

Bio-informatics: Evolutionary and Quantitative Genetics I0D53A

Population genomics

Class 4: Geo-genomics

(Prepared by dr. Joost Raeymaekers)

Prof. Filip Volckaert (lecturer)

Janne Swaegers (tutor)

Laboratory of Biodiversity and Evolutionary Genomics

KU Leuven

Ch. Deberiotstraat 32, 3000 Leuven

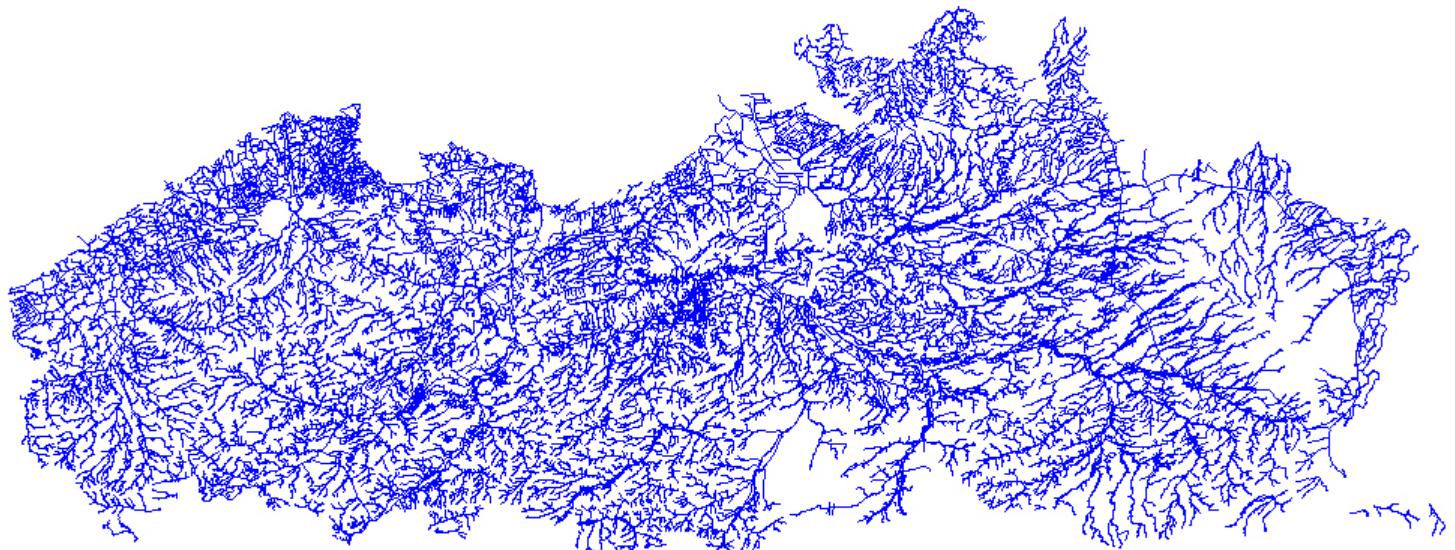
Phone: 016 32 39 72,

E-mail: filip.volckaert@kuleuven.be; janne.swaegers@kuleuven.be;

URL: <http://bio.kuleuven.be/eeb/lbeg>

Introduction

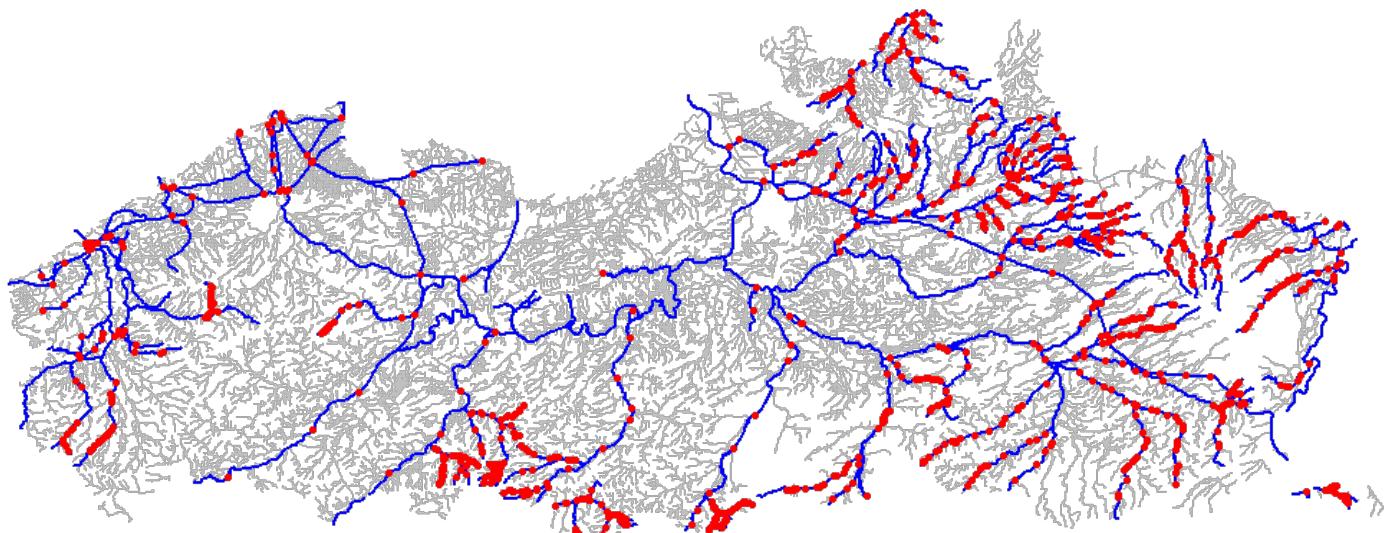
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



Map of river systems in Flanders

Introduction

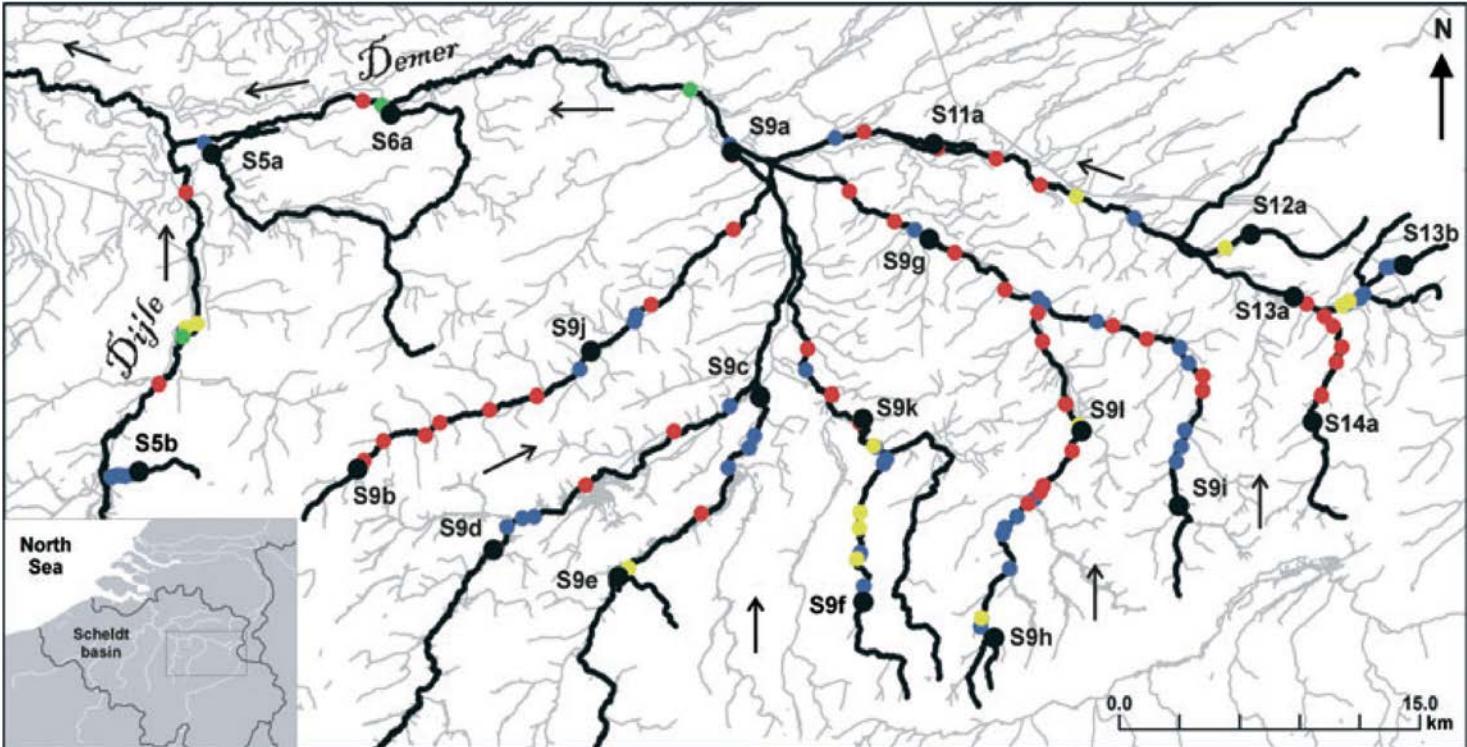
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



Map of barriers of all sorts in Flemish rivers

Introduction

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



● mills ● weirs ● tunnels ● sluices

Map of barriers in the Demer basin

Introduction

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

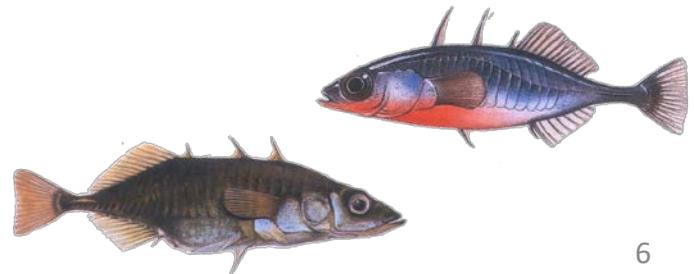


A wide diversity of barriers:
sluices, culverts, water mills and weirs

Introduction

Barriers have impact ...

- **introduction** - directly affect the reproductive success of fish populations
- **aims & tools** - may affect the global fitness of fish populations on the long term
- **geographical connectivity** ⇒ ecological (!) and evolutionary (?) consequences
- **genetics**
- **correlation analyses** **Sticklebacks have some advantages ...**
 - no stocking; preserved population structure
 - most common species
- **conclusion**

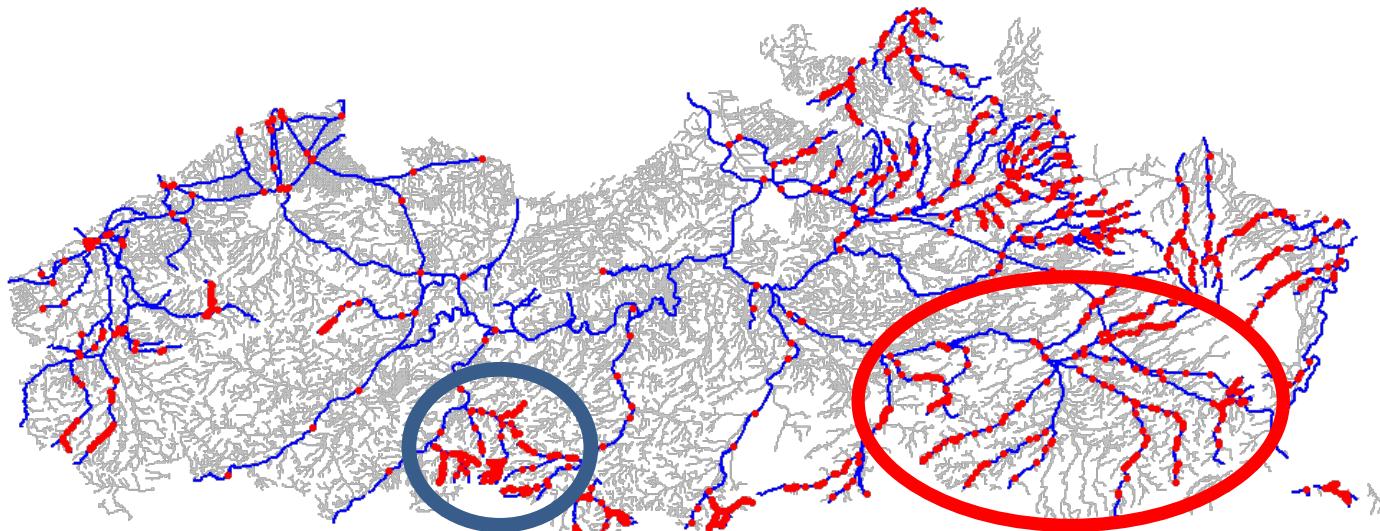


Introduction

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

Focus on two basins:

- Demer
- Zwalm



Introduction

Modeling genetic connectivity in sticklebacks as a guideline for river restoration

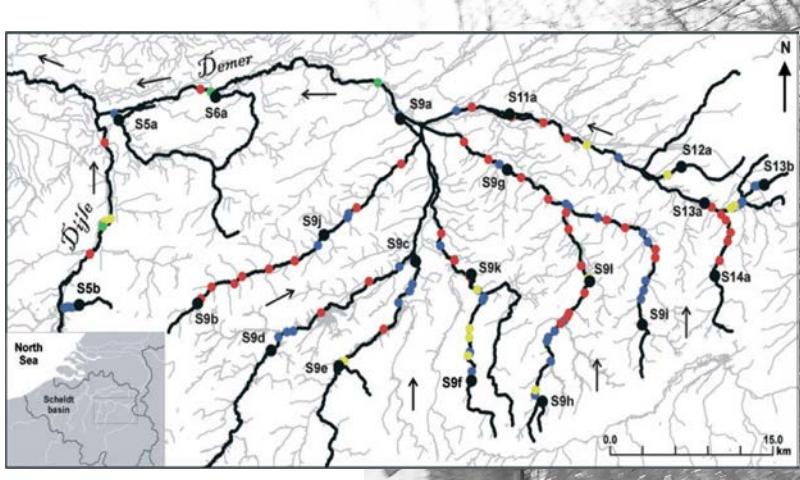
Joost A. M. Raeymaekers,¹ Gregory E. Maes,¹ Sarah Geldof,¹ Ingrid Hontis,¹ Kris Nackaerts² and Filip A. M. Volckaert¹

¹ Laboratory for Animal Biodiversity and Systematics, Katholieke Universiteit Leuven, Leuven, Belgium

² InterGraph Belgium NV, Tennessee House, Riverside Business Park, Brussels, Belgium



Raeymaekers et al *Evolutionary Applications* 2008



Journal of Applied Ecology 2009, **46**, 562–571

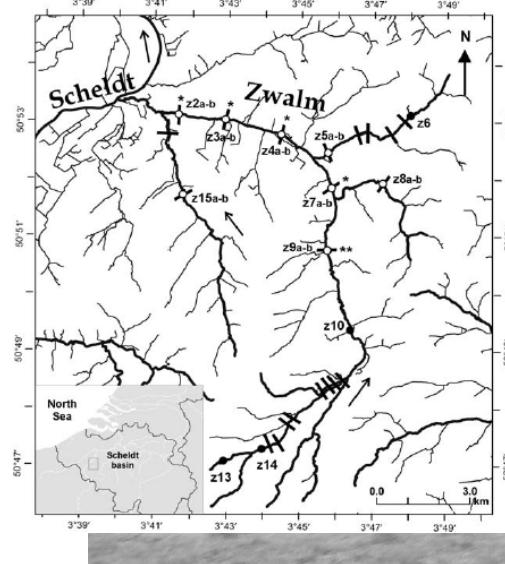
doi: 10.1111/j.1365-2664.2009.01652.x

Guidelines for restoring connectivity around water mills: a population genetic approach to the management of riverine fish

Joost A. M. Raeymaekers^{1*}, Dries Raeymaekers², Itsuro Koizumi³, Sarah Geldof¹ and Filip A. M. Volckaert¹



Raeymaekers et al *Journal of Applied Ecology* 2009



Modeling genetic connectivity in sticklebacks as a guideline for river restoration

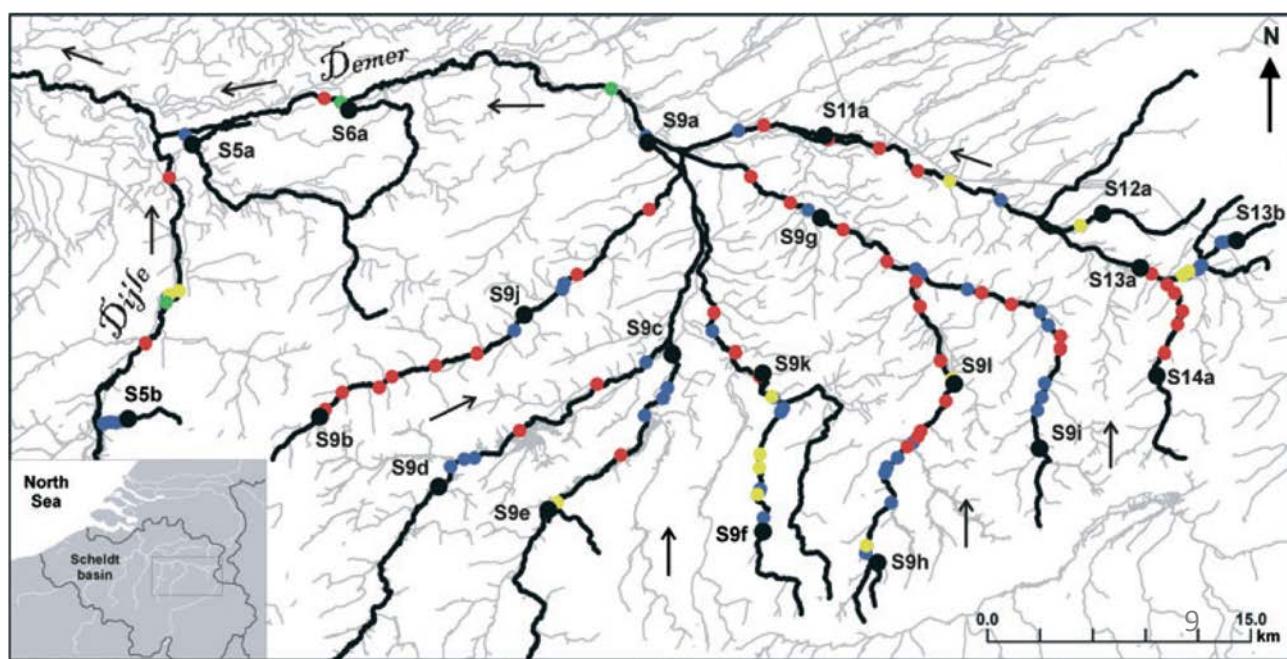
Joost A. M. Raeymaekers,¹ Gregory E. Maes,¹ Sarah Geldof,¹ Ingrid Hontis,¹ Kris Nackaerts² and Filip A. M. Volckaert¹

¹ Laboratory for Animal Biodiversity and Systematics, Katholieke Universiteit Leuven, Leuven, Belgium

² InterGraph Belgium NV, Tennessee House, Riverside Business Park, Brussels, Belgium



Raeymaekers et al EVA 2008



Aims & Tools

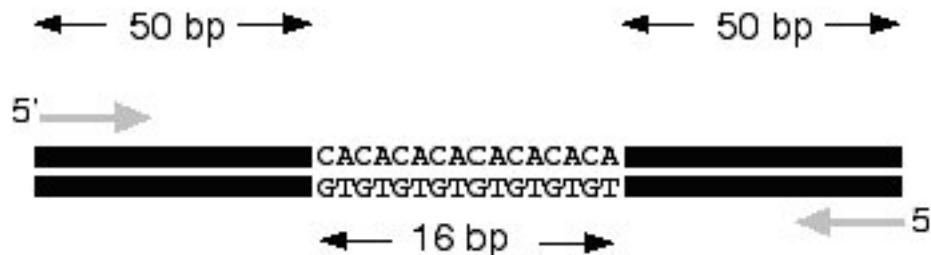
1) What is the connectivity between stickleback populations?

- **introduction**
- **aims & tools**
- **geographical connectivity**
- **genetics**
- **correlation analyses**
- **conclusion**

⇒ (telemetry, capture/mark/recapture)

⇒ Genetic relatedness as a measure of connectivity

⇒ Tool: microsatellite markers:



High mutation rate and highly polymorphic

Aims & Tools

- **introduction**
 - **aims & tools**
 - **geographical connectivity**
 - **genetics**
 - **correlation analyses**
 - **conclusion**
- Genetic information
- ⇒ genetic variability: allelic richness & heterozygosity
- ⇒ inbreeding; bottlenecks
- ⇒ genetic differentiation
- ! gene flow \neq dispersal/number of migrants!

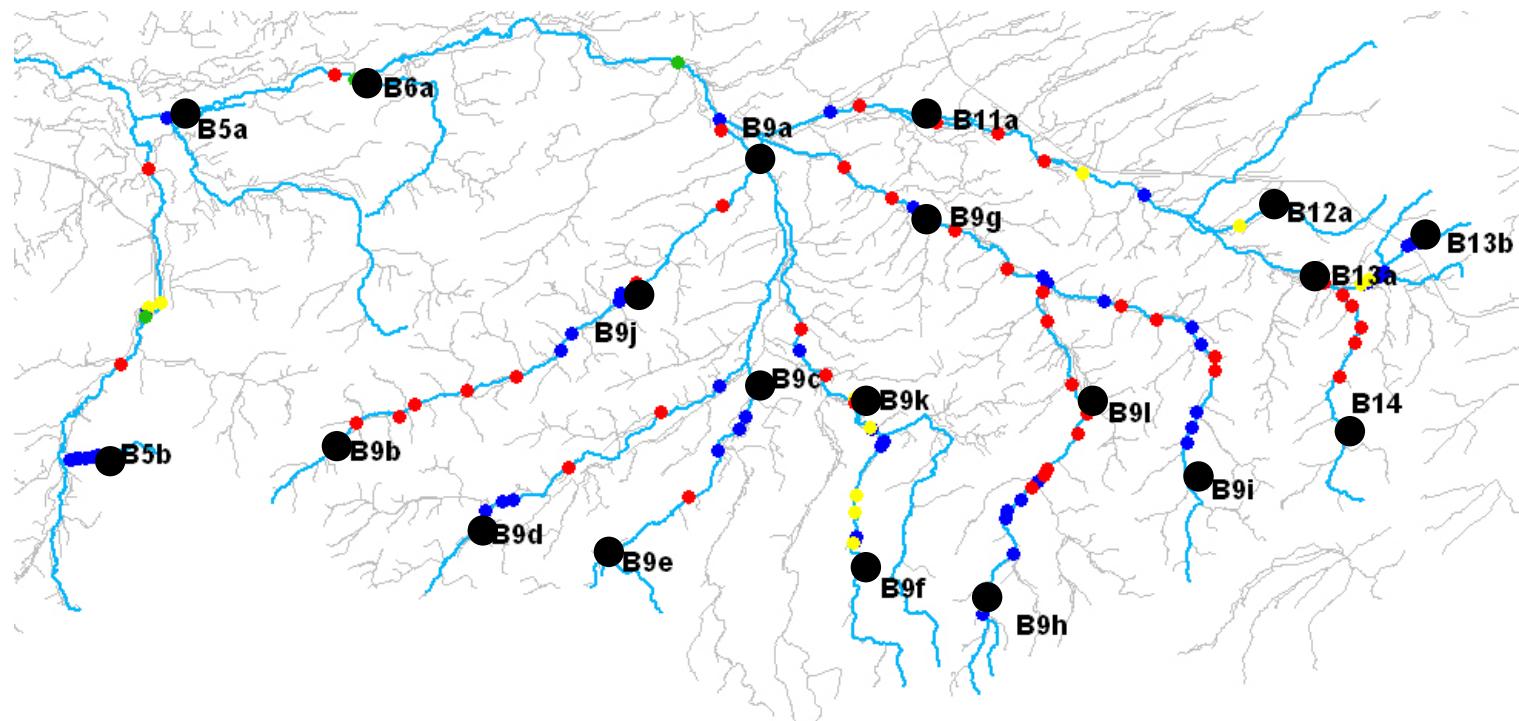
Aims & Tools

2) Which geographical features affect connectivity?

- **introduction**
- **aims & tools**
- **geographical connectivity**
 - distance
 - presence of migration barriers
 - age/type/severity of migration barriers
- **genetics**
- **correlation analyses**
 - ⇒ Different geographical scenarios
 - ⇒ Geographical information system (GIS)
- **conclusion**

Geographical connectivity

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

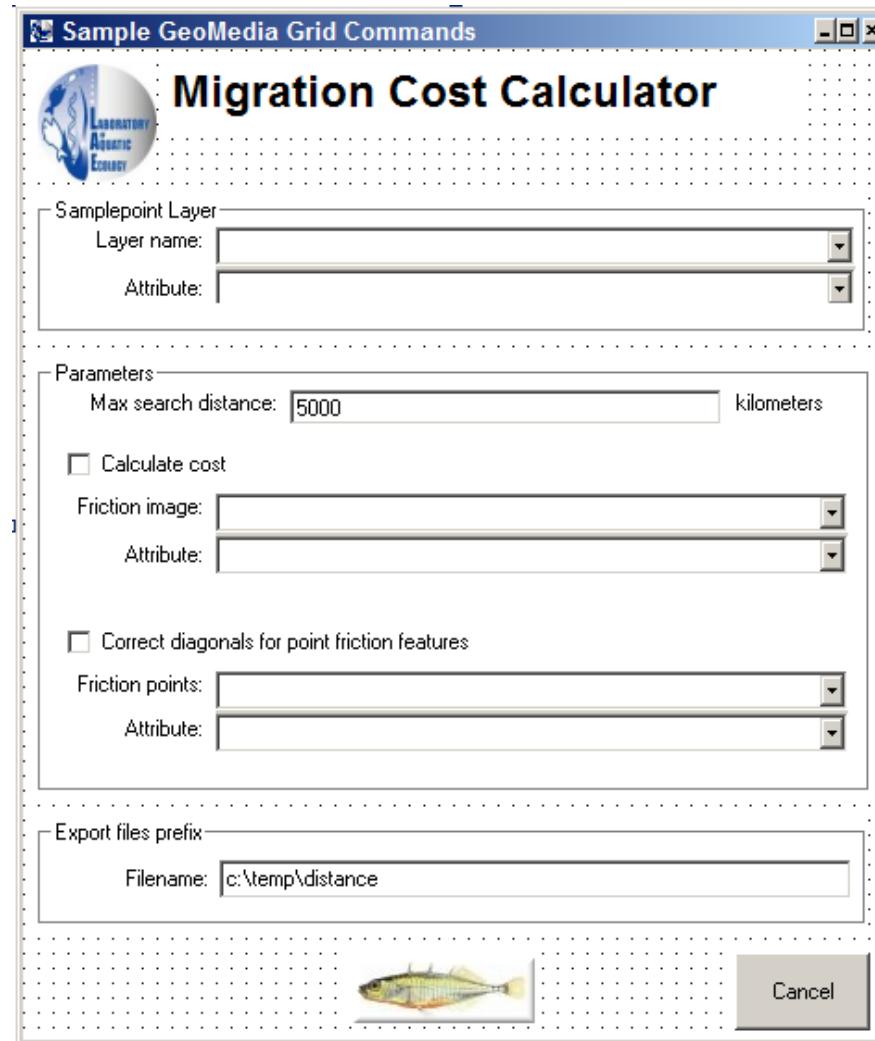


⇒ 21 populations

⇒ 210 pairwise connectivities → GIS

Geographical connectivity

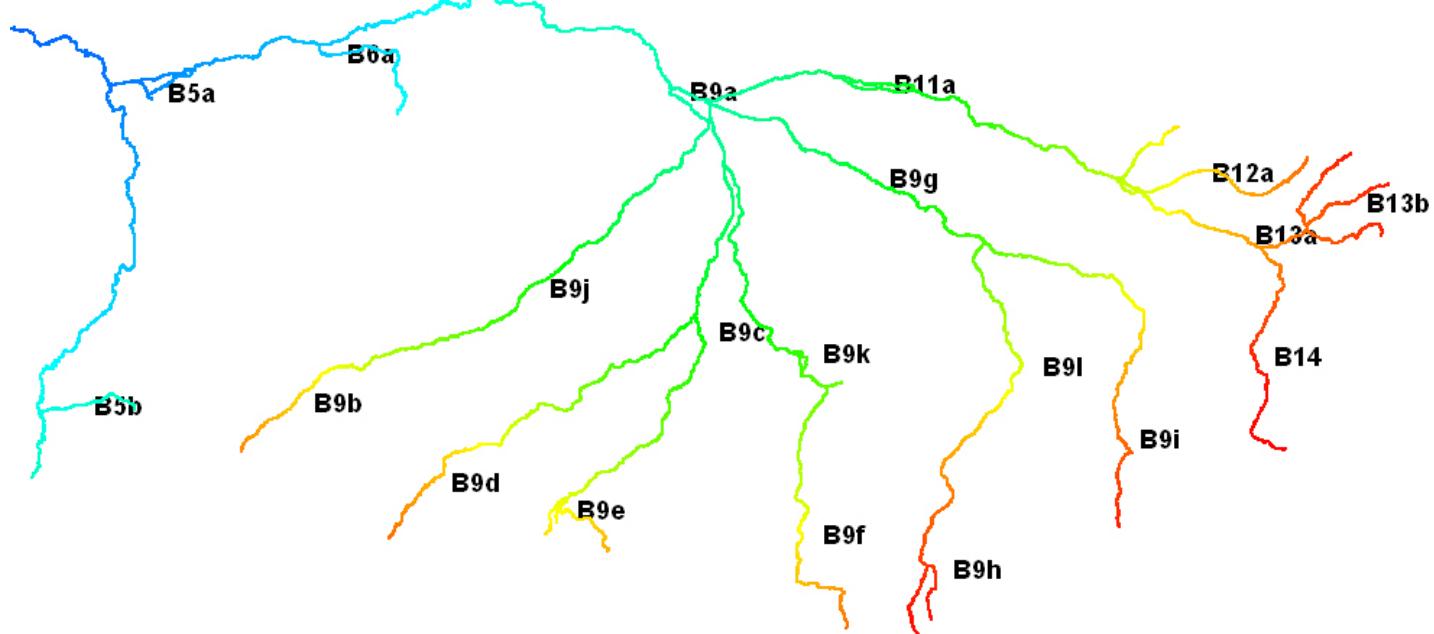
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



GeoMedia 5.1

Geographical connectivity

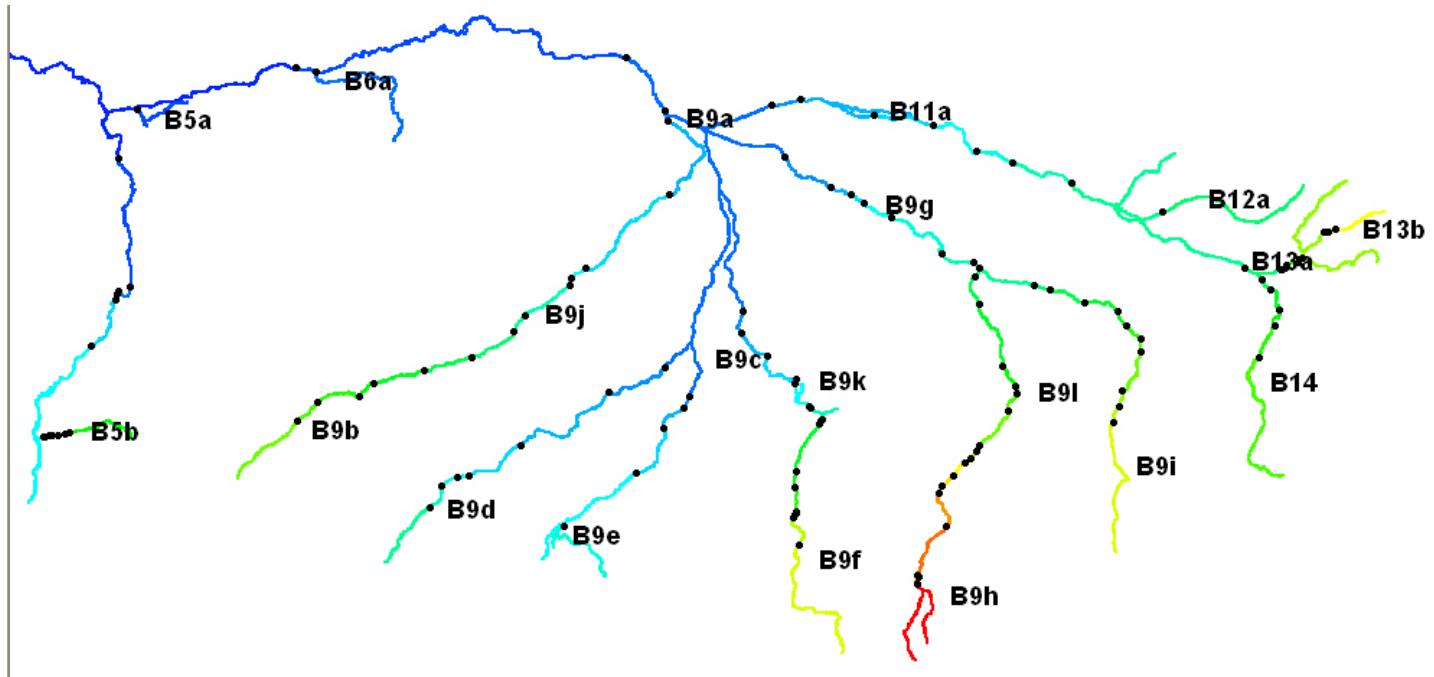
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



distance (115 km)

Geographical connectivity

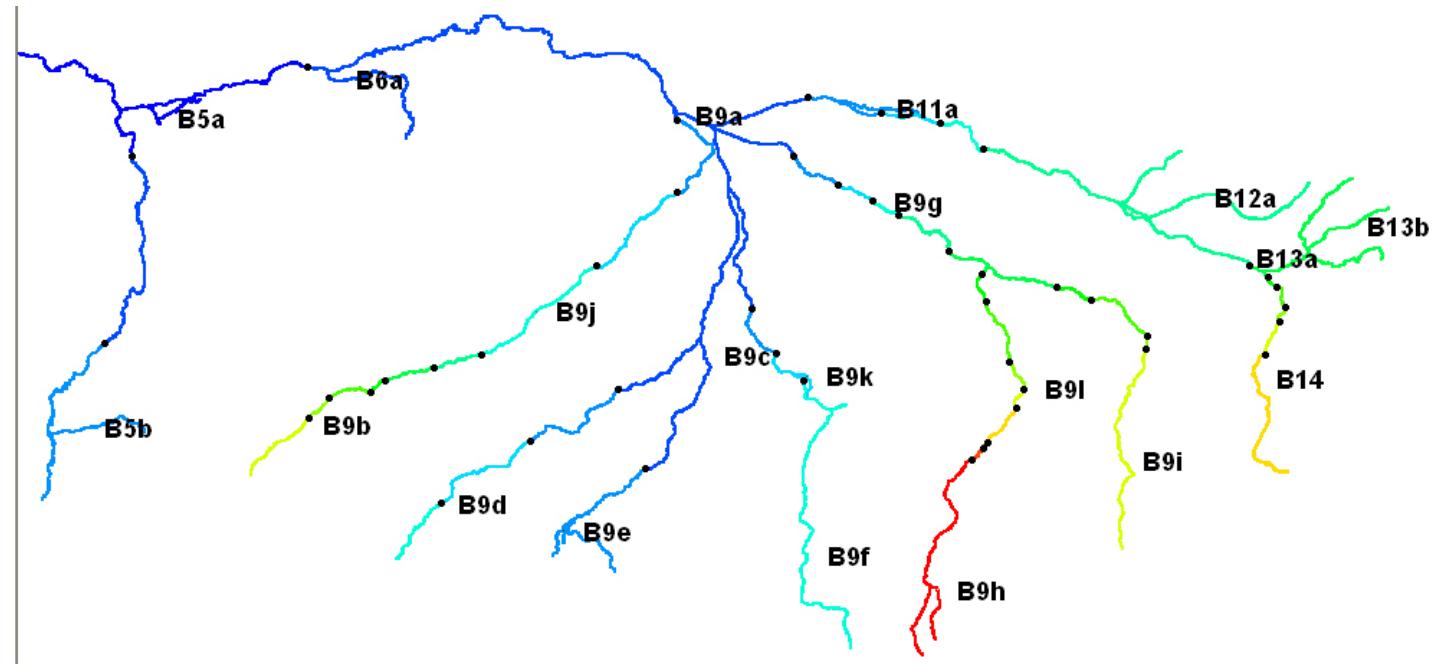
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



total number of barriers (28)

Geographical connectivity

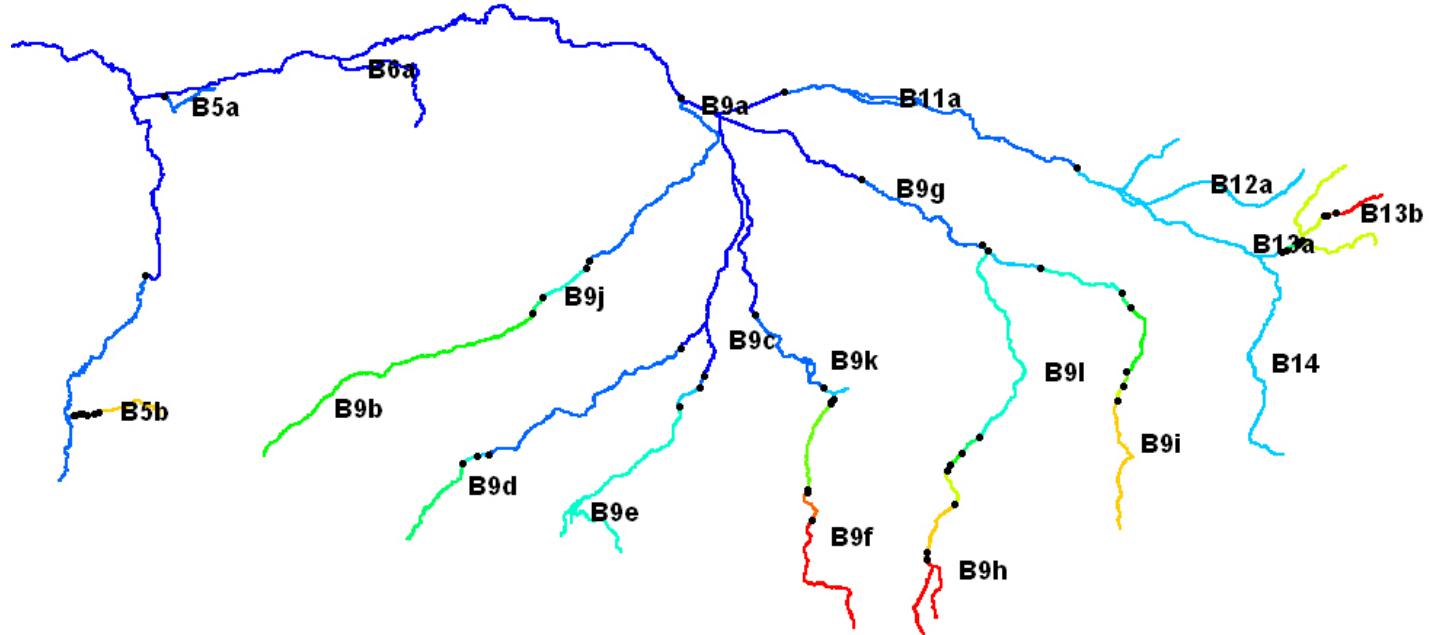
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



number of mills (14)

Geographical connectivity

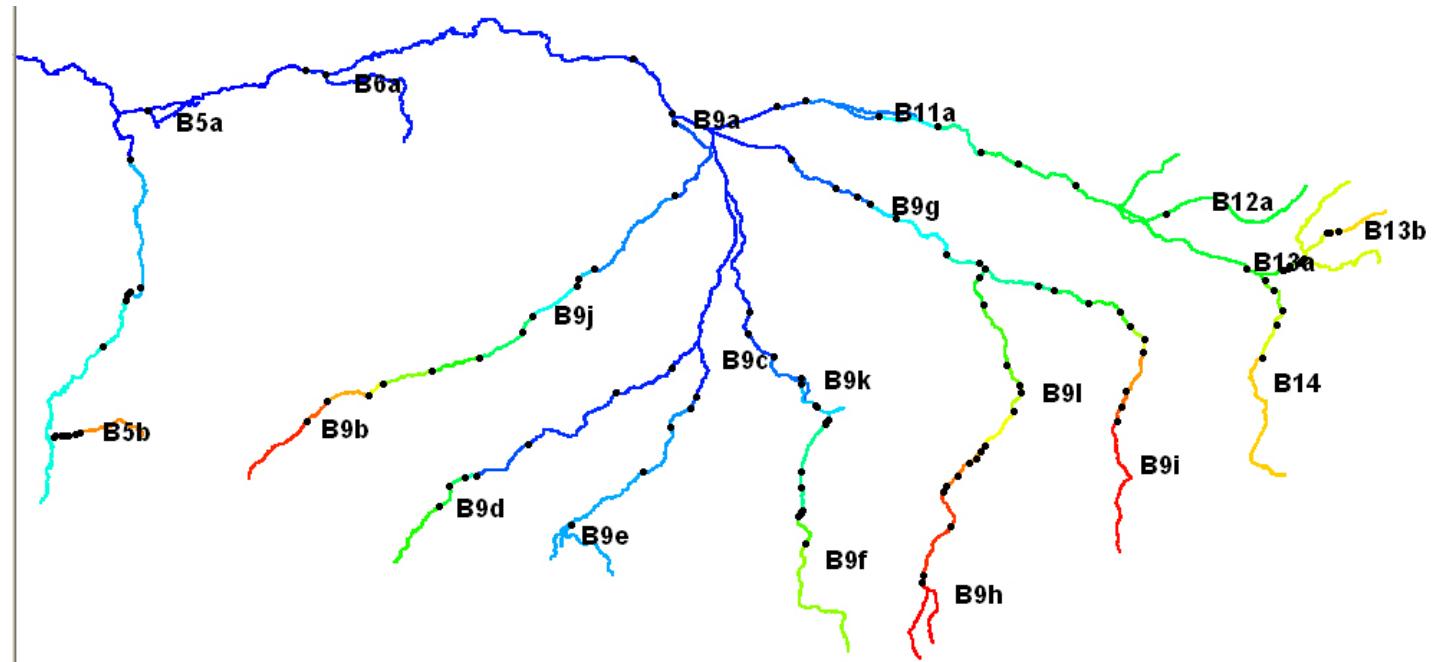
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



number of weirs (10)

Geographical connectivity

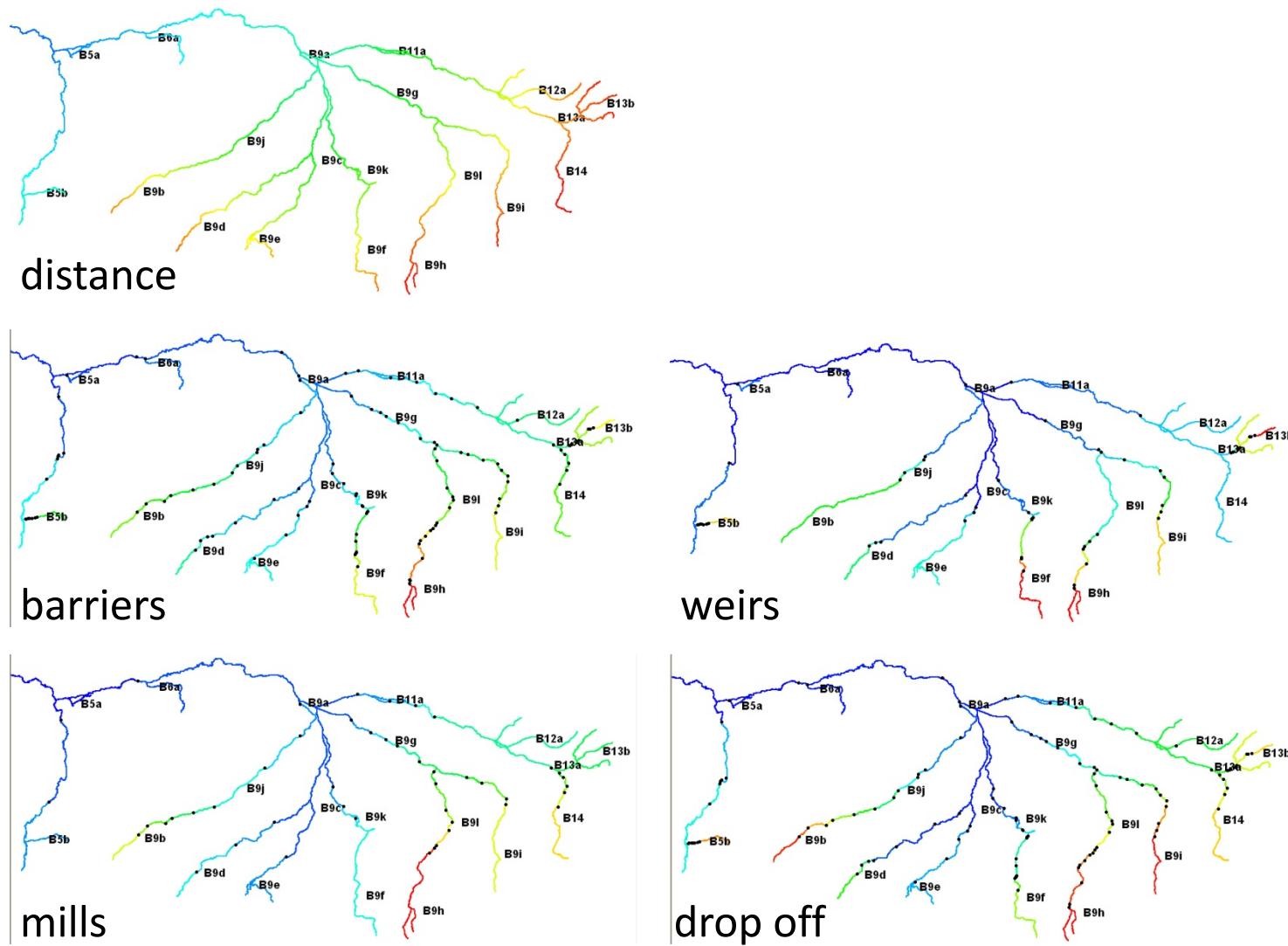
- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



vertical drop off (16.7 m)

Geographical connectivity

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



A scenario?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

- postglacial colonization of freshwater habitats by anadromous sticklebacks with reduction of the genetic variability
(Raeymaekers *et al.* 2005)
 - further reduction of the genetic variability in upstream populations
 - construction of mills (? 500 years); construction of weirs and tunnels (? 100 years)
 - genetic drift: upstream alleles only can get lost
 - genetic differentiation
- ! Natural + induced reduction of genetic variability

Genetic connectivity

Genetic variability

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

Table 1. Characteristics of 21 sampling locations of three-spined stickleback (*Gasterosteus aculeatus*) populations in Belgium, genotyped at six microsatellite loci.

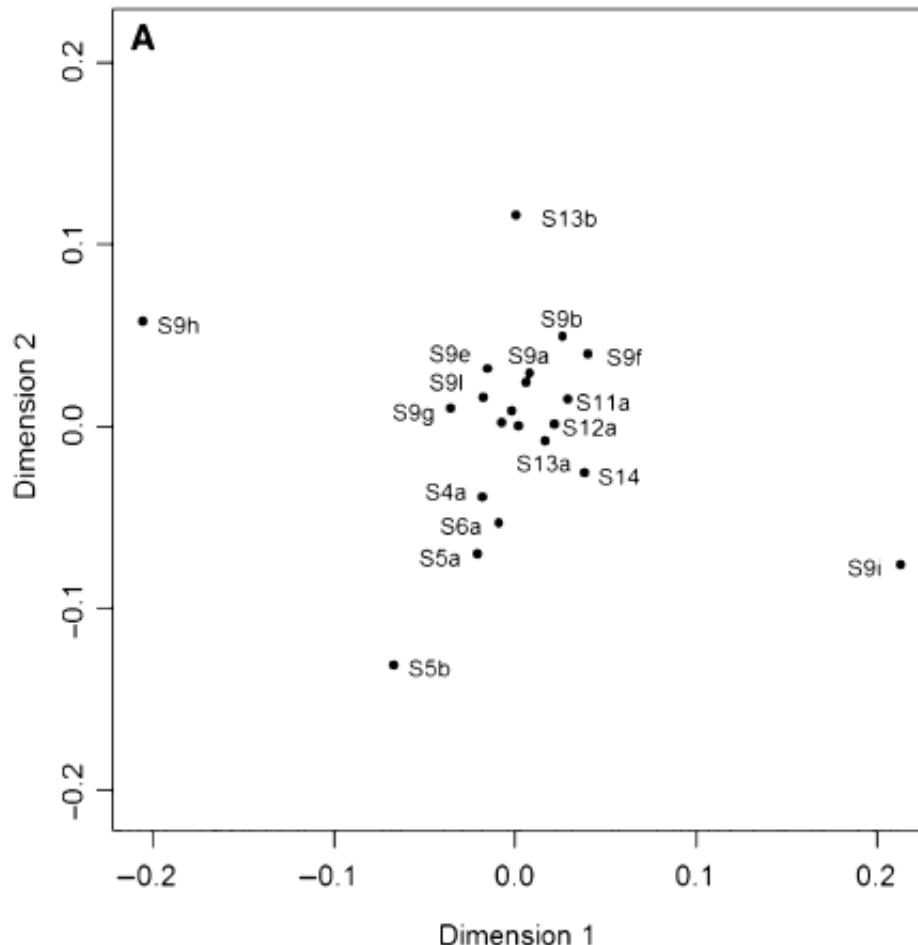
Population	Code	Basin	Latitude	Longitude	H_E	H_O	AR
Mechelen	S4a	Nete	51°03.734'	4°27.588'	0.83	0.83	10.34
Werchter	S5a	Dijle	50°57.806'	4°43.276'	0.78	0.77	9.46
Vaalbeek	S5b	Dijle	50°49.466'	4°40.105'	0.65	0.63	6.22
Aarschot	S6a	Demer	50°58.831'	4°50.898'	0.79	0.76	8.20
Zelem	S9a	Demer	50°57.761'	5°05.431'	0.73	0.72	7.66
Boutersem	S9b	Demer	50°49.506'	4°49.405'	0.66	0.68	5.78
Zoutleeuw	S9c	Demer	50°51.310'	5°6.531'	0.78	0.75	8.89
Hoegaarden	S9d	Demer	50°47.355'	4°55.150'	0.79	0.77	9.471
Landen	S9e	Demer	50°46.613'	5°00.441'	0.74	0.76	6.50
Gingelom	S9f	Demer	50°45.882'	5°10.842'	0.69	0.74	5.35
Stevoort	S9g	Demer	50°55.393'	5°13.823'	0.78	0.80	7.02
Mechelen-Bovelingen	S9h	Demer	50°44.904'	5°16.350'	0.55	0.56	3.67
Borgloon	S9i	Demer	50°48.318'	5°24.287'	0.44	0.43	3.46
Kortenaken	S9j	Demer	50°52.513'	4°59.953'	0.75	0.74	7.74
St-Truiden	S9k	Demer	50°50.702'	5°10.900'	0.80	0.81	7.32
Wellen	S9l	Demer	50°50.312'	5°20.196'	0.77	0.78	7.28
Kerm	S11a	Demer	50°57.900'	5°14.043'	0.73	0.72	7.56
Diepenbeek	S12a	Demer	50°55.409'	5°27.509'	0.79	0.73	7.59
Bilzen	S13a	Demer	50°53.749'	5°29.290'	0.81	0.83	7.73
Zutendaal	S13b	Demer	50°54.539'	5°34.006'	0.64	0.66	4.37
Alt-Hoeselt	S14	Demer	50°50.479'	5°30.031'	0.79	0.71	7.39

H_E , expected (unbiased) heterozygosity; H_O , observed heterozygosity; AR, allelic richness standardized for 18 diploid individuals.

Genetic connectivity

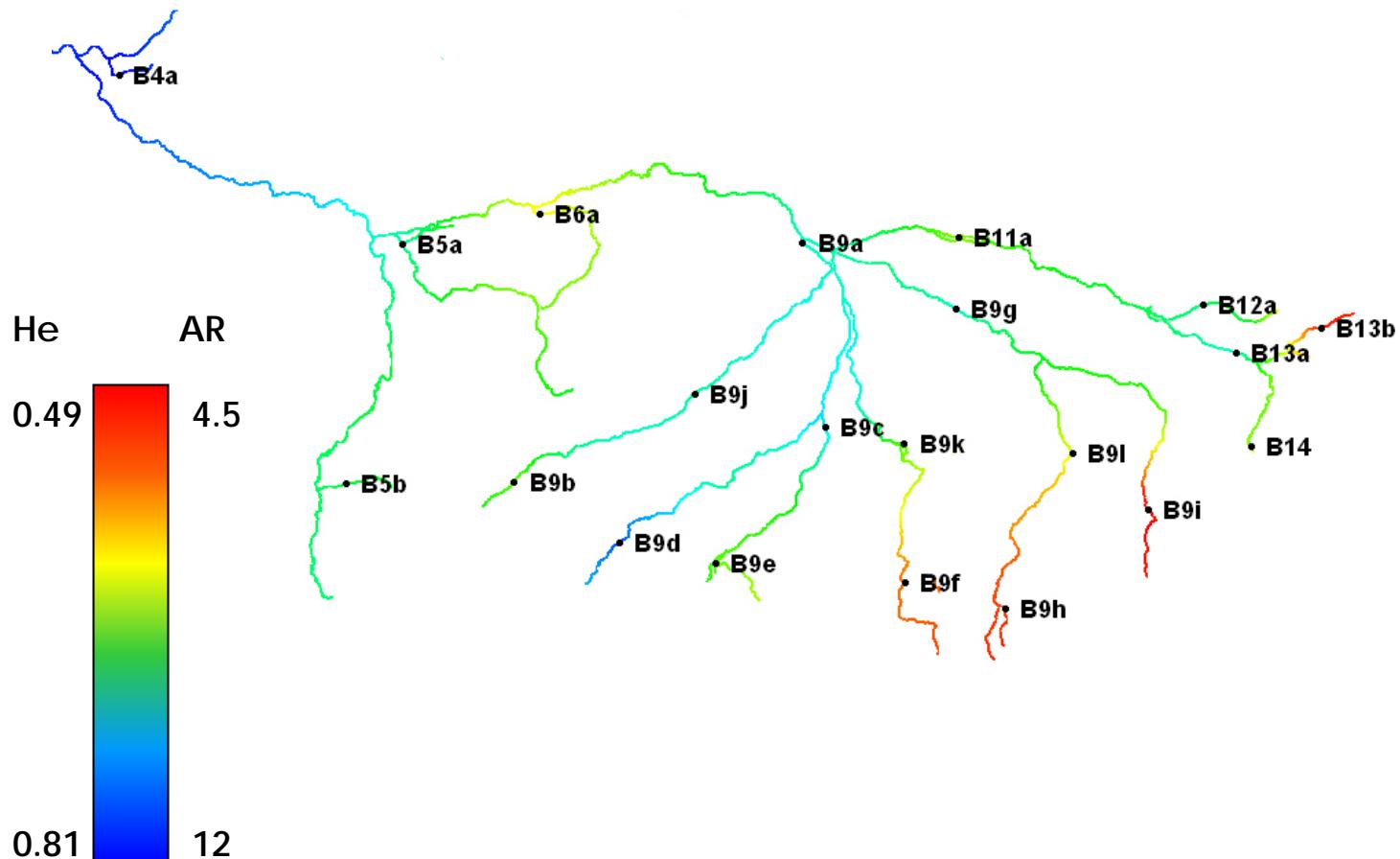
Genetic differentiation

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



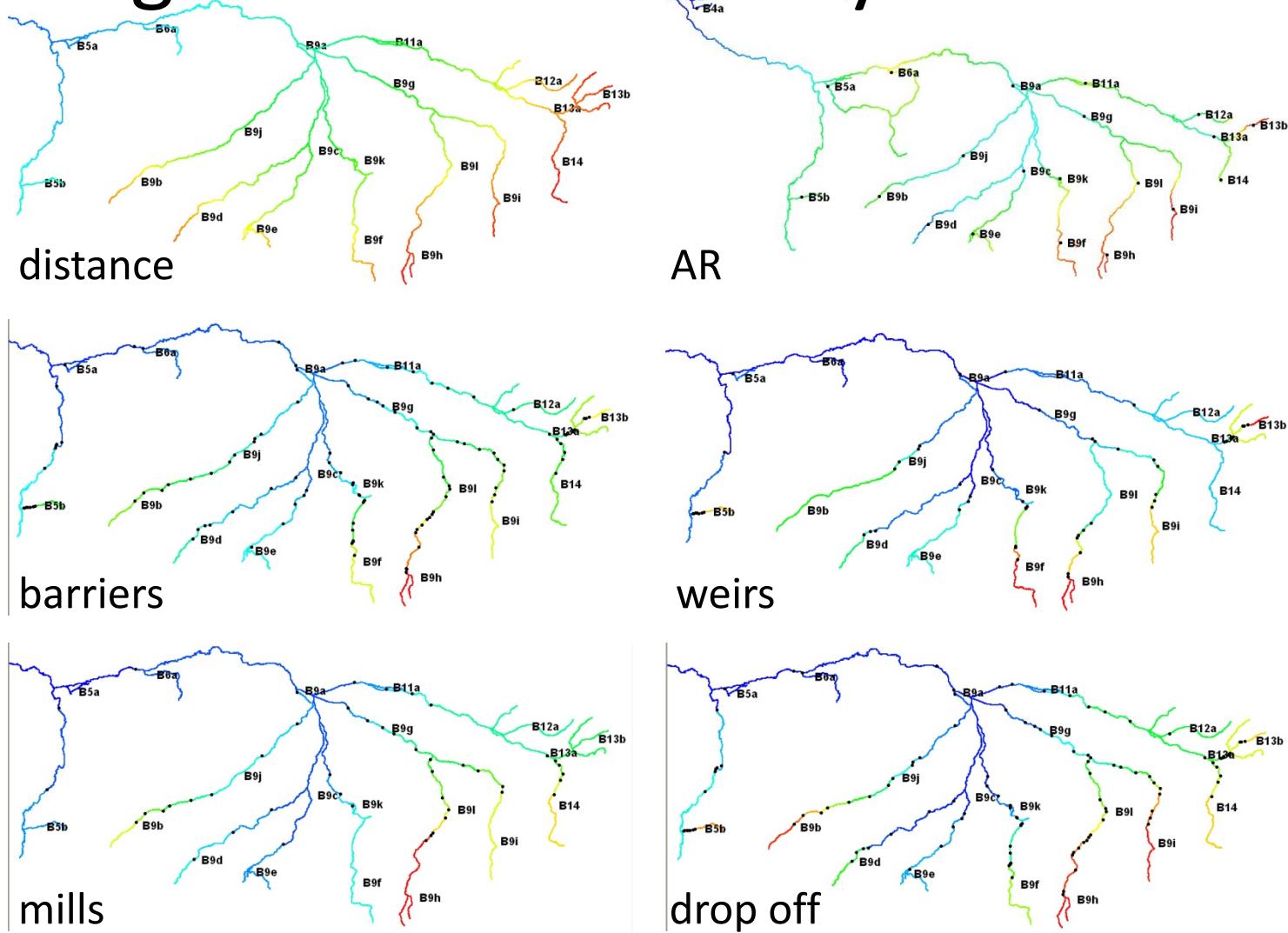
Linking geographic connectivity with genetic connectivity?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



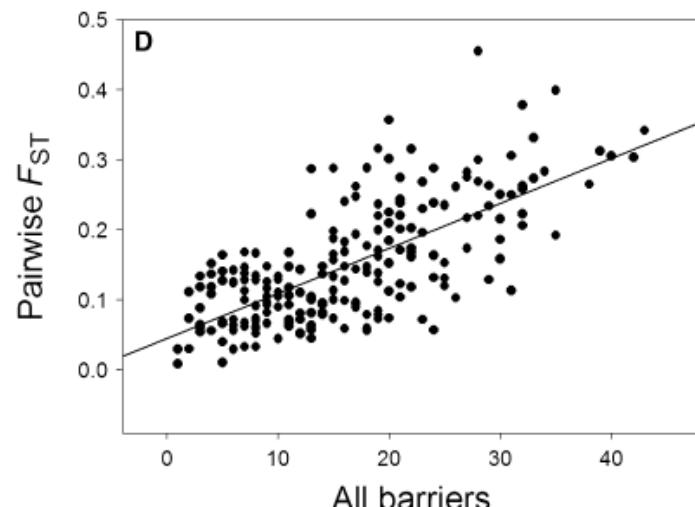
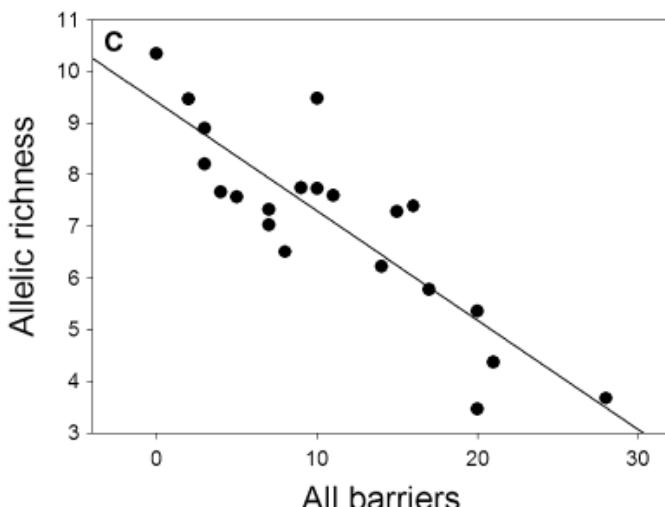
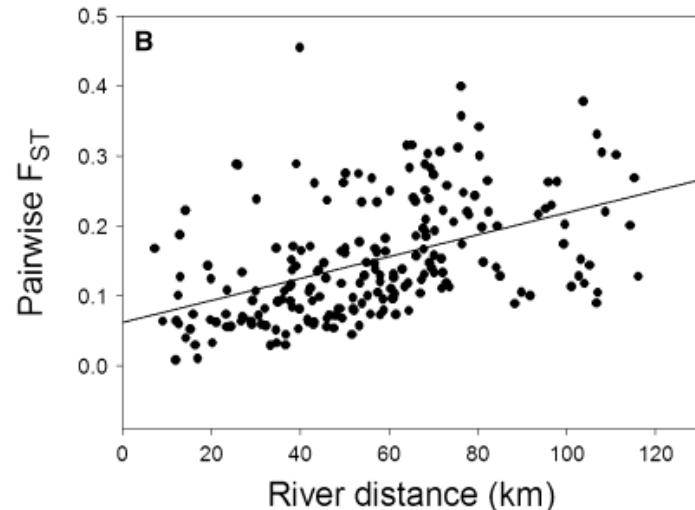
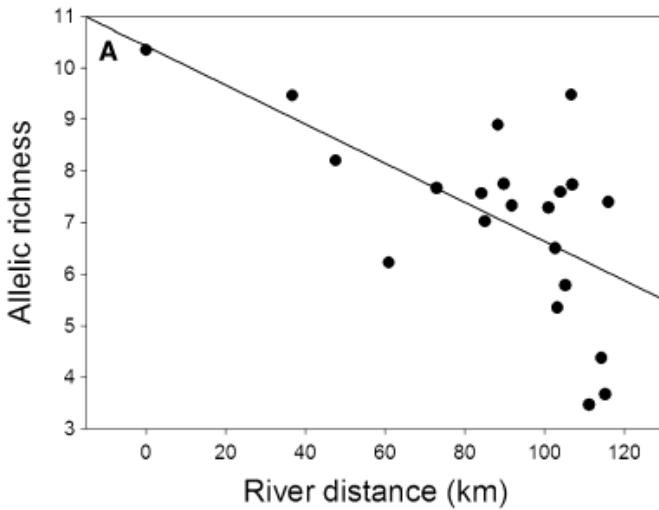
Linking geographic connectivity with genetic connectivity?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



Linking geographic connectivity with genetic connectivity?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion



Linking geographic connectivity with genetic connectivity?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

- We need to make sure that the explanatory variables are not too much intercorrelated. To do so, we check correlation coefficients and calculate the “variance inflation factor” (see publication and R script).

Table 4. Correlations among the explanatory variables used in the landscape models from Table 3. (A) Pearson correlations among the geographical features from Table 3A. (B) Mantel matrix correlations among the pairwise geographical features from Table 3B.

(A)	All barriers	$\text{Log}_{10}(\text{habitat width})$	$\text{Log}_{10}(\text{watershed position})$	
All barriers	–	–0.5701	–0.6858	
$\text{Log}_{10}(\text{habitat width})$	0.007	–	0.5592	
$\text{Log}_{10}(\text{watershed position})$	0.001	0.008	–	
(B)†	All barriers	$\text{Log}_{10}(\text{habitat width})$	$\text{Log}_{10}(\text{watershed position})$	River distance
All barriers	–	–0.5553	–0.5447	0.5170
$\text{Log}_{10}(\text{habitat width})$	0.0008	–	0.4261	–0.4390
$\text{Log}_{10}(\text{watershed position})$	0.0004	0.02310	–	0.0146
River distance	0.0004	0.0024	0.4502	–

Correlation coefficients are given above the diagonal, P -values below the diagonal. Significant P -values are in bold.

† P -values are given after 10 000 randomizations.

Linking geographic connectivity with genetic connectivity?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

- Finally, we can build and compare landscape genetic models

Table 3. (A) Multiple regression analysis of number of barriers, $\log_{10}(\text{habitat width})$ and $\log_{10}(\text{watershed position})$ on allelic richness (AR; $R^2 = 0.80$; adjusted $R^2 = 0.76$; $F_{3,17} = 22.30$; $P < 0.0001$). (B) Nonparametric multiple regression of number of barriers, $\log_{10}(\text{pairwise average habitat width})$, $\log_{10}(\text{pairwise average watershed position})$ and river distance on pairwise F_{ST} ($R^2 = 0.54$; adjusted $R^2 = 0.53$).

Effect	(A) AR				(B) †Pairwise F_{ST}	
	df	MS	F	P-value	β	P-value
Intercept	1	44.60	56.51	<0.0001	–	–
All barriers	1	11.98	15.17	0.0012	0.00490	0.0001
$\log_{10}(\text{habitat width})$	1	0.39	0.49	0.4935	-0.09611	0.0048
$\log_{10}(\text{watershed position})$	1	2.51	3.18	0.0926	0.00551	0.7742
River distance	–	–	–	–	0.00027	0.3776
Error	17	0.79				

Significant P-values are in bold.

†P-values are given after 10 000 randomizations.

Linking geographic connectivity with genetic connectivity?

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

- For model comparison, we use AIC, Akaike's Information Criterium
- AIC is calculated from the least-squares regressions

$$AIC = 2K + n \times \ln(RSS/n)$$

where K is the number of parameters, n is the number of populations, and RSS is the residual sum of squares

- The smaller AIC, the better the model
- Models with differences in AIC values > 2 are considered “substantially different”

Conclusion

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

- We have to compare highly congruent alternative scenarios
- These scenarios are to be tested with strongly correlated and dependent geographical information
- Scenarios are also confounded with the evolutionary history of sticklebacks (anadromous ancestry + differences in N_e)
- Landscape genetic models help us to deal with these issues statistically

Conclusion

- introduction
- aims & tools
- geographical connectivity
- genetics
- correlation analyses
- conclusion

- the number of migration barriers (rather than other factors) affects levels of genetic variability and genetic differentiation
- different types of migration barriers have a different impact on genetic differentiation
 - Barriers underlie migration-drift equilibrium more than geographical distance
 - Barrier-induced effects likely apply to other riverine fish species as well
- Humans can have a dramatic impact on the viability of riverine fishes

Conclusion

- **introduction**
- **aims & tools**
- **geographical connectivity**
- **genetics**
- **correlation analyses**
- **conclusion**

- Landscape genetic models obviously are also constructed for terrestrial organisms, but tend to explain a lower proportion of genetic variation than “riverscape” genetic models
- The sampling design need to be considered carefully in the planning phase of a landscape genetic study

* * *

Literature

- Manel, S., M. K. Schwartz, G. Luikart, and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* 18:189–197.
- Raeymaekers, J.A.M., Maes, G.E., Geldof, S., Hontis, I., Nackaerts, K., and Volckaert, F.A.M. (2008). Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evolutionary Applications* 1, 475-488.
- Raeymaekers, J.A.M., Raeymaekers, D., Koizumi, I., Geldof, S., and Volckaert, F.A.M. (2009). Guidelines for restoring connectivity around water mills: a population genetic approach to the management of riverine fish. *Journal of Applied Ecology* 46, 562-571.

Task #5:

You are asked to prepare landscape genetic models for allelic richness and pairwise F_{ST} .

1. Which statistical analyses do you propose to investigate barrier impact? Do you think there is a straightforward way to measure barrier impact? Why or why not?
2. Please provide me a table with the output of the statistical analysis, choose the best model and interpret the results in a few lines. What do the results actually mean?
3. Do you have suggestions for improving the study?