NORWEGIAN UNIVERSITY OF LIFE SCIENCES

Development of a traceability tool using SNPs

FHF prosjekt 900706:

”Sporing av laks: SNP-tilnærming”

Matthew Kent, Harald Grove, Teresa Andersstuen, Mariann Arnyasi, Kristil Sundsaasen, Sigbjørn Lien

CIGENE, Dept Animal and Aquaculture Sciences, Norwegian University of Life Sciences UMB, Ås, NORWAY

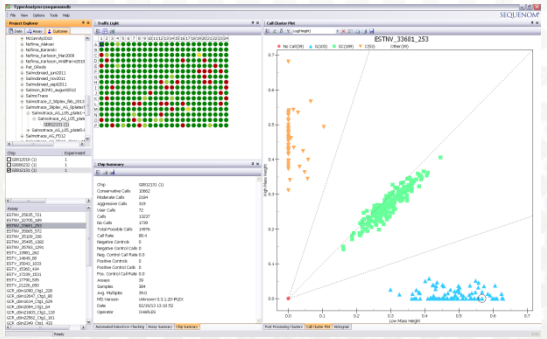
SNP marker panel development

- ca. 6,000 SNPs genotyped in 756 samples belonging to MOWI, Salmobreed and AquaGen breeding populations was analysed.
- A subset of SNPs (n=114) was identified using the following criteria:
  - SNPs must have high minor allele frequency (MAF >0.45) in all three breeding populations
  - 3-4 SNPs from each chromosome and a wide physical distribution
  - 3 SNPs from mitochondrial genome to provide extra assurance for female assignment

Genotyping protocol

- A panel of 60 SNPs was genotyped using Sequenom MassArray 4 instrument.

- Genotype assignment is automatic but usually requires manual inspection

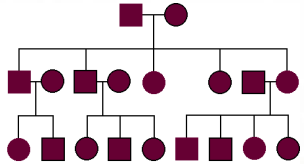


- DNA extraction from fin clips has been semi-automated using a BioMek3000 robot. Process includes tissue digestion, protein – nucleotide separation, and DNA precipitation.

Genotyping protocol - Challenges

- DNA extraction is producing sufficient amounts of high quality material, however contaminant in the DNA extract seems to inhibit genotyping. This can be overcome by dilution.
- Genotyping 60 markers in a single reaction has not been robust, and the set is run as 2 separate reactions.
- "Holes" in the genotyping dataset weakens the power to assign. Missing parental genotypes is the most damaging situation.

Assignment principles



- Assignment script looks for incorrect Mendelian inheritance (MI) patterns to eliminate impossible offspring-parent pairings. For example:

Offspring 1 (AA)			→ Possible match
Offspring 2 (AB)	✚ Parent1 (AA)	?	→ Possible match
Offspring 3 (BB)			→ Impossible match

- <u>After</u> identifying *candidate parents*, software considers MI patterns within all possible parent-parent-offspring to eliminate incorrect trios. For example:

Offspring 2 (AB)	?	P1 (AA) + P2 (BB)	→ Possible trio
		P1 (AA) + P3 (AA)	→ Impossible trio

Assignment software - Challenges

- An incomplete set of genotypes (ie not all 60 SNPs) can lead to offspring being (i) unassigned, (ii) assigned to 1 parent only, (iii) assigned to multiple trios
- Additional information can be used to increase certainty and reduce trio combinations. Examples of additional information are:
  - Sex of parents
  - Parental crossings (ie mating scheme)
  - Relationships between parents

Validation study 1 – many offspring, few parents

- Samples from AquaGen breeding program, selected by NVH and sent with anonymous IDs to CIGENE for assignment. Includes:
  - 230 Parents
  - 520 Offspring
  - 40 unrelated offspring
- Require minimum 40 genotypes for assignment, allow 1 Mendelian mismatch :
  - 16 offspring (2.8%) failed to produce >40 genotypes

Validation study 1 - assignment

- 230 Parents
- 520 Offspring
- 40 unrelated offspring

Category	Description	Number
Unique – full	Unique Assignment, valid parent couple	496
Uncertain / unrelated / missing	One parent only / no parents / insufficient genotypes	64

Assignment validation rate = 97%

Validation study 2 – Many parents, few offspring

- Samples selected by NVH and sent with anonymous IDs to CIGENE for assignment. Includes:
  - 496 Parents
  - 279 Offspring
- Require minimum 40 genotypes for assignment, allow 1 Mendelian mismatch:
  - 10 offspring (3.5%) failed to produce >40 genotypes

Validation study 2 - assignment

- 496 Parents
- 279 Offspring

Category	Description	Number
Unique – couple	Unique Assignment, parent couple	185
Unique – single	Only one parent, other unknown	33
Multiple - couple	More than one valid couple	9
Uncertain / Missing	Multiple parent options and no valid couple/ insufficient genotypes	52

Assignment validation rate = 98%

Validation study 3 – Wild fish

- 95 wild fish (5 fish x 19 rivers) provided by NINA

Category	Description	Number
Mismatch – A	Mismatches ≧ 3	87
Mismatch - B	2 Mismatches	6
Uncertain / Missing	Multiple parent options and no valid couple/ insufficient genotypes	2

- No wild fish assigned to known parents

Extraction protocols

- To assess issues related to practical implementation of genetic tracing, samples were stored, extracted and genotyped under different conditions.

- Variables included:
  - Tissue type (bukfinne, fettfinne, skjell)
  - Preservation Method (ethanol, red-spirit, frozen)
  - Extraction (chelex, precipitation)
  - Operator (3 sites)

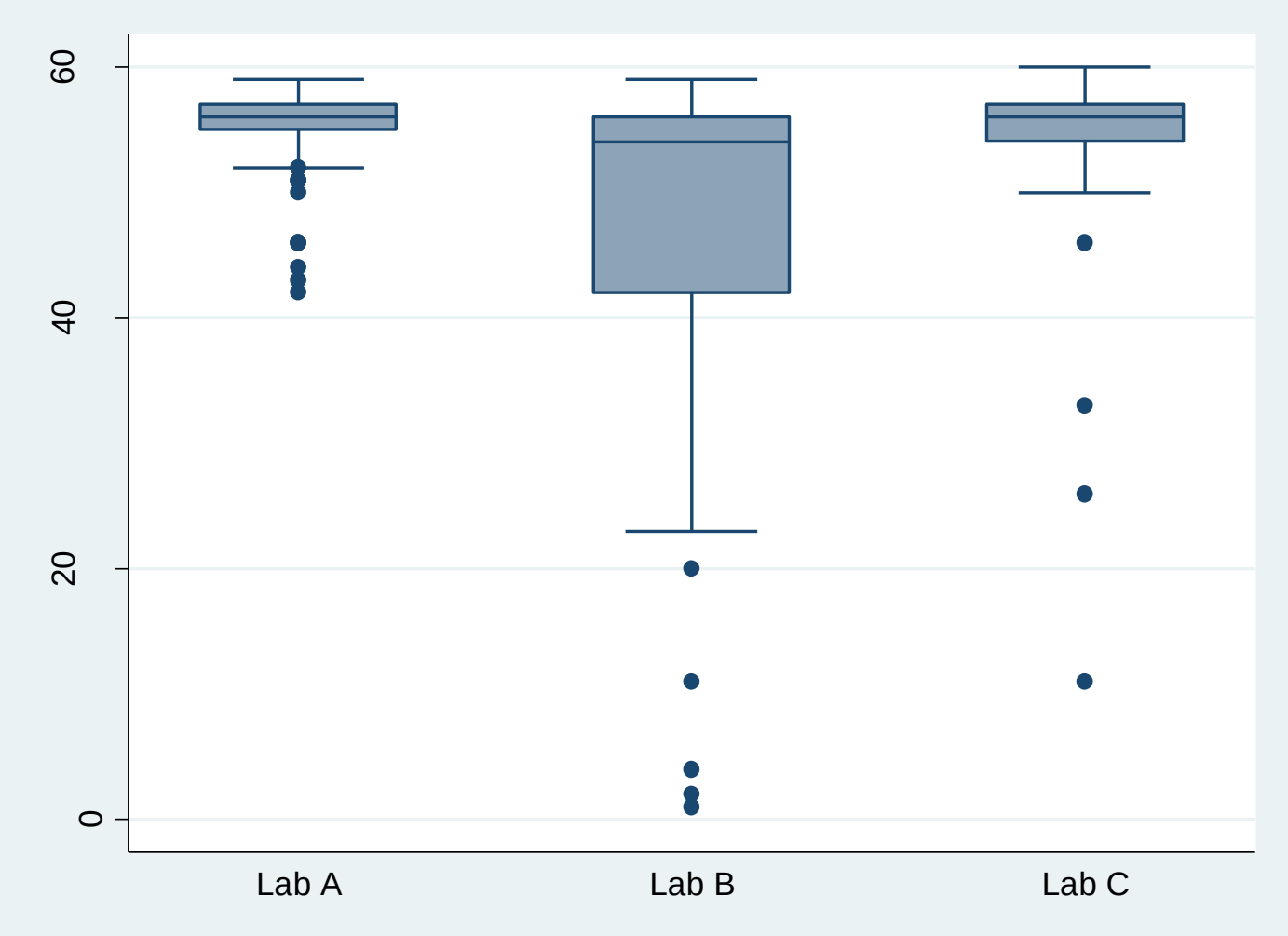
- Samples genotyped with SNPs, "number of genotypes" used to represent DNA quality.

Significant effects

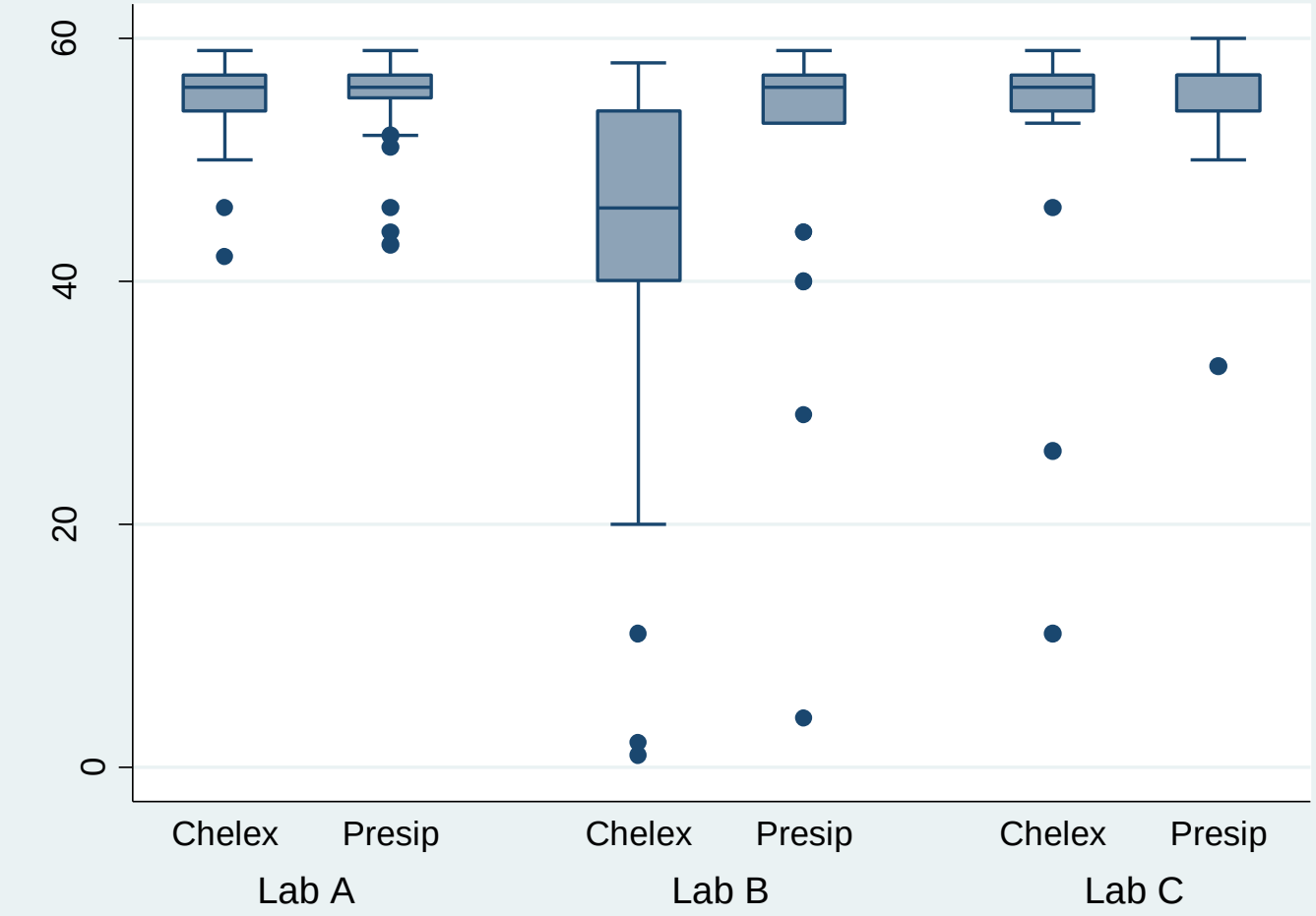
- Statistical analysis
  - effect of tissue, preservation, extraction, and site on # genotypes reported
  - generalized linear models with gamma distribution and identity link function
- # genotypes ranges from 1 – 60, mean 52, median 56

Effect of …	… on number of genotypes	P value
Tissue	Not significant	0,7
Preservation	Not significant	0,8
Extraction	Significant (precip gave 3.7 more genotypes than chelex)	0,018
Site	Significant (Sites A and C gave 7-8 more genotypes than B)	0,001

Site vs. #genotypes



Site and Extraction vs. #genotypes



Conclusion

- SNP based technology demonstrates good ability to assign farmed fish to their parents.
- The current SNP set has the power to differentiate wild and farmed.
- Standardization (and optimization) of the DNA extraction methodology is important.
- Further implementation of SNP-tracing would benefit from a redesign of the SNP set to include more markers and/or achieve single assay throughput.

Synergies

- Routine testing of salmon DNA using SNP technology is performed large scale today, including:
  - Sample collection,
  - sample storage,
  - DNA extraction,
  - genotyping and reporting of data
- Thoughtful design of SNP panels can create added value, ie markers can be included that not only allow for tracing, but can provide broodstock information to producers.