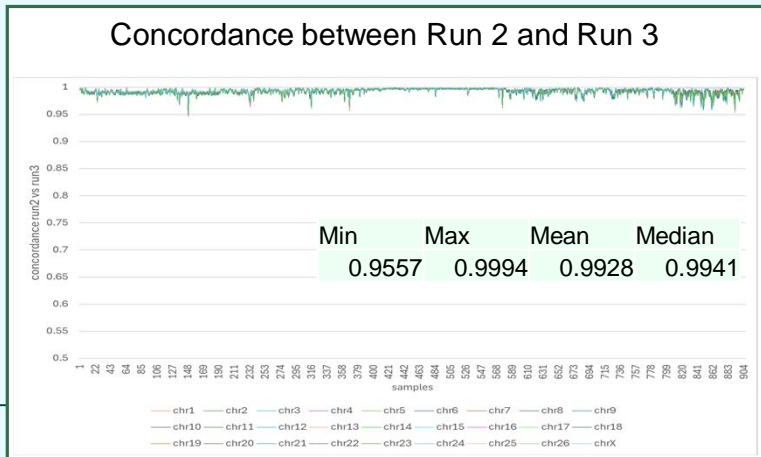
- Run 3 "1000" genome project completed; ~3600 individuals



ISGC "3000 genome" project



- Metadata for all 3522 animals available <https://doi.org/10.6084/m9.figshare.21775100.v1>
- Aligned against ARS-UI_Ramb_v3.0 reference genome
- Run3 processing was done via GATK best practise as implemented in nf_core's sarek pipeline.
- Due to the volume of data only individual gVCFs were produced through this pipeline before following GATK's "Biggest Practices" for joint calling supersized cohorts (GnarlyGenotyper).

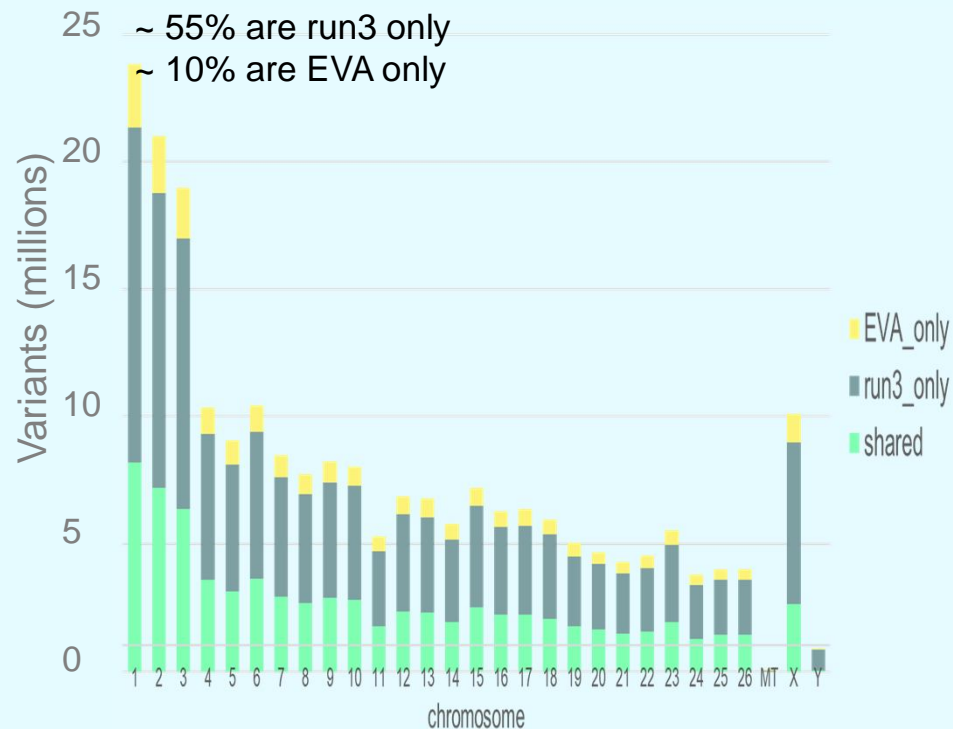


	Genus Ovis	Ovis aries
Variant count	251.6M	202.1M
File size	6.6 TB	3.3 TB
VCF files population-wide summaries		11 GB
VCF files with sample-level data		3 Tb
VCF files with sample-level data + filtering (MAF > 0.01 and missingness <0.9)		3 Gb



Positional comparison with EVA

- ~ 35% of the combined variants in EVA and run3
- ~ 35% are shared
- ~ 55% are run3 only
- ~ 10% are EVA only



Ensembl Working with Ensembl to make the variants via EVA

Aotearoa Genomic Data Repository (AGDR)

"Genomic Insights into Ovis: Detecting Variation Using Illumina Data and the nf-core Sarek Pipeline"

<https://doi.org/10.57748/sgr6-9k42>

Run3 versions

Large VCF files- provide multiple versions:

- Ovis aries OR complete ovis genus
- VQSR filtering (Variant Quality Score Recalibration)
- Population level details OR sample level details
- MAF and missingness
- "HD array" variants only

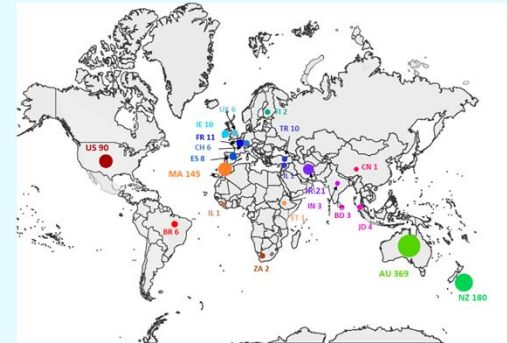


Genotyping tools: low-pass sequencing & 80K custom chip

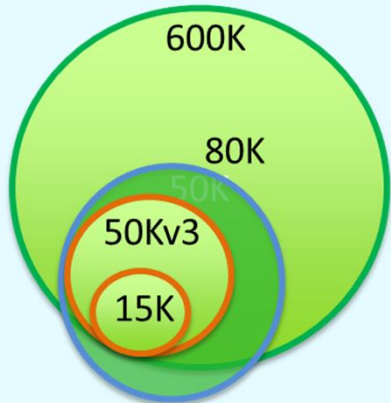


low-pass sequencing

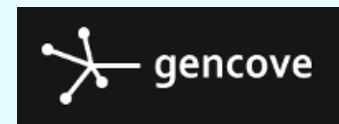
- Reference panel for imputation is ~1000 animals
- ~65 USD for the “package”
 - ~40 mil imputed variants (+ raw sequence data ~1x)
 - ISGC Illumina Ovine HD and 50K chip variants
 - *Includes running on new 80K custom array*
 - *Provides imputation concordance check*



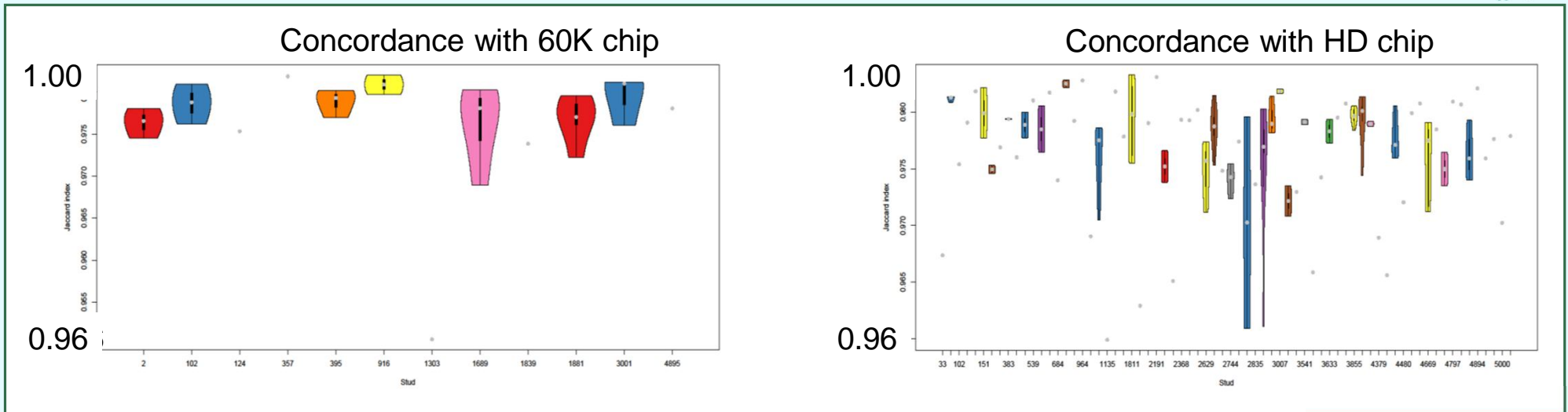
<https://www.sheepmap.org/>



	Min	Median	Mean	Max
All sites	92.5%	97.8%	97.7%	93.3%
PASS sites	98.4%	99.3%	99.2%	99.4%



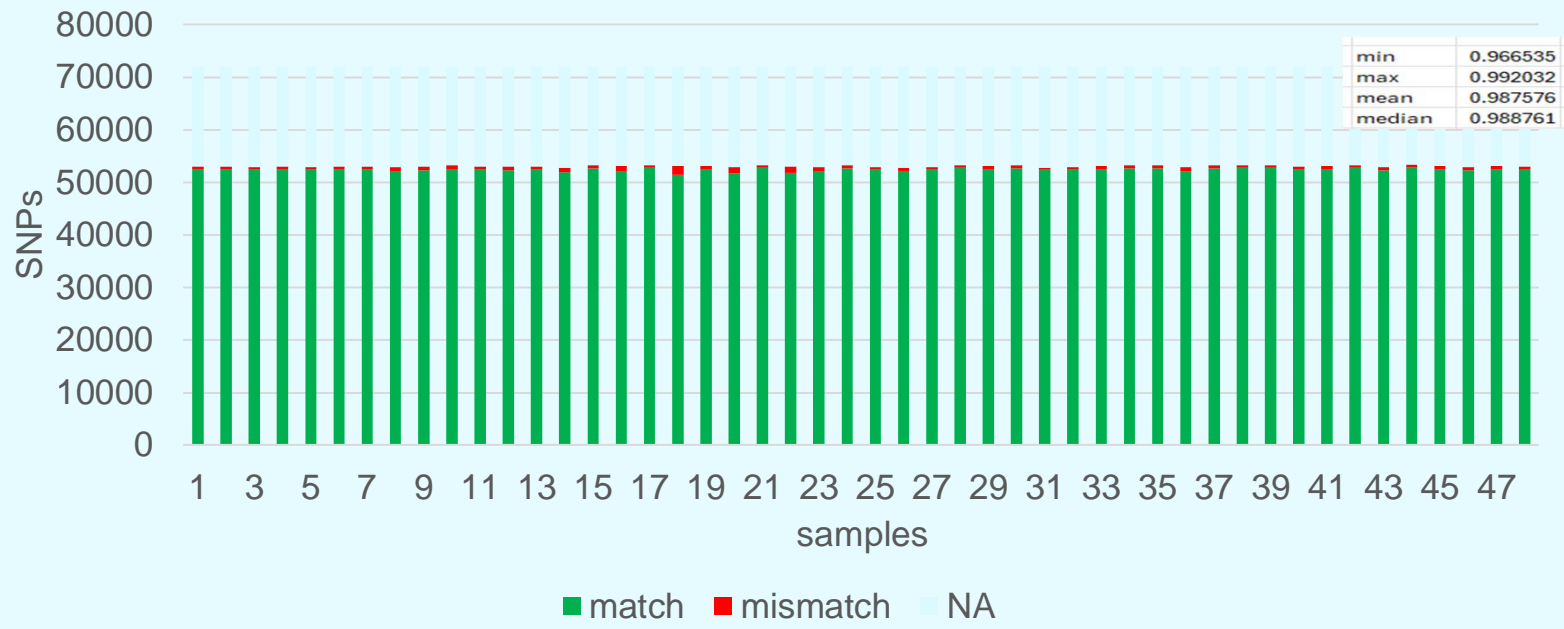
<https://doi.org/10.6084/m9.figshare.8424935.v4>



- * NZ updated ~2000 influential sires from past ~30 years
- * UK and Norwegian Sires
- * meat, wool, milk, shedding, hair breeds



48 US submitted samples (Merino and Rambouillet)



We can test your breed(s) on a few samples prior to submitting larger samples sizes

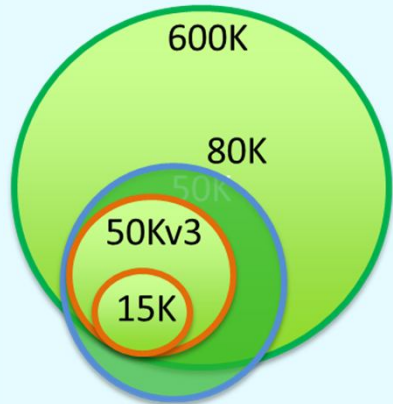




80K custom chip -Illumina XT



- AgResearch Sheep Genomic 80K plus array available (addon to AgR 60K)
- ~15-17 USD (includes DNA extraction, processing and uploading genotypes)
 - 724 *new* probes (612 variants)
 - 9276 *existing* (selected from the ovine HD) probes
 - 4583 putative CNVs
 - 4693 putative epistasis probes



Genomic tools: Demonstration in practice



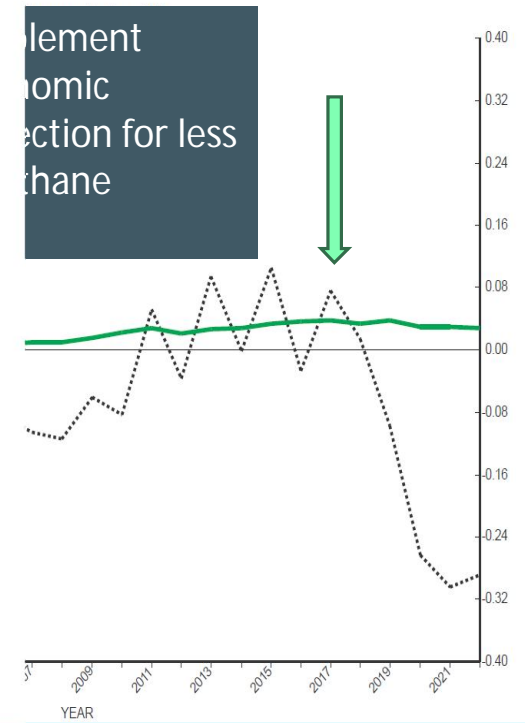
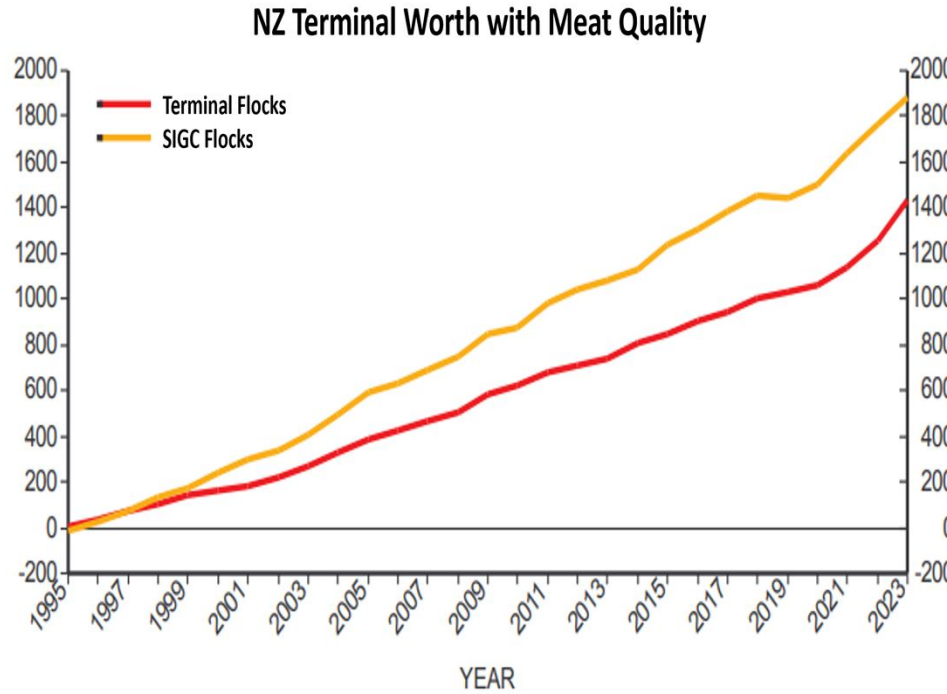
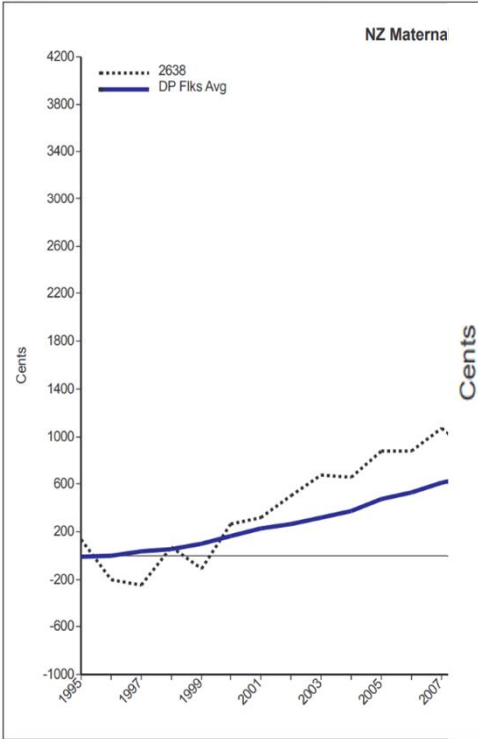
Genetic Trends for Flock 2638
 GE Analysis #42666 1/8/2025 Analysis Flocks: 1,2,7,8 + 1468 others

Page 1



Genetic Trends for Flock 2638
 GE Analysis #40757 11/8/2023 Analysis Flocks: 1,2,7,8 + 1410 others

Page 39

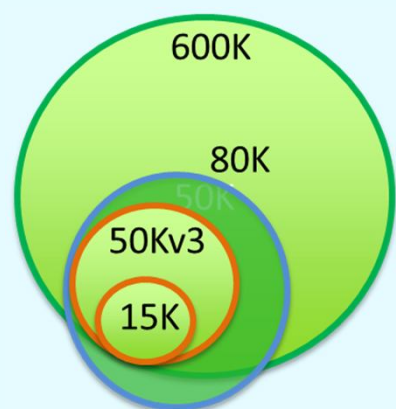




<https://www.sheephapmap.org/>

Shannon.Clarke@agresearch.co.nz

Rudiger.Brauning@agresearch.co.nz



Thank You!



<https://doi.org/10.6084/m9.figshare.8424935.v4>