Building genomic resources for breeding in a new species
– mink as an example

Zexi Cai
Center for Quantitative Genetics and Genomics
Department of Molecular Biology and Genetics
Aarhus University

Mink breeding tools before genomic

- Conventional genetic evaluation
  - Pedigree-based breeding value estimation

- Molecular markers
  - Sparse microsatellite map of 103 markers
  - A few QTL information but with wide confidence intervals
  - SNP (single nucleotide polymorphism) identified by RAD (Restriction site Associated DNA) sequencing without flanking information

Benefits of genome-wide markers and genome assembly

- Genomic prediction (like other livestock species)

- Genome-wide association analysis

- Genetic research on fur color, fur quality, disease-resistance, reproduction, feed efficiency and animal behavior

- Extract information from other closely related species

Development of genomic tools in mink

1. The *de novo* genome assembly of mink
2. Identification of genome-wide markers

- The *de novo* genome assembly of mink

- Identification of genome-wide markers

Mink (*Neovison vison*) genome assembly

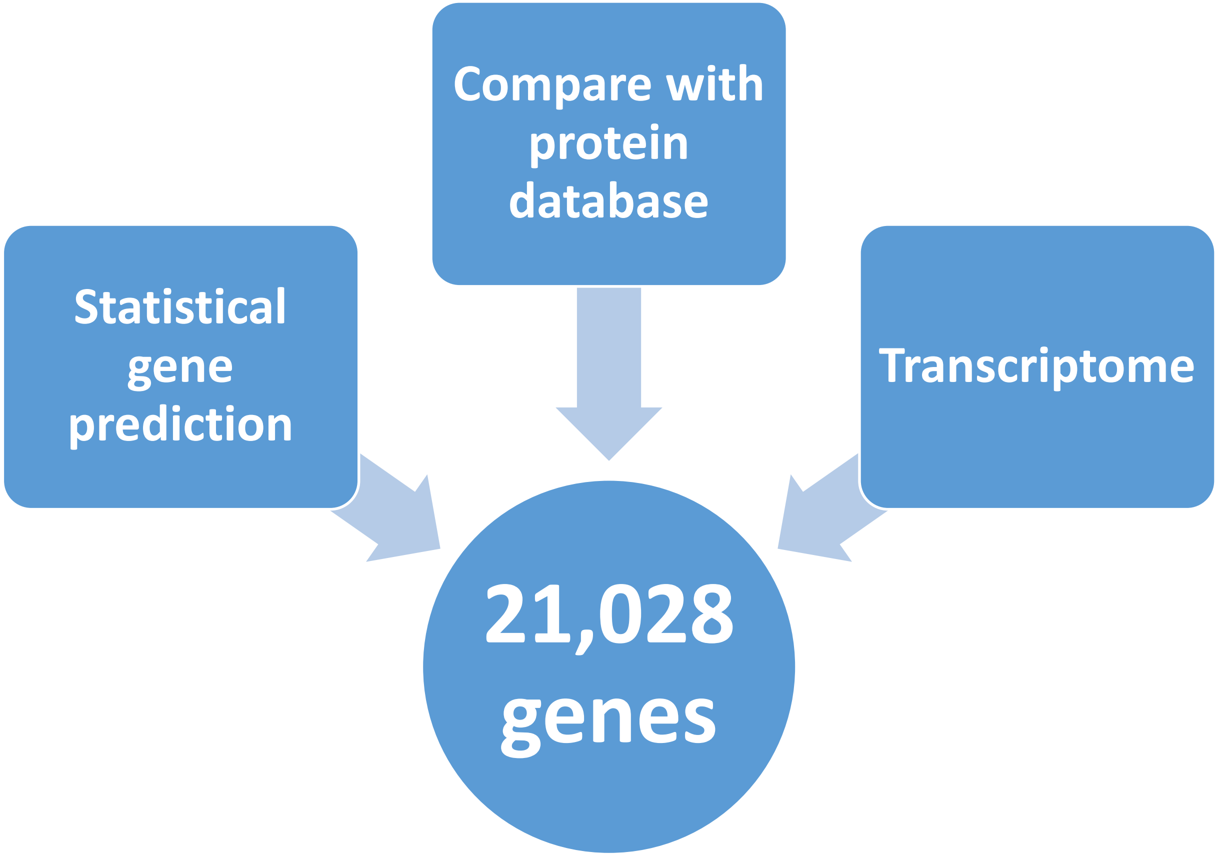
Whole genome sequence

1. A male pearl mink
2. A brown mink female

	Mink	**Ferret (Peng et al. 2014)**
Estimated Genome Size	2.7 Gb	2.7 Gb
Total length including gaps	2.45 Gb	2.41 Gb
Scaffolds (Fragments)	7175	7741
Scaffolds N50	6.8 Mb	9.3 Mb

Cai et al. Scientific Reports 2017 **7**:14564

Genes in the mink genome



Compare with protein database

Statistical gene prediction

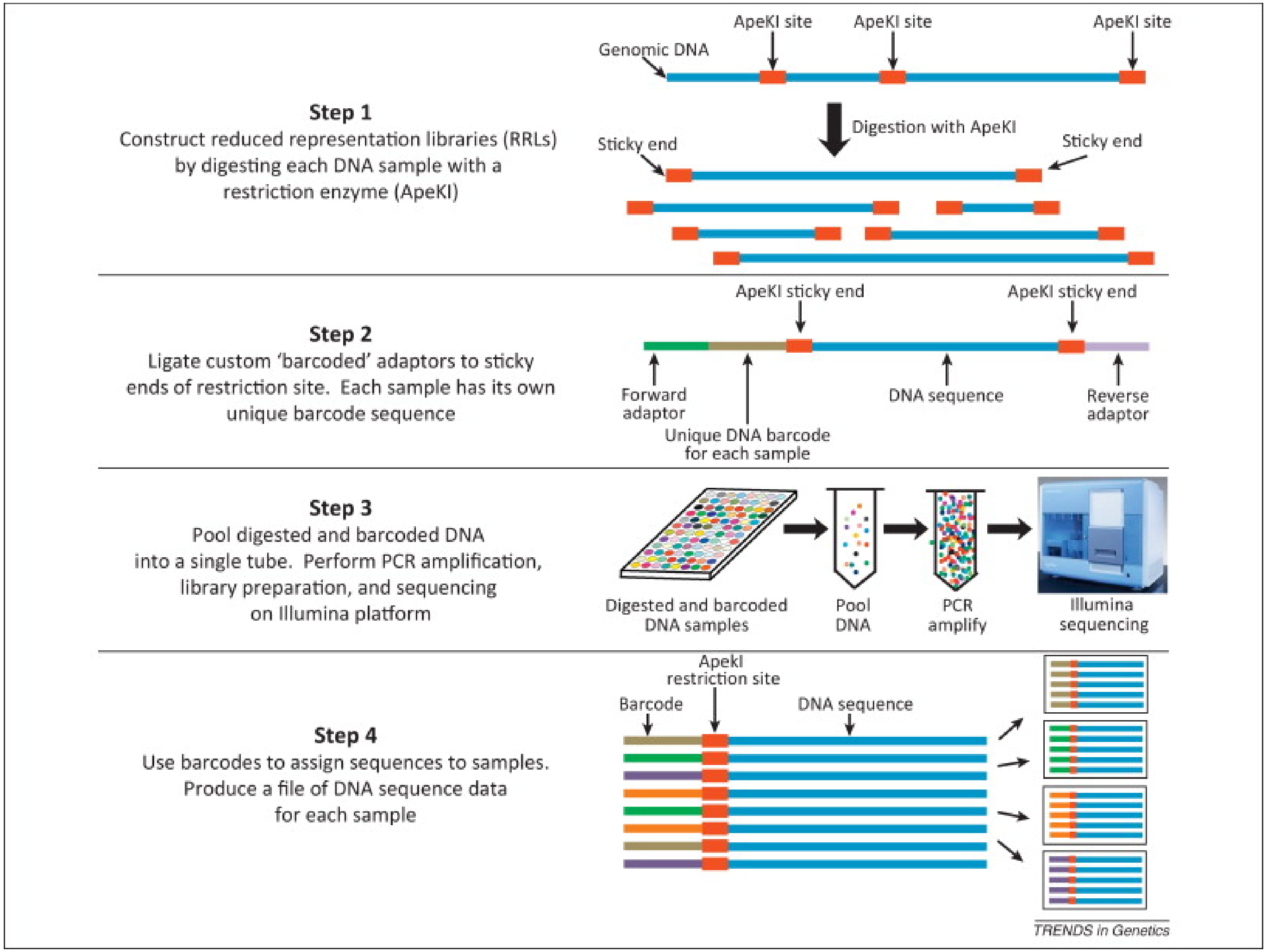
Transcriptome

21,028 genes

- The *de novo* genome assembly of mink

- Identification of genome-wide markers

GBS – genotyping by sequencing



Step 1
Construct reduced representation libraries (RRLs) by digesting each DNA sample with a restriction enzyme (ApeKI)

Step 2
Ligate custom 'barcoded' adaptors to sticky ends of restriction site. Each sample has its own unique barcode sequence

Step 3
Pool digested and barcoded DNA into a single tube. Perform PCR amplification, library preparation, and sequencing on Illumina platform

Step 4
Use barcodes to assign sequences to samples. Produce a file of DNA sequence data for each sample

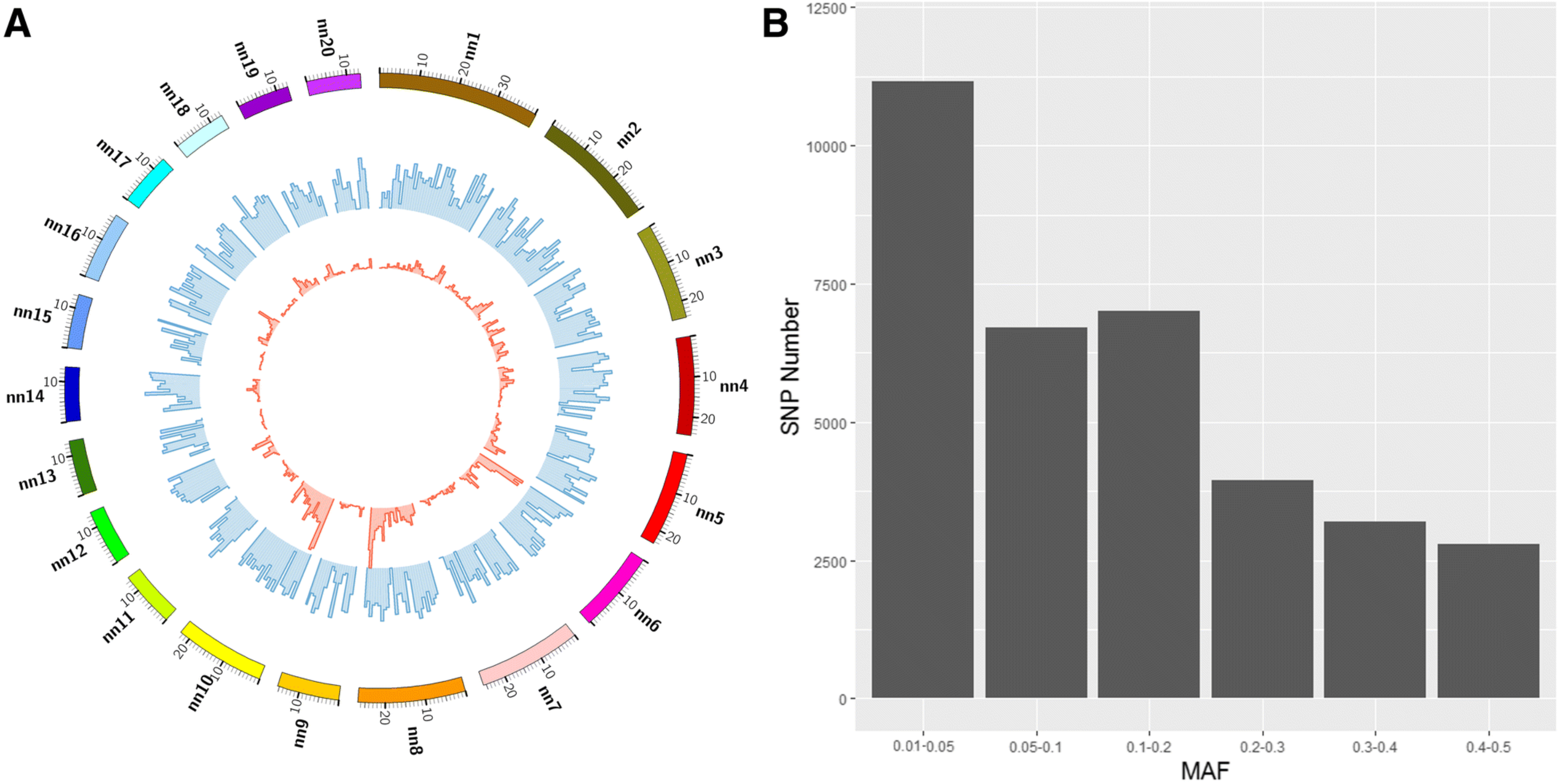
TRENDS in Genetics

(Sean Myles, 2013)

SNP calling statistics

Total number of samples	2,451
Total number of SNPs	34,816
Average distance between adjacent SNPs	57,265 bp
Scaffold number with at least one SNP	945
Scaffold number with at least two SNP	674
Scaffold number with at least ten SNP	335
Gene number with at least one SNP	7,589

Gene density vs. SNP density



Cai et al. BMC Genetics 2018 **19**:103

Markers capture the additive genetics variance

Trait	Unit	Records	Mean (std. error)	Heritability	Genomic heritability
Body weight	gram	2,265	74.59 (5.98)	0.49	0.48
Pelt density	1-3 (flat-filling)	1,995	0.11 (0.01)	0.22	0.21
Pelt length	cm	1,960	0.06 (0.09)	0.42	0.44
Pelt quality	1-12 (worst-best)	1,995	0.13 (0.05)	0.31	0.32

Application of genomic tools

1. Genomic selection

2. Genome-wide association study

Application of new tools

- Genomic selection

- Genome-wide association study

BLUP vs. SSGBLUP evaluation

- Evaluate the effect of genomic selection by comparing accuracy of breeding values estimated from BLUP and SSGBLUP evaluation.

- Genetic gain by introduction of genomic selection in mink will primarily have to come from increased accuracy of selection, it is difficult to:
    - Reduce generation interval
    - Increase selection intensity

Models

- Single trait BLUP vs. Single trait SSGBLUP
- The SSGBLUP incorporate marker information into the relationship matrix
- 5 fold cross validation

BLUP vs. SSGBLUP evaluation

Trait	h^2	Accuracy BLUP	Accuracy SSGBLUP	% Increase in accuracy
Live grading				
Body weight	0.532	0.255	0.426	67.1
Quality	0.303	0.548	0.615	12.2
Density	0.163	0.215	0.283	31.6
Silky	0.297	0.595	0.686	15.3
Guard Hair length	0.292	0.522	0.532	1.9
Guard hair thick	0.357	0.512	0.618	20.7
Clarity	0.202	0.273	0.287	5.1
Skin grading				
Fur length	0.457	0.371	0.451	21.6
Quality	0.327	0.315	0.390	23.8
Density	0.222	0.140	0.202	44.3
Silky	0.181	0.299	0.281	-6.0

Application of new tools

- Genomic selection

- Genome-wide association study

Candidate genes for body weight and pelt length

- *WWC3*: control organ size in mammals
- *MAP2K4*: mouse knockout lines showed reduced body weight
- *SLC7A1*: this gene plays a critical role in growth control during mouse development.
- *USP22*: mouse knockout lines showed decreased number of lumbar vertebrae and decreased body weight.

Cai et al. BMC Genetics 2018 **19**:103

Summary

- We build a high quality genome assembly for mink.

- Using GBS, we obtained genome-wide markers set which showed the potential for acceleration of mink research and breeding application.

- By incorporating marker information, we gained higher prediction accuracy which showed the potential in mink breeding.

- The marker set along with the reference genome help us to find candidate genes for body weight and pelt length.

Acknowledgement

Mogens Sandø Lund

Bernt Guldbrandtsen

Goutam Sahana

Trine Michelle Villumsen

Torben Asp

GenSAP

InnovationsFonden
FORSKNING, TEKNOLOGI & VÆKST I DANMARK

KOPENHAGEN
FUR

AARHUS UNIVERSITY