1. Genetic analysis of brown trout in the headwater tributary streams of large northern river basins

Authors: Tuomas Leinonen[1], Ari Huusko[2] and Erkki Jokikokko[3]

[1]Natural Resources Institute Finland (Luke), Latokartanonkaari 9, 0790 Helsinki

[2]Natural Resources Institute Finland (Luke), Manamansalontie 90, 88300 Paltamo

[3]Natural Resources Institute Finland (Luke), Laivurintie 6, 94450 Keminmaa

1.1. Introduction

The main aim of the brown trout genetic analyses was to characterize the genetic structure and the extent of genetic diversity of the brown trout populations in the main River Kemijoki basin. River Kemijoki is a regulated river with five power stations located in the lower reach of the river between the Bothnian Bay and the town of Rovaniemi. There are also several other power stations in the middle reach of the river and upper river reaches of the river system, especially in the River Kitinen basin (see https://www.kemijoki.fi/toimintamme/voimalaitokset ja -tuotanto.html). There is a fishway in the Isohaara power station located nearest to the Bothnian Bay. However, the upriver fish migration has been blocked off since 1946. In addition to River Kemijoki basin, brown trout samples were collected from the free-flowing River Tornionjoki and River Kalix, and the regulated River Luleå basins to reveal potential genetic differences in stream-specific brown trout populations between large river basins with differences in upriver migration possibilities of salmonids.

In the analyses, it was first determined whether the sampled sites contained genetically discreet brown trout populations. This included analyses of genetic mixing among the sampling sites and the effect of hatchery stocks. Second, the extent of genetic variation in each brown trout population was quantified to provide an estimate of viability of each population. This included the estimation of genetic diversity and effective population sizes in each population as well as relatedness among individuals at each sampling site.

All the sampled populations were first analyzed separately and then samples from each river basin were combined to compare genetic structure and variation among different river basins.

1.2. Material and Method

2. Samples and sampling areas

A total of 2255 brown trout were genotyped in 2022 for the analysis of genetic diversity and differentiation of the brown trout populations in the target river basins. Genotypes were obtained for 2244 samples. In addition, 350 trout samples from earlier genetic analyses done in the joint molecular genetic laboratory of the University of Helsinki and Natural Resources Institute Finland (Luke) were used in the analyses as reference samples. In total, 2594 trout samples genotyped at 16 microsatellite loci were used in the analyses. The samples comprised 17 groups: 9 river basins (Fig. 1) and 8 sets of reference samples (Table 1). Reference sample sets were 1) River Tornionjoki (a known sea-migrating brown trout population), 2) River Kemijoki (samples consisting of brown trout ascending into the river for spawning, sampled below and from the fishway at Isohaara power station (i.e. sea-migrating brown trout, originally of hatchery origin)), 3) River Iijoki (sea-migrating brown trout, a hatchery population), 4) River Kemihaara I (samples from a hatchery population of stream resident population collected in 2014, the origin of the brood stock being from the River Kemihaara basin), 5) River Kemihaara II (stream resident population, similar to 4) but sampled in 2021 from the hatchery brood stock), 6) Rautalampi water course (hatchery population of lake-run (adfluvial) brown trout, used widely in stocking in lakes and rivers in Finland), 7) River Oulujoki water course (a hatchery population of lake-run (adfluvial) brown trout), 8) Lake Inari basin (a hatchery population of lake-run (adfluvial) brown trout, a combined sample set of River Juutuanjoki, Ivalojoki and Siuttajoki populations, used in stocking in Lake Inari and in Porttipahta, a large reservoir located in the upper reach of River Kitinen basin).

3. Statistical analyses

To get an overview of the genetic diversity among the brown trout populations and river basins, the data were first analysed with a discriminant analysis of principal components (DAPC)(Jombart et al., 2010) using the adegenet package (Jombart, 2008) for R (R Core Team, 2019). The number of distinct genetic groups was determined by the k-means clustering method implemented in the DAPC analysis with the adegenet package. DAPC was also done separately for each river basin. Genetic structure and the effect of hatchery stocking of the brown trout populations were also tested with STRUCTURE software using the admixture model (Pritchard et al., 2000). Because the number of samples has a strong effect on the results of the clustering analyses, a maximum of 20 randomly selected individuals from each sampling site were included in the analysis. Genetic structure was tested for up to 20 genetic clusters (k = 20). For each k from 1 to 20, ten replicates of 500,000 iterations with a burn-in of 100,000 iterations were run, and the results from the replicates were combined and reported using the CLUMPAK server (Kopelman et al., 2015). The extent of mixing among different sampling streams was also tested with a reassignment test in adegenet, where each individual fish is reassigned to its original sampling stream. The probability of reassignment is higher for clear-cut genetic clusters, while low probability of reassignment is an indication of gene flow and admixture among different populations (Jombart and Collins, 2017). Genetic diversity within each brown trout population and river basin was

analysed using diveRsity package (Keenan et al., 2013) for R (R Core Team, 2019). Two different indices of genetic differentiation among the brown trout populations and among different river basins were calculated: DA (Nei et al., 1983) with populations software (Langella, 2000), and FST (Wright, 1949) with the diveRsity package for R. DA provides a 'raw' genetic distance, while FST is a fixation index that scales the among population genetic differentiation with the genetic variation within populations.

Relatedness among individuals in each sampling stream was tested with demerelate package for R (Kraemer and Gerlach, 2017). The effective population size (Ne) for each brown trout population was estimated linkage disequilibrium method as implemented in NeEstimator v2 (Do et al., 2014). Confidence intervals for the Ne estimates were calculated by jacknifing on samples and rare alleles with a frequency lower than 0.02 were screened out (Waples and Do, 2008).

3.1. Results

4. General patterns

Genetic differentiation among the brown trout samples from different river basins was generally relatively low (Fig. 2; Table 1), meaning that there have been no clear barriers for gene flow among the different river basins. The brown trout populations from different river basins formed three clear clusters (Fig. 2). Most of the samples fell into the first group: brown trout from the rivers flowing to the Bothnian Bay and the populations from the headwater tributary streams. In the second group were the samples from Lake Inari basin and River Kirakkajoki basin, which run east into the Barents Sea. The third group consisted of the hatchery populations of adfluvial brown trout.

Within each river basin, there were individual populations that diverged from the rest (Table 2.) However, many of the individual populations that showed clear differentiation from the rest had high relatedness among the samples consisting of several age groups, reaching even the level of relatedness among full siblings (r = 0.5; Table 3). This indicates that either the brown trout populations in these rivers or streams are extremely small and isolated and consist of close relatives, or a family group has been caught when sampling the location. Relatedness was quite high in many of the brown trout populations, and the effective population size (Ne; Box 1) was low, below the bare minimum threshold value (Ne > 50) for a population viable in short term (i.e., five generations; Table 3). This is quite common for small, isolated populations and is a cause of concern not only in terms of short-term persistence but also in terms of adaptation to changes in the environment (e.g., temperature, predation, food, water quality and discharge).

In the analysis of genetic structure and mixing of brown trout populations, the simplest model divided the brown trout samples into two genetically distinct groups (k=2; Fig 3A). The smaller group comprised the samples from rivers running into the east: River Kirakkajoki basin and Lake Inari reference samples, while the rest of the samples grouped into the second, larger group. The only exception was the stream Kuorajoki population from the River Luiro-Kitinen

basin, which was clearly similar to the populations in the Lake Inari group (Fig. 3A). The model with three genetic clusters (k=3; figure not shown) separated individual, genetically most differentiated populations (streams Haarainoja, Jäkälähaara, Lehto-oja 2, and Kulvakko-oja) from the rest. These populations are small and suffer from high number of relatives (Table 3). Presence of family groups in analyses of genetic structure is problematic and can produce strong patterns of genetic structure in cases where there is no actual genetic structure (Anderson & Dunham, 2008). Therefore, only gross patterns of genetic structure are discussed here.

Grouping the samples into four genetic clusters (k=4) showed that there are clear differences among different river basins. (Fig. 3B). Brown trout populations closer to the Bothnian Bay are closer to the sea-migrating brown trout (reference populations of River Kemijoki, Iijoki and Tornionjoki) in their genotype than brown trout populations further from the sea, which share a larger proportion of their genome with the resident brown trout (reference population of Kemihaara I and Kemihaara II; Fig. 3). The hatchery populations of adfluvial (lake-run) brown trout (reference populations of Rautalampi water course and River Oulujoki water course) showed a mixed similarity with the sea-migrating brown trout and the group of Lake Inari basin, but this is likely an artefact of the method used. In the analysis of genetic structure, the samples are forced into the set number of genetic groups, which can cause small groups (here the hatchery groups of adfluvial brown trout with a combined n = 40) to show similarity with the larger clusters. Considering the rest of the genetic analysis, this is the case for the adfluvial brown trout hatchery samples, which in all the other genetic analyses show clear genetic differentiation with the rest of the samples.

Population mixing and the lack of clear genetic differentiation was also evident in the dendrogram of genetic relationships among the populations (Fig. 4). Some of the populations from the same river basin clustered clearly together, but there were a lot of populations with an uncertain position in the dendrogram. This result can be partly explained by the low sample sizes at some of the sampling streams. Grouping the samples from each river basin together provided a clearer picture of the patterns of genetic differentiation among the analyzed brown trout populations. The hatchery samples of adfluvial brown trout (Rautalampi water course and River Oulujoki water course) grouped together with samples from the Lake Inari and River Kirakkajoki basins (Fig. 5). They were clearly different from the main River Kemijoki samples, which were divided into two groups, the samples from the upper river basins (River Kemihaara and Luiro-Kitinen basin) and the samples from the other river basins. The samples of potential sea-migrating brown trout (River Tornionjoki and River Kalix basins) did not fall into a clear group nor did the sea-migrating reference population of River Tornionjoki. The other sea trout reference samples from River Kemijoki and Iijoki formed a clear group together. The same pattern was evident in the DAPC of the river basins, which also gave an indication of the extent of genetic differentiation among the main River Kemijoki basin’s brown trout and the samples from other river basins (Fig. 2). The resident brown trout samples (Kemihaara I and II) grouped together with the samples from River Kemihaara and River Kitinen-Luiro basins (Fig. 5).

Analyses without the reference samples showed more clearly that the River Raudanjoki basin samples were the most clearly differentiated from each other as well as the samples from the other river basins, which might be explained by the high level of relatedness in the samples from this river basin (Table 3). The samples from the upper reaches of the main River Kemijoki basin (River Luiro-Kitinen and Kemihaara basins) were also clearly differentiated from the samples from the other river basins of River Kemijoki and Swedish river basins (Fig. 2B). There was less differentiation among the samples from the lower and middle reaches of the River Kemijoki and River Ounasjoki basins, which are closer to the sea-migrating brown trout samples (Fig. 2B). Moving further from the sea, connectivity to the sea-migrating brown trout populations got weaker, and there was more genetic differentiation. The relative roles of genetic drift (small, isolated populations) and local adaptation in the increase of population differentiation with the increasing distance from the sea clearly deserves further study.

In most of the populations, the effective population sizes were low and clearly under the bare minimum rule-of-thumb level for a population viable in short term (Ne > 50) (Table 3). The relatedness was also relatively high in many of the samples (Table 3), characteristic to small populations. There were also samples (streams Haarainoja, Kutuoja, Naarmajoki), where the relatedness was at the level of full siblings, meaning extremely close relatedness in the population, or an artefact of all the samples coming from the same family (i.e. meaning for example that sampling in a stream was done in a section of too limited areal coverage to get a representative sample of the population).

5. River basins separately

The reassignment tests, where the probability of reassignment of each individual to a source population is tested, were done for all the samples together. This resulted in 71 possible source populations. For clarity, the legend with all the 71 possible source populations is presented separately (Fig. 6).

In the River Tornionjoki basin, the populations from streams Kutuoja and Särkijoki were clearly different from all the others (Fig. 7A). Both the Kutuoja and Särkijoki populations are, however, very small as indicated by the Ne (Table 3). All the populations used for comparison, especially the hatchery stocks from Rautalampi watercourse and River Oulujoki watercourse, were also clearly divergent from the River Tornionjoki basin's populations (Fig. 7A). The reassignment tests showed that there was more mixing in the streams Ahmajoki, Alanen Kihlankijoki, and Nivunkijoki, while the rest of the sampling sites – with a high proportion of siblings - appeared more distinct (Fig. 7B, Table 3). Because of the high proportion of relatives in many of the samples, the results should be interpreted with caution.

In the River Ounasjoki basin, the most divergent population was again the clearly smallest one, the population from stream Palontaustan latvaoja, with Ne of only 4 (Fig. 8A, Table 3). Again, the hatchery populations of adfluvial (lake-run) brown trout were clearly divergent from the rest of the populations. The reference samples – River Kemijoki and Iijoki sea-migrating brown trout, and resident brown trout (Kemihaara I and II) were also clearly different from the

River Ounasjoki basin's populations, of which stream Kienajaoja clustered closest to the reference samples (Fig. 8A). There was some mixing in the streams Karhuoja, Liivajoki, and Toto-oja populations (Fig. 8B). Interestingly, Liivajoki is a tributary flowing to stream Perttausjoki, the samples from which were mostly reassigned to the own population of origin.

In the river basin consisting of lower and middle reaches of Kemijoki, there were more clearly differentiated populations than in the other river basins (Fig. 9A). Only the population from stream Korkiamaanoja and naturally the River Kemijoki sea-migrating brown trout population reached the recommended level of Ne (Ne > 50; Table 3). The population from stream Ala-Runkausjoki clustered together with the stream Korkiamaanoja population, both being clearly divergent from the rest of the populations (Fig. 9A). Korkiamaanoja is a tributary flowing into Ala-Runkausjoki. Samples from the stream Konttijoki were also clearly different from the rest, while the samples from the stream Juujoki could not be clearly distinguished from the reference samples (Fig. 9A). There was very little mixing among the lower and middle reaches of River Kemijoki populations (Fig. 9B).

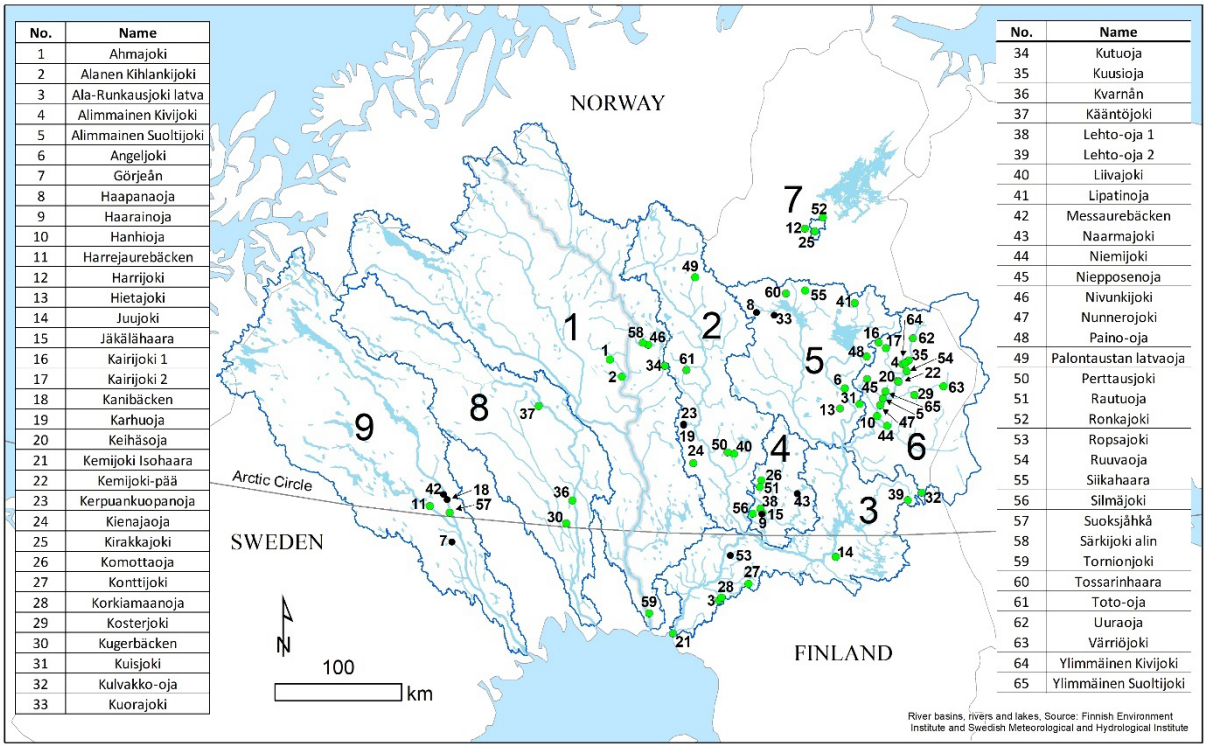
The River Raudanjoki basin also contained a number of clearly differentiated populations, which formed two clear clusters (Fig. 10A). In one cluster were the streams Haarainoja, Jäkälähaara, and Lehto-oja populations, and in the other the samples from streams Komottaoja, Naarmajoki, and Silmäjoki, which clustered together with the River Kemijoki and Iijoki sea-migrating brown trout reference samples (Fig. 10A). With the exception of the stream Haarainoja population, all Ne's were below the viable level of Ne > 50 (Table 3). There was some mixing in the stream Komottaoja population, and there was a clear-cut distinction of mixture vs. no mixture in the samples from the lower and upper parts of the stream Silmäjoki, respectively (Fig. 10B).

The samples from the River Kemihaara basin were analysed without the samples from stream Kulvakko-oja, due to the high number of full-sib families and clear geographic separation in the sample, which might severely bias the analyses of genetic differentiation. Most of the samples clustered together with the Kemihaara resident brown trout hatchery samples, suggesting an influence of the hatchery stocks in the area or because the brood stock for the hatchery population (Kemihaara I and II) is collected from this river basin (Fig. 11A). The samples from streams Alimmainen and Ylimmäinen Kivijoki clustered close together with the samples from stream Niepposenoja (Fig. 11A). The samples from the stream Värriojoki were closer to the sea-migrating brown trout samples than to the rest of the River Kemihaara basin samples (Fig. 11A). Overall, the populations from River Kemihaara basin were less well defined and there was more indication of gene flow than among the populations from the other river basins (Fig. 11B). There was a clear distinction between the samples from the lower and upper parts of the stream Kairijoki, the samples from the upper being mostly reassigned to their own population of origin, while the samples from the lower parts showed more mixing, especially with the Kemihaara resident brown trout hatchery samples (Fig. 11B). Only two of the 18 populations from River Kemihaara basin had Ne > 50 (streams Alimmainen Suoltijoki and Niemijoki; Table 3).

Among the populations from the River Luiro-Kitinen basin, samples from the stream Hietajoki were clearly genetically different from the rest, as were the samples from the stream Kuisjoki (Fig. 12A). The samples from the stream Kuorajoki clustered close to and showed some mixing with the River Oulujoki hatchery population (Fig. 12A). As was evident also from the analysis of genetic structuring and admixture (Fig. 3), the stream Kuorajoki population has been influenced by stocking from Lake Inari basin reference population (Fig. 12B). There was also a significant amount of mixing in the streams Angeljoki, Paino-oja, and Tossarinhaara populations (Fig. 12B). All the River Luiro-Kitinen basin's populations were below Ne > 50 (Table 3).

The populations from River Luleå basin (n=5) were analysed together with the populations from the River Kalix basin (n=3). Samples from the streams Messaurebäcken and Suoksjåhkå clustered close to populations from the streams Kääntöjoki and Kugerbäcken populations from the River Kalix basin (Fig. 13A). The stream Görjeån population was close to the stream Kvarnån population from the River Kalix basin, while the samples from the stream Kanibäcken clustered clearly separately from all the others (Fig. 13A). Most of the samples from the Swedish populations fell into their source populations in the reassignment tests, although there was some mixing with the sea-migrating brown trout reference populations, especially in streams Messaurebäcken, Suoksjåhka and Kugerbäcken (Fig. 13B). The Ne's in the Swedish populations were clearly higher than those in the Finnish rivers (Table 3).

6. Figures and tables

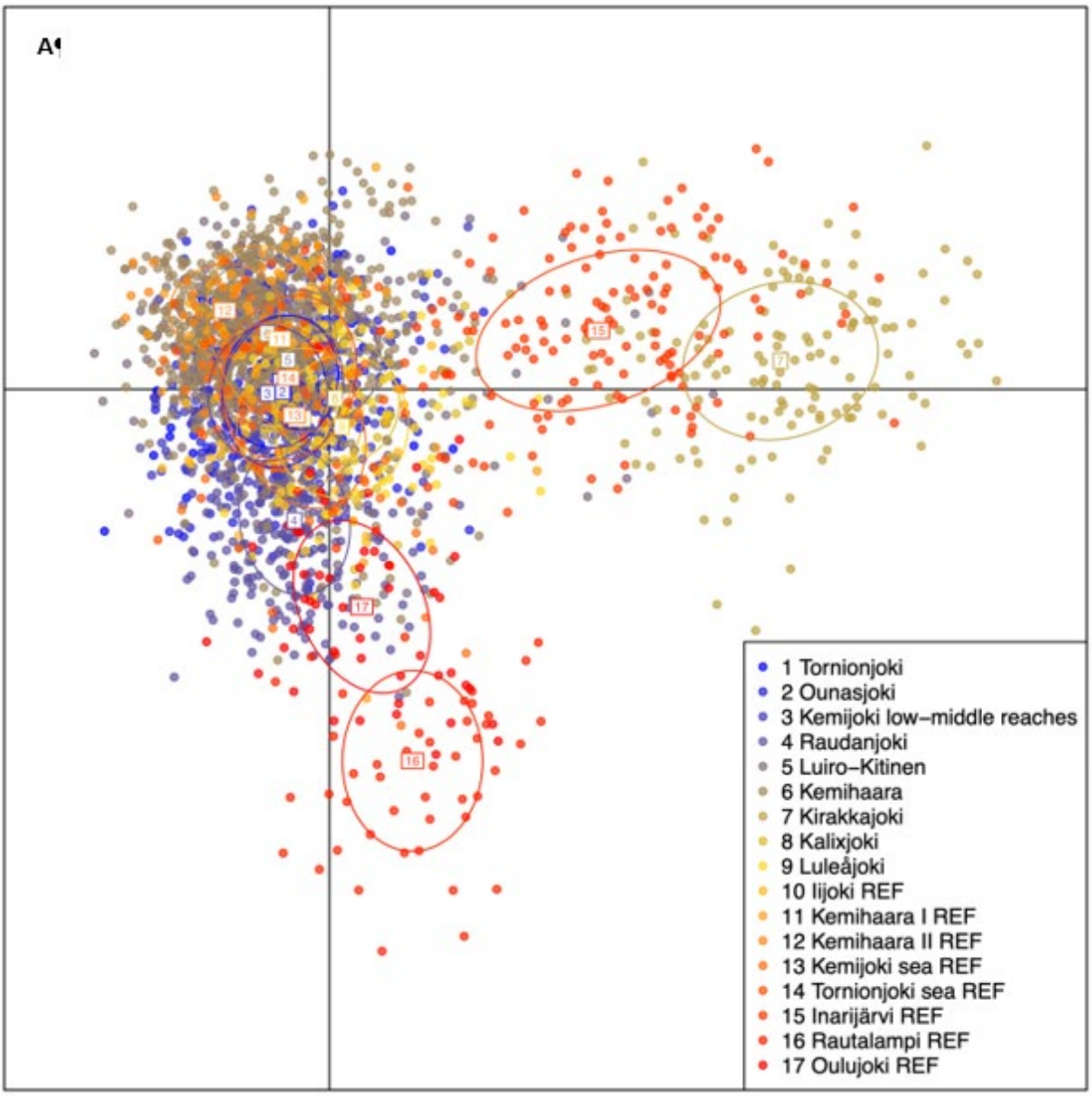


No.	Name
1	Ahmajoki
2	Alanen Kihlankijoki
3	Ala-Runkausjoki latva
4	Alimmainen Kivijoki
5	Alimmainen Suoltijoki
6	Angeljoki
7	Görjeån
8	Haapanaoja
9	Haarainoja
10	Hanhioja
11	Harrejaurebäcken
12	Harrijoki
13	Hietajoki
14	Juujoki
15	Jäkälähaara
16	Kairijoki 1
17	Kairijoki 2
18	Kanibäcken
19	Karhuoja
20	Keihäsoja
21	Kemijoki Isohaara
22	Kemijoki-pää
23	Kerpuankuopanoja
24	Kienajaoja
25	Kirakkajoki
26	Komottaoja
27	Konttijoki
28	Korkiamaanoja
29	Kosterjoki
30	Kugerbäcken
31	Kuisjoki
32	Kulvakko-oja
33	Kuorajoki

No.	Name
34	Kutuoja
35	Kuusioja
36	Kvarnån
37	Kääntöjoki
38	Lehto-oja 1
39	Lehto-oja 2
40	Liivajoki
41	Lipatinoja
42	Messaurebäcken
43	Naarmajoki
44	Niemijoki
45	Niepposenoja
46	Nivunkijoki
47	Nunnerojoki
48	Paino-oja
49	Palontaustan latvaoja
50	Perttausjoki
51	Rautuoja
52	Ronkajoki
53	Ropsajoki
54	Ruuvaoja
55	Siikahaara
56	Silmäjoki
57	Suoksjåhkå
58	Särkijoki alin
59	Tornionjoki
60	Tossarinhaara
61	Toto-oja
62	Uuraoja
63	Värriöjoki
64	Ylimmäinen Kivijoki
65	Ylimmäinen Suoltijoki

Figure 1. A map illustrating the river basins used in the brown trout genetic analyses: 1) River Tornionjoki basin, 2) River Ounasjoki basin, 3) River Kemijoki basin including lower and middle reaches of the river, 4) River Raudanjoki basin, 5) River Kitinen and Luiro basin, 6) River

Kemihaara basin, 7) Lake Inari basin, 8) River Kalix basin, and 9) River Luleå basin. Basins 2-6 form the main River Kemijoki basin. Sampled streams and they locations are indicated by numbers and dots (green dot: sample size >= 20 individuals, black dot sample size <20).



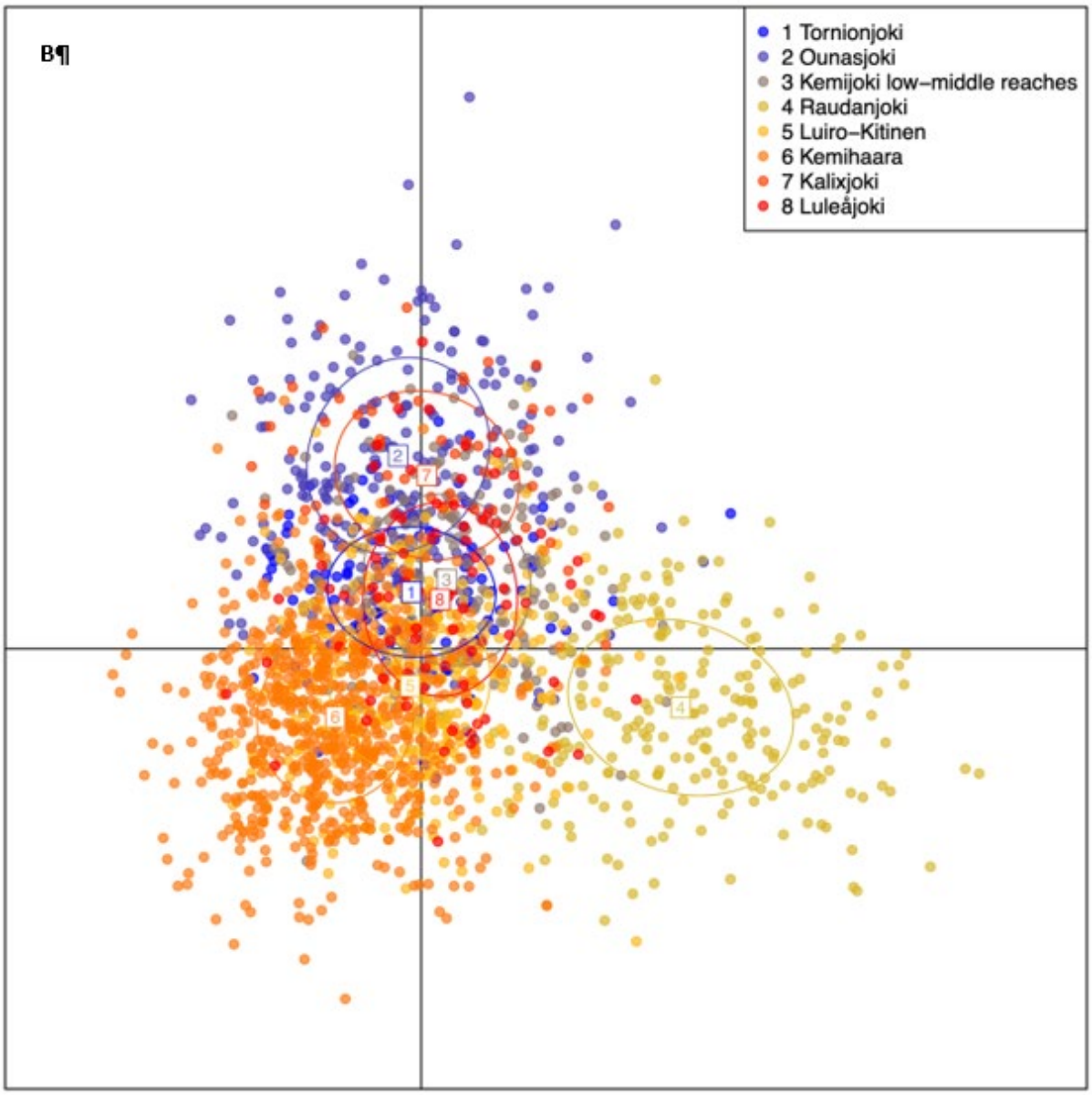


Figure 2. A) Genetic differentiation among the brown trout samples from different river basins based on the discriminant analysis of principal components (DAPC). In the scatter plot, each dot depicts an individual brown trout sample. B) Scatterplot of the DAPC scores without the reference groups.

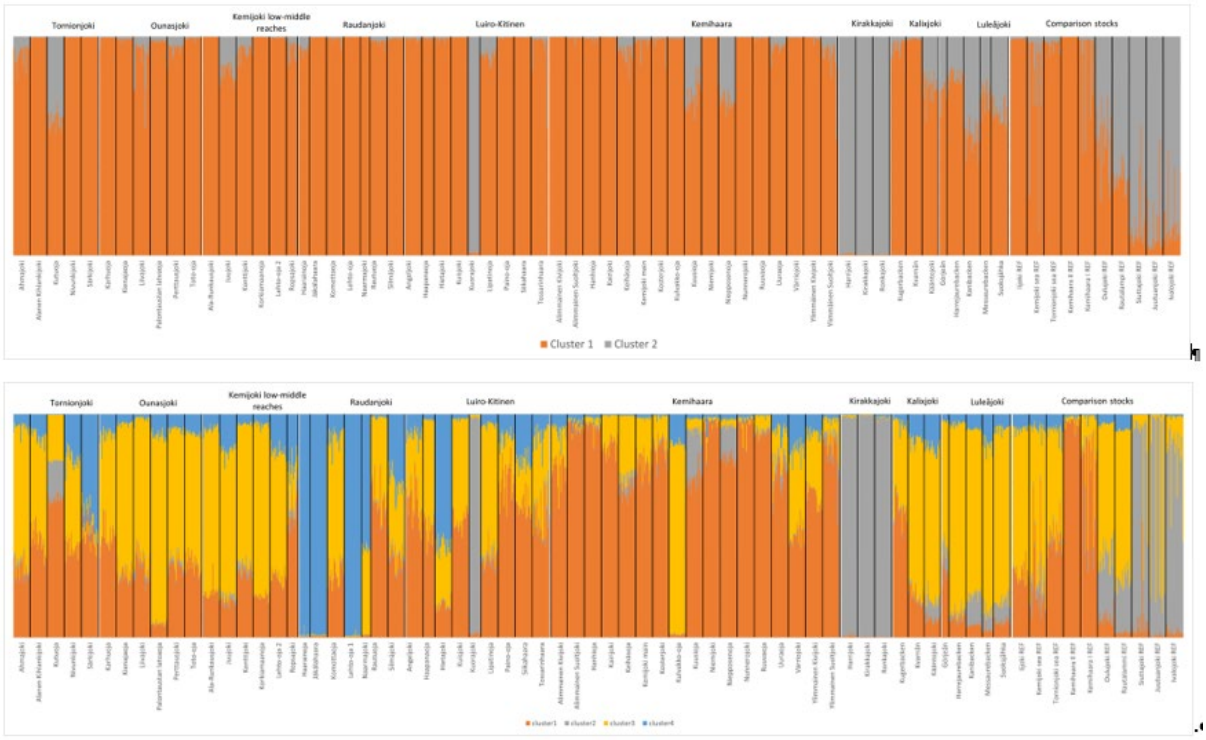


Figure 3. Genetic structure of the sampled brown trout populations based on similarity of their genotypes. The sampled populations are separated with vertical black lines and each bar represents an individual fish and the proportion of its genotype that is similar to the others in the genetic cluster with a corresponding colour. A) the samples are divided into two genetic clusters (k=2), B) the samples are divided into four genetic clusters (k=4).

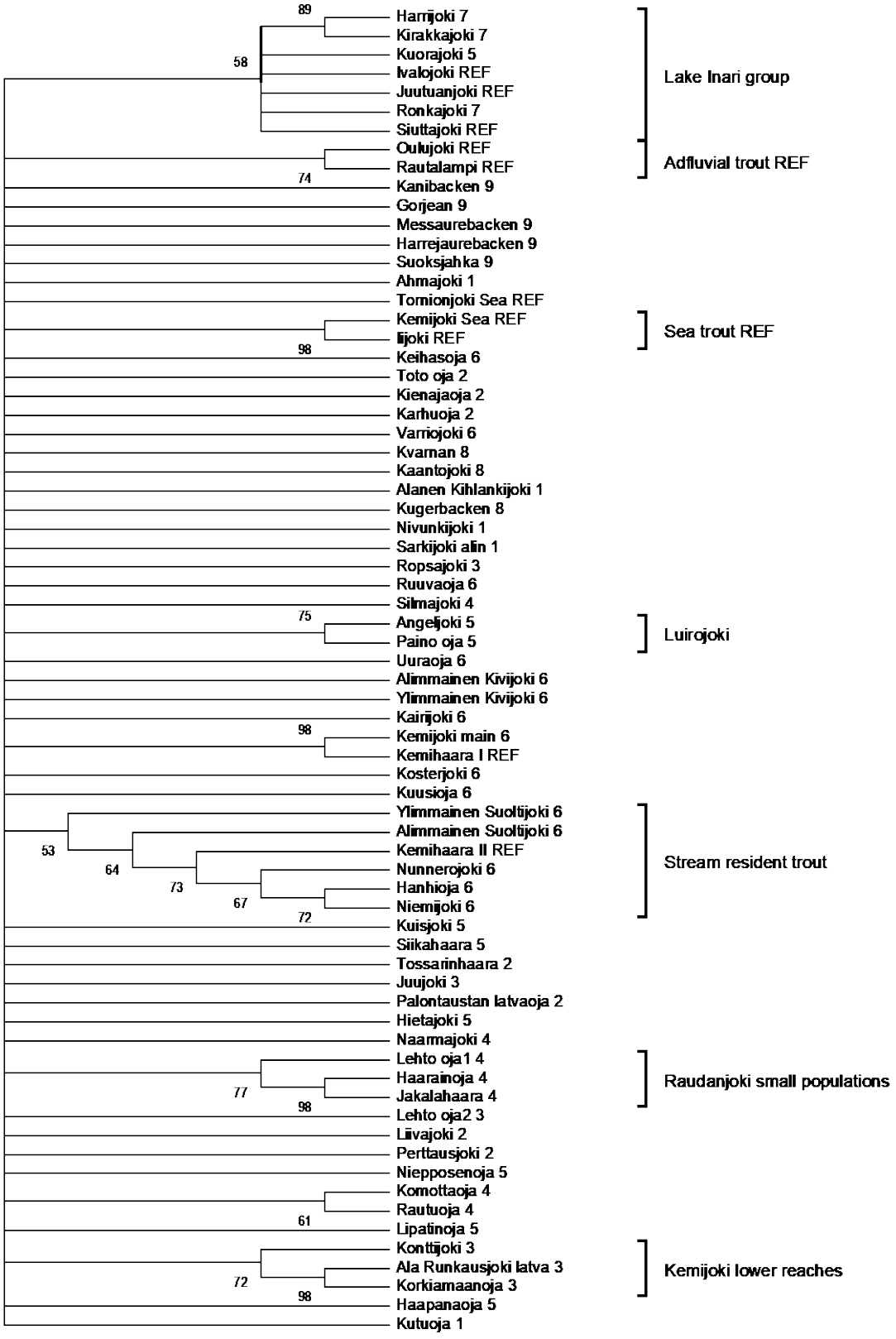

89
58
74
98
75
98
53
64
73
67
72
77
98
61
72
98
Harrijoki 7
Kirakkajoki 7
Kuorajoki 5
Ivalojoki REF
Juutuanjoki REF
Ronkajoki 7
Siuttajoki REF
Oulujoki REF
Rautalampi REF
Kanibacken 9
Gorjean 9
Messaurebacken 9
Harrejaurebacken 9
Suoksjahka 9
Ahmajoki 1
Tornionjoki Sea REF
Kemijoki Sea REF
Iijoki REF
Keihasoja 6
Toto oja 2
Kienajaoja 2
Karhuoja 2
Varriojoki 6
Kvarnan 8
Kaantojoki 8
Alanen Kihlankijoki 1
Kugerbacken 8
Nivunkijoki 1
Sarkijoki alin 1
Ropsajoki 3
Ruuvaoja 6
Silmajoki 4
Angeljoki 5
Paino oja 5
Uuraoja 6
Alimmainen Kivijoki 6
Ylimmainen Kivijoki 6
Kairijoki 6
Kemijoki main 6
Kemihaara I REF
Kosterjoki 6
Kuusioja 6
Ylimmainen Suoltijoki 6
Alimmainen Suoltijoki 6
Kemihaara II REF
Nunnerojoki 6
Hanhioja 6
Niemijoki 6
Kuisjoki 5
Siikahaara 5
Tossarinhaara 2
Juujoki 3
Palontaustan latvaoja 2
Hietajoki 5
Naarmajoki 4
Lehto oja1 4
Haarainoja 4
Jakalahaara 4
Lehto oja2 3
Liivajoki 2
Perttausjoki 2
Niepposenoja 5
Komottaoja 4
Rautuoja 4
Lipatinoja 5
Konttijoki 3
Ala Runkausjoki latva 3
Korkiamaanoja 3
Haapanaoja 5
Kutuoja 1
Lake Inari group
Adfluvial trout REF
Sea trout REF
Luirojoki
Stream resident trout
Raudanjoki small populations
Kemijoki lower reaches

Figure 4. Dendrogram showing genetic relationships among the brown trout populations. The numbers after the names of the sampling sites indicate the river basin. All the nodes with bootstrap values <50 have been collapsed. The bootstrap values next to the nodes give the percentage of 1000 replicate trees, where the branch is in the same position.

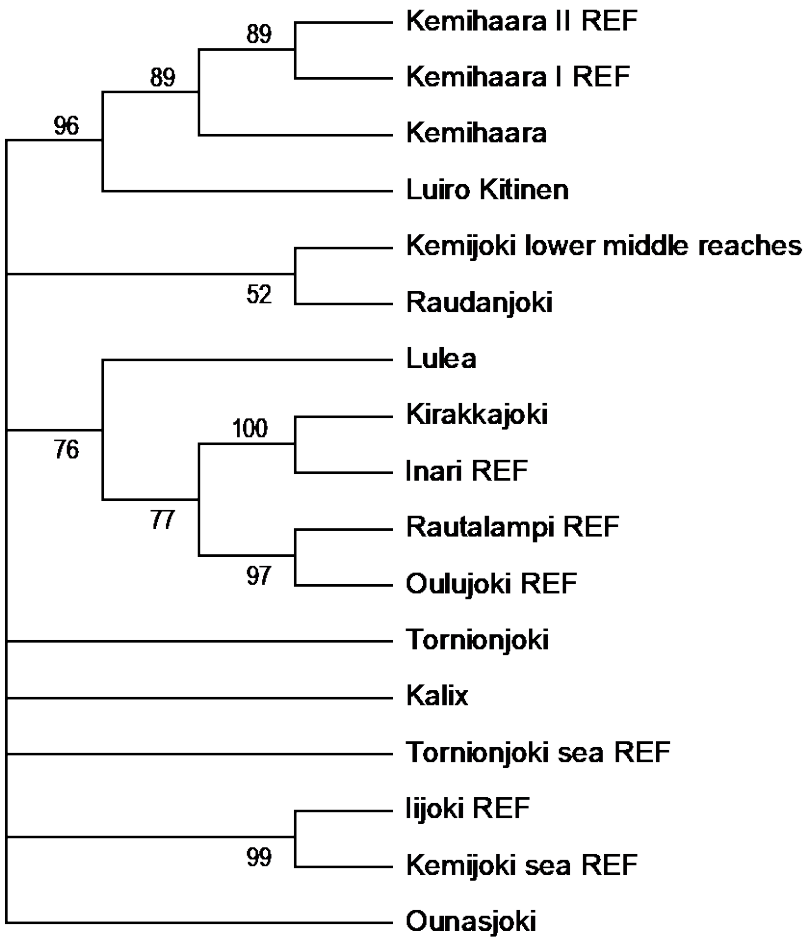


Figure 5. Dendrogram showing genetic relationships among different river basins and reference populations. All the nodes with bootstrap values <50 have been collapsed.

Figures 7-13 show the results for the discriminant analysis of principal components (DAPC) for each river basin separately as scatterplots (A), where each dot depicts an individual fish and as bar plots of the posterior probabilities (B) with which each individual fish (vertical bar) was reassigned to one of the 71 sampling sites (including the reference samples). Because of the large number of groups, the colors corresponding to each of the sampling sites are shown in a separate plot (Figure 6).

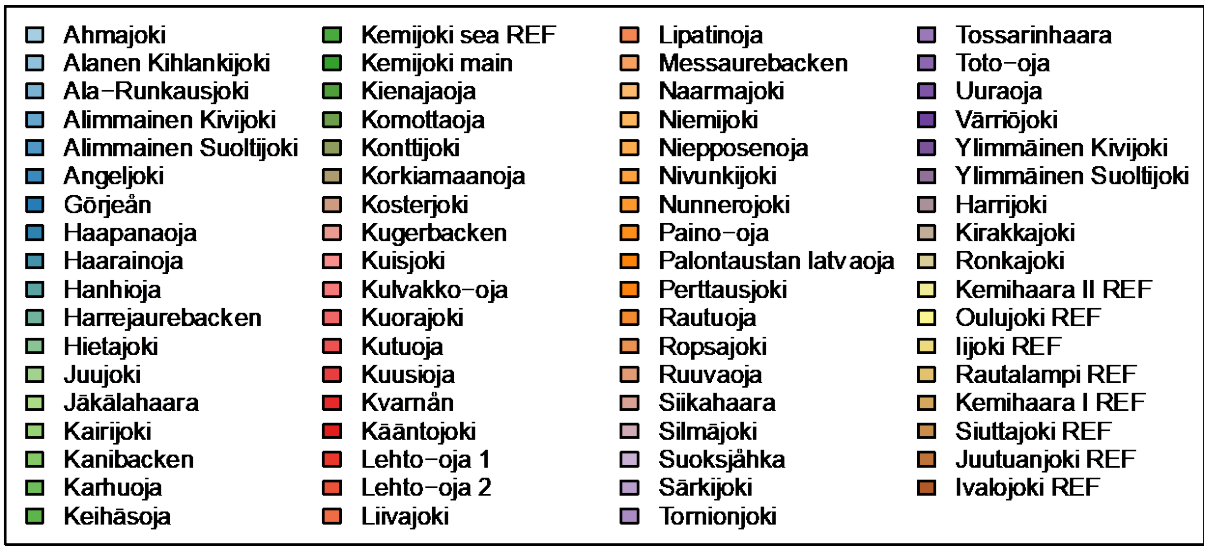


Figure 6. Legend of colours for the reassignment tests (Figs. 7B-13B).

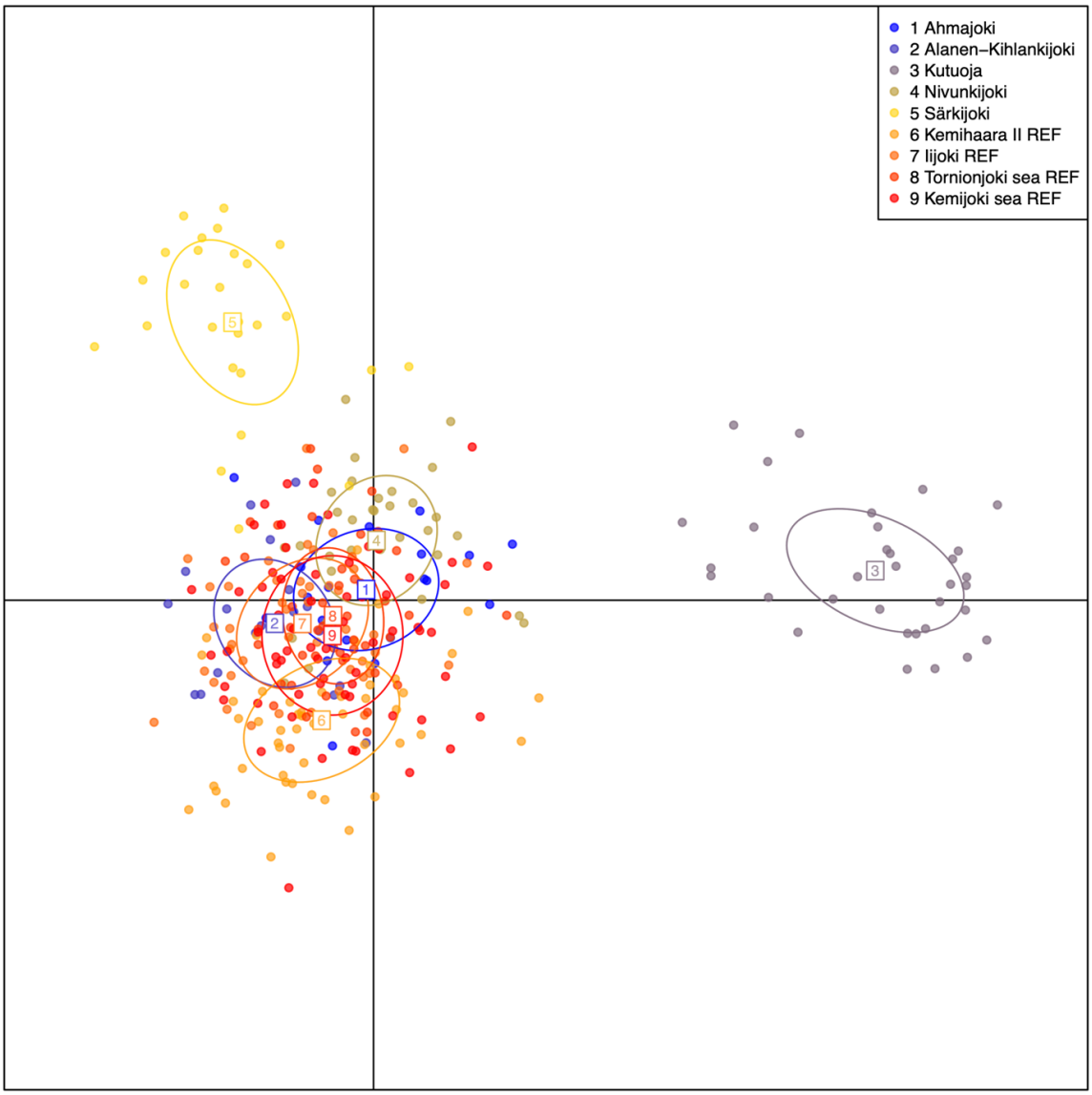


Figure 7A. Tornionjoki basin DAPC scatterplot.

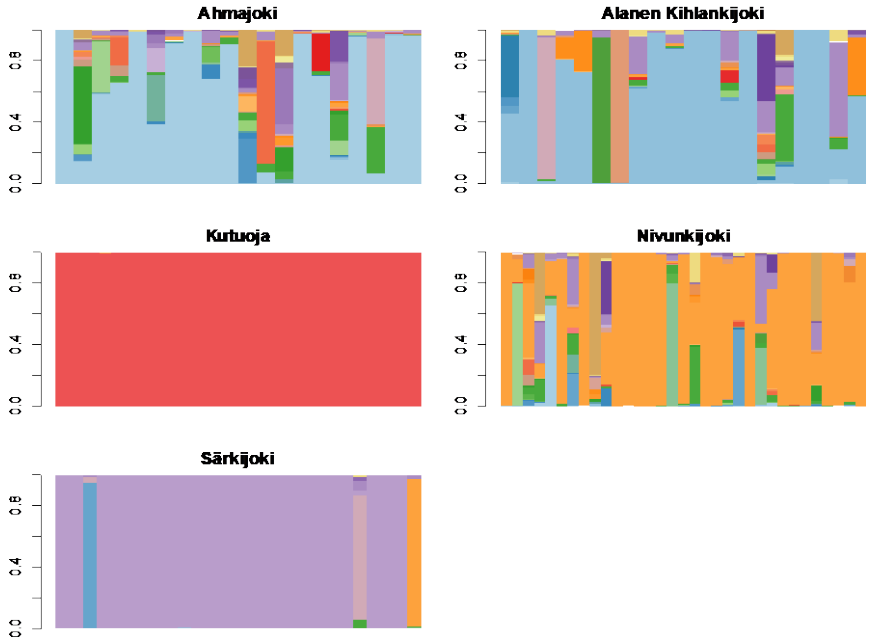


Figure 7B. Tornionjoki basin reassignment plot.

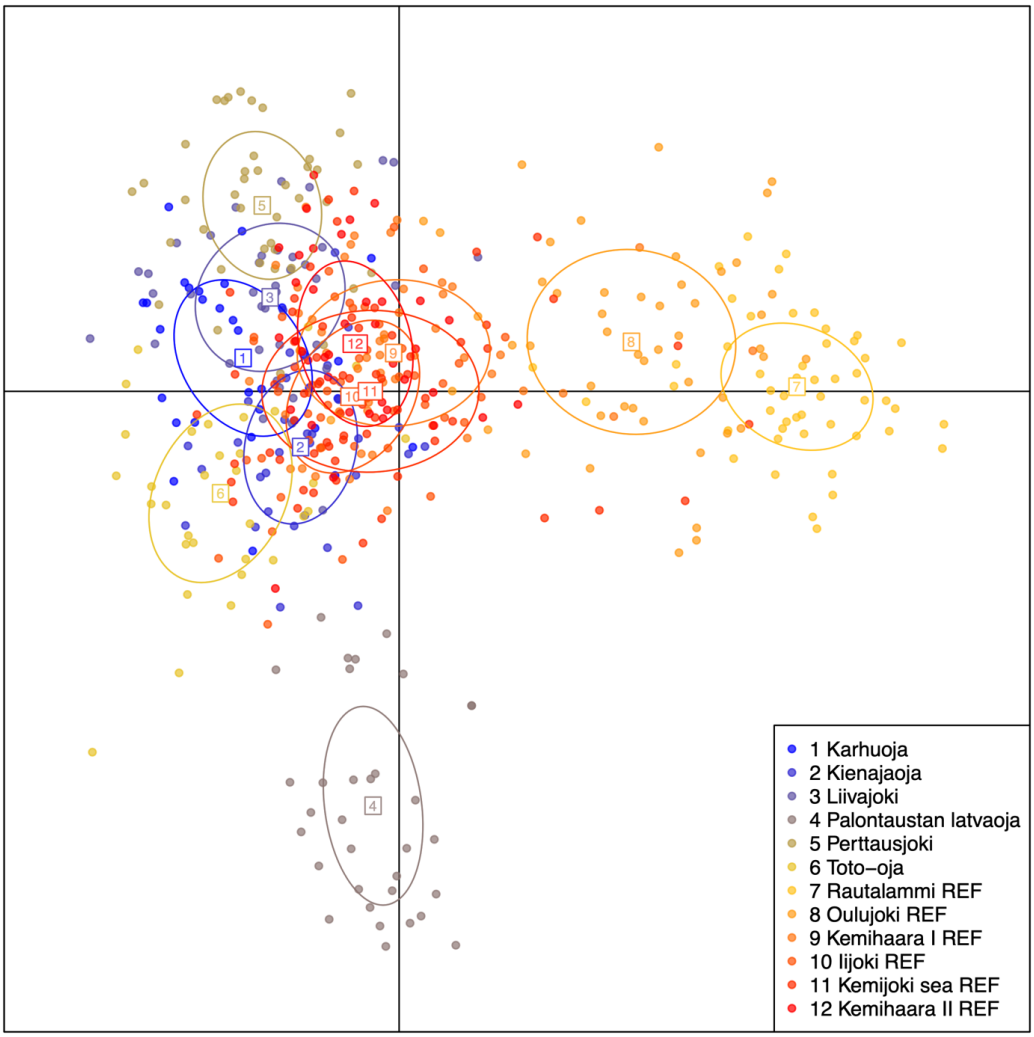


Figure 8A. Ounasjoki basin DAPC scatterplot.

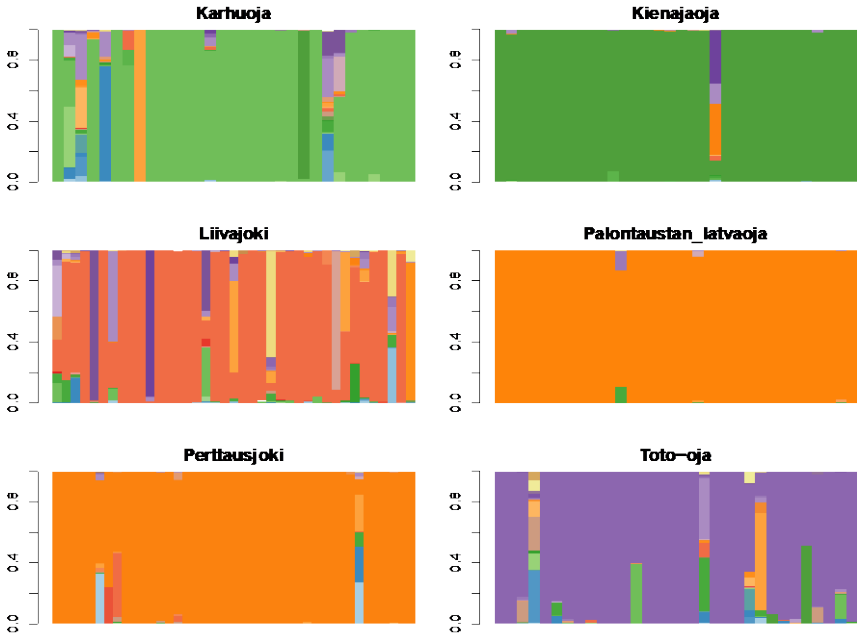


Figure 8B. Ounasjoki basin reassignment plot.

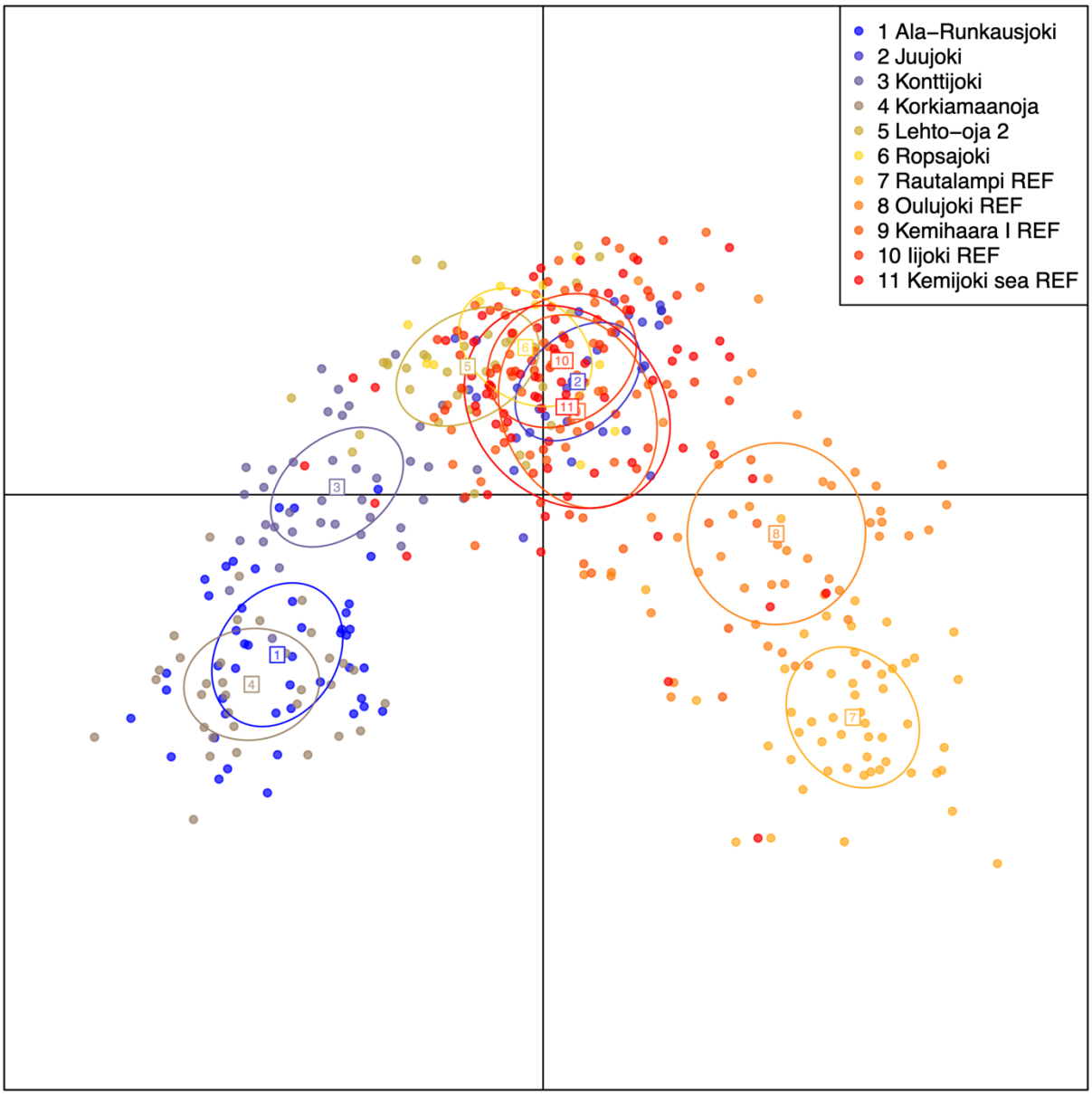


Figure 9A. Kemijoki low-middle reaches DAPC scatterplot.

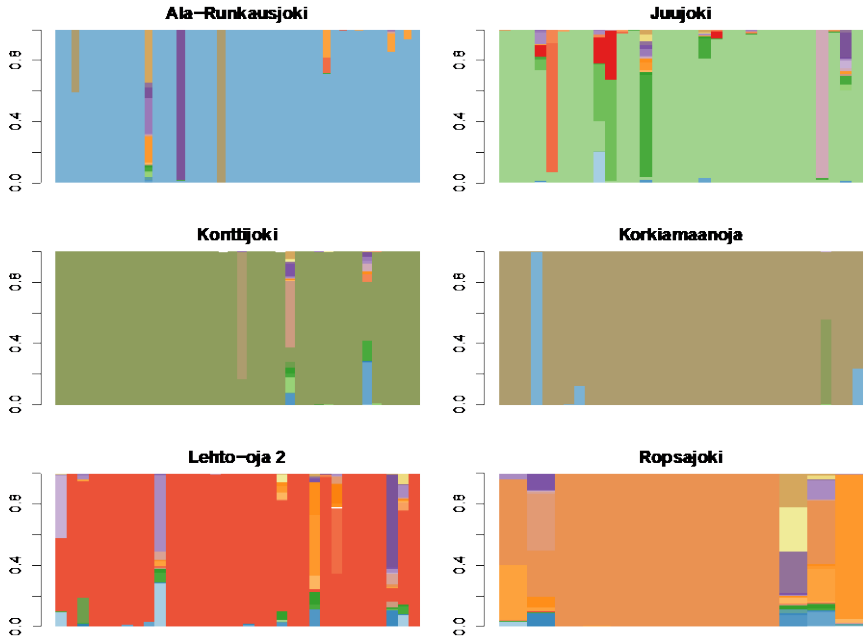


Figure 9B. Kemijoki low-middle reaches reassignment plot.

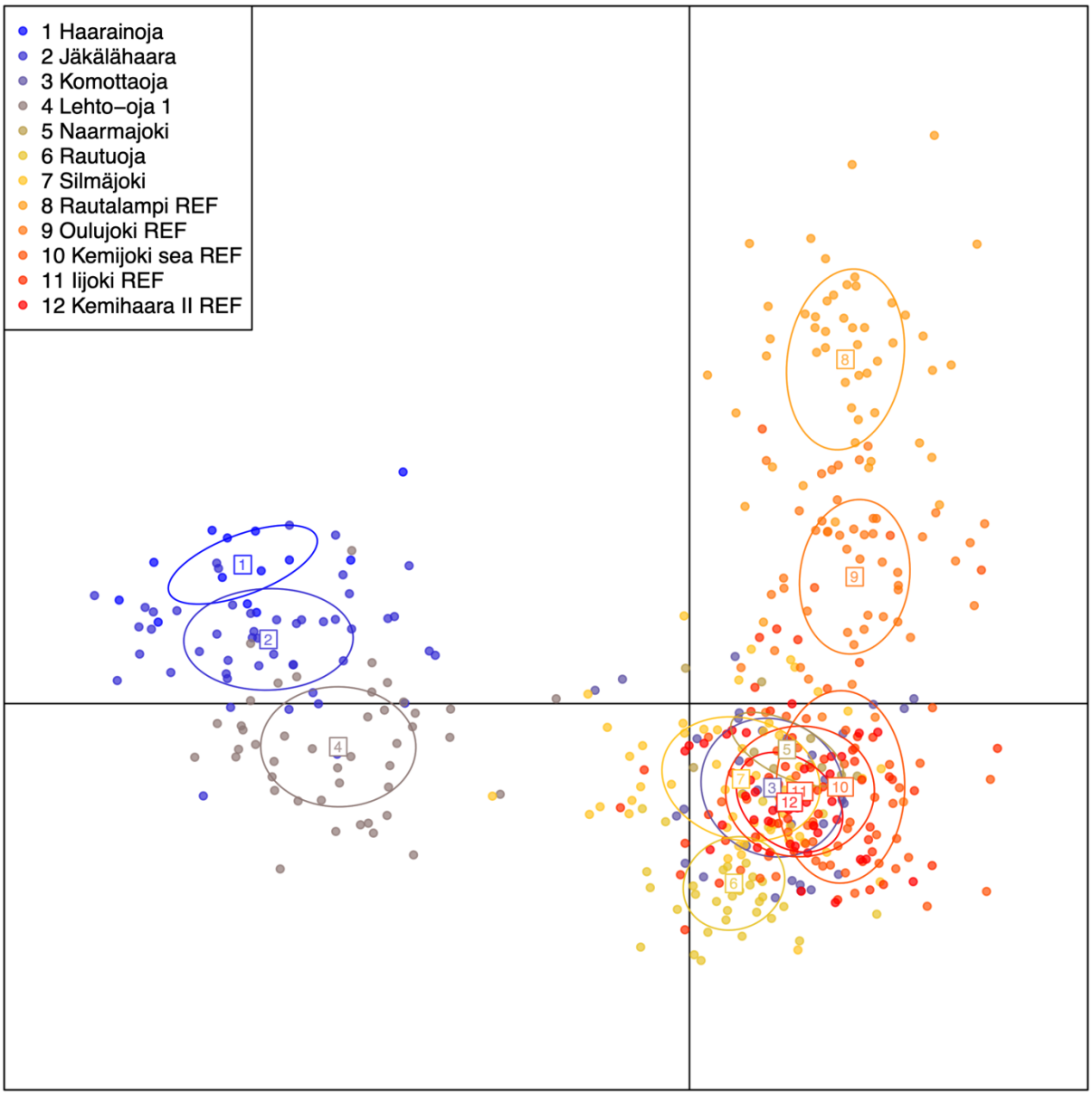


Figure 10A. Raudanjoki basin DAPC scatterplot.

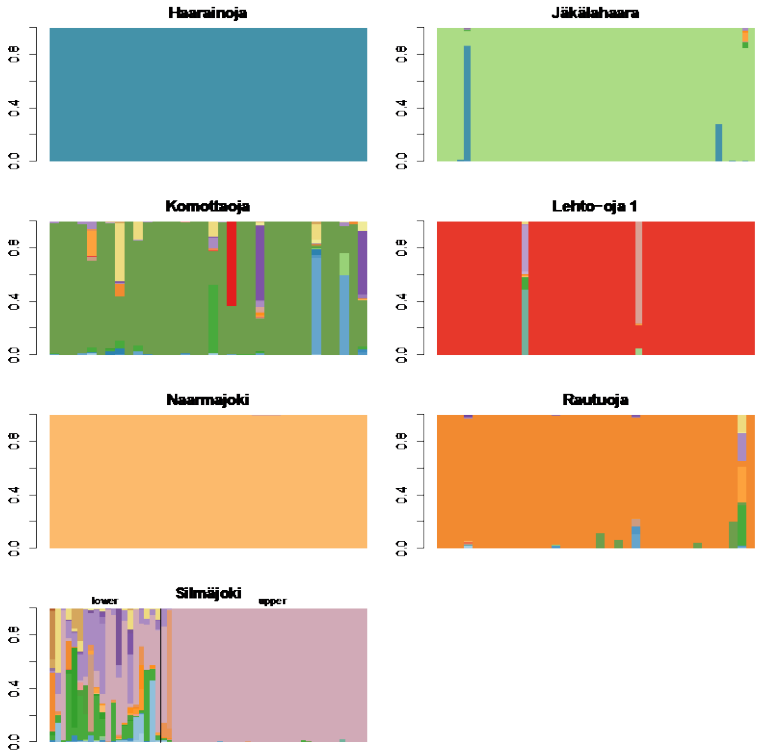


Figure 10B. Raudanjoki basin reassignment plot.

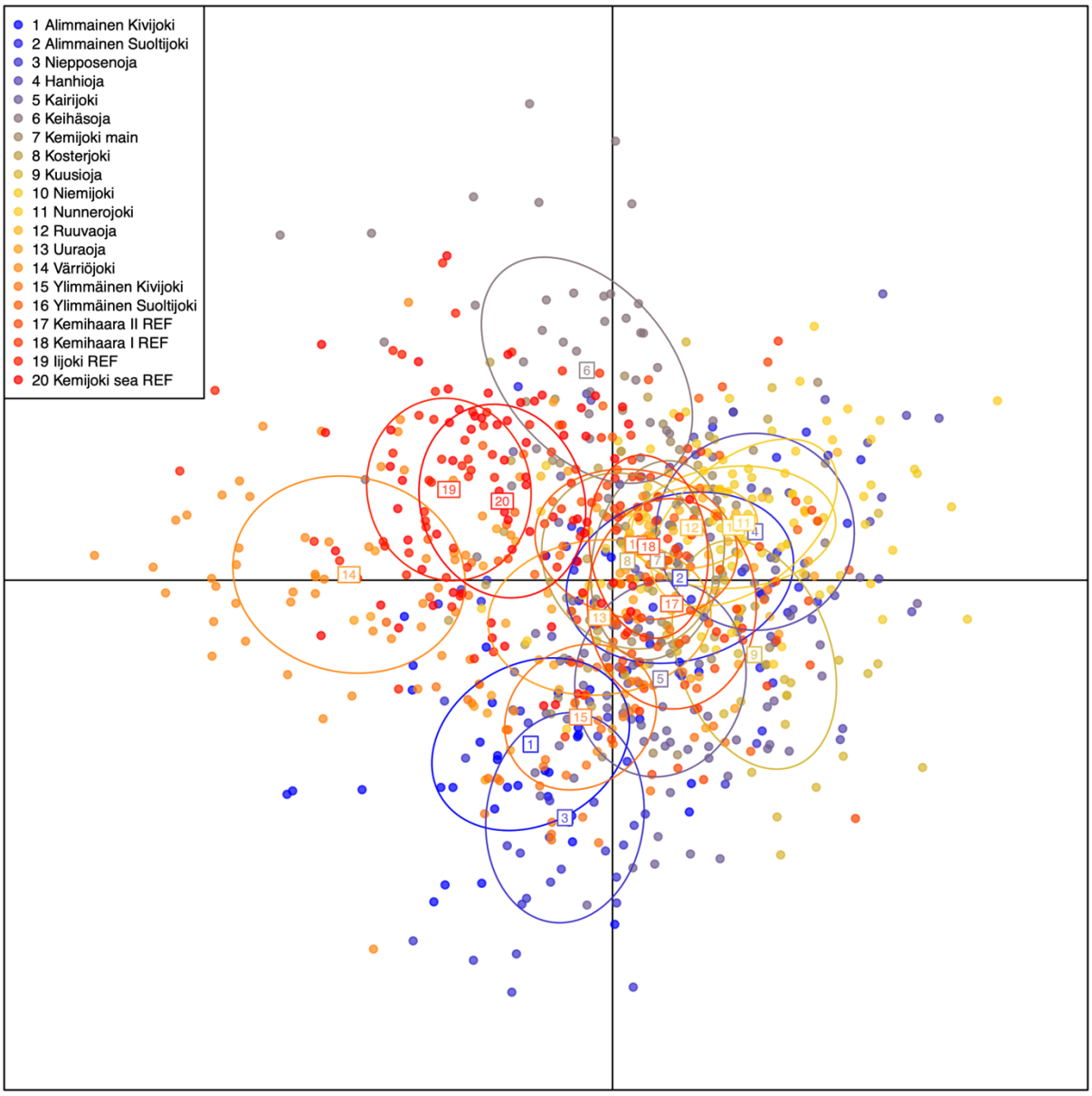


Figure 11A. Kemihaara basin DAPC scatterplot.

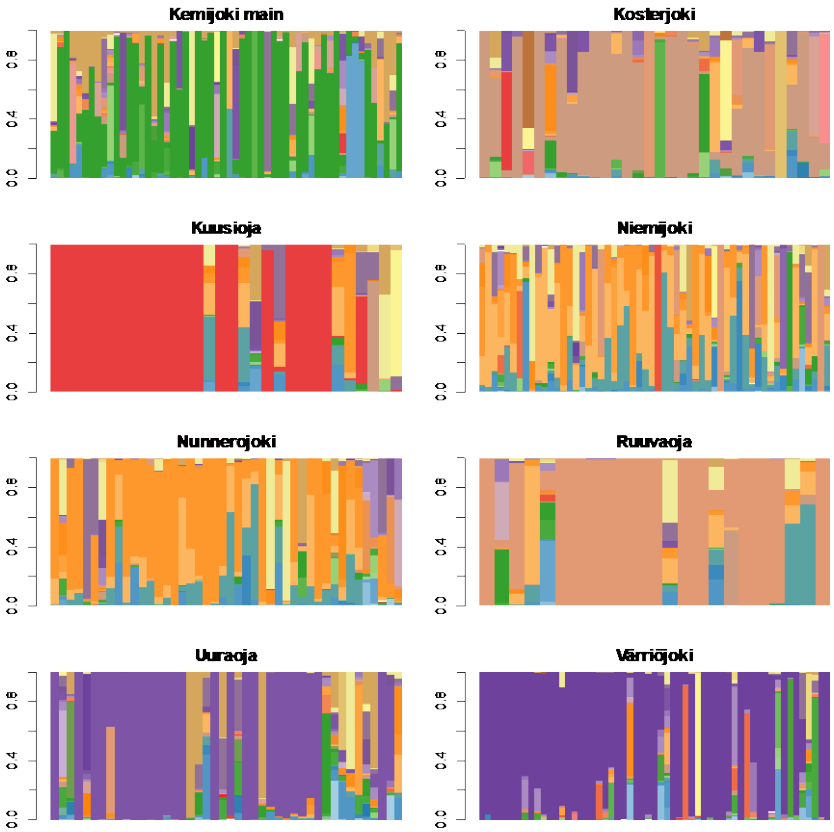


Figure 11B. Kemihaara basin reassignment plot.

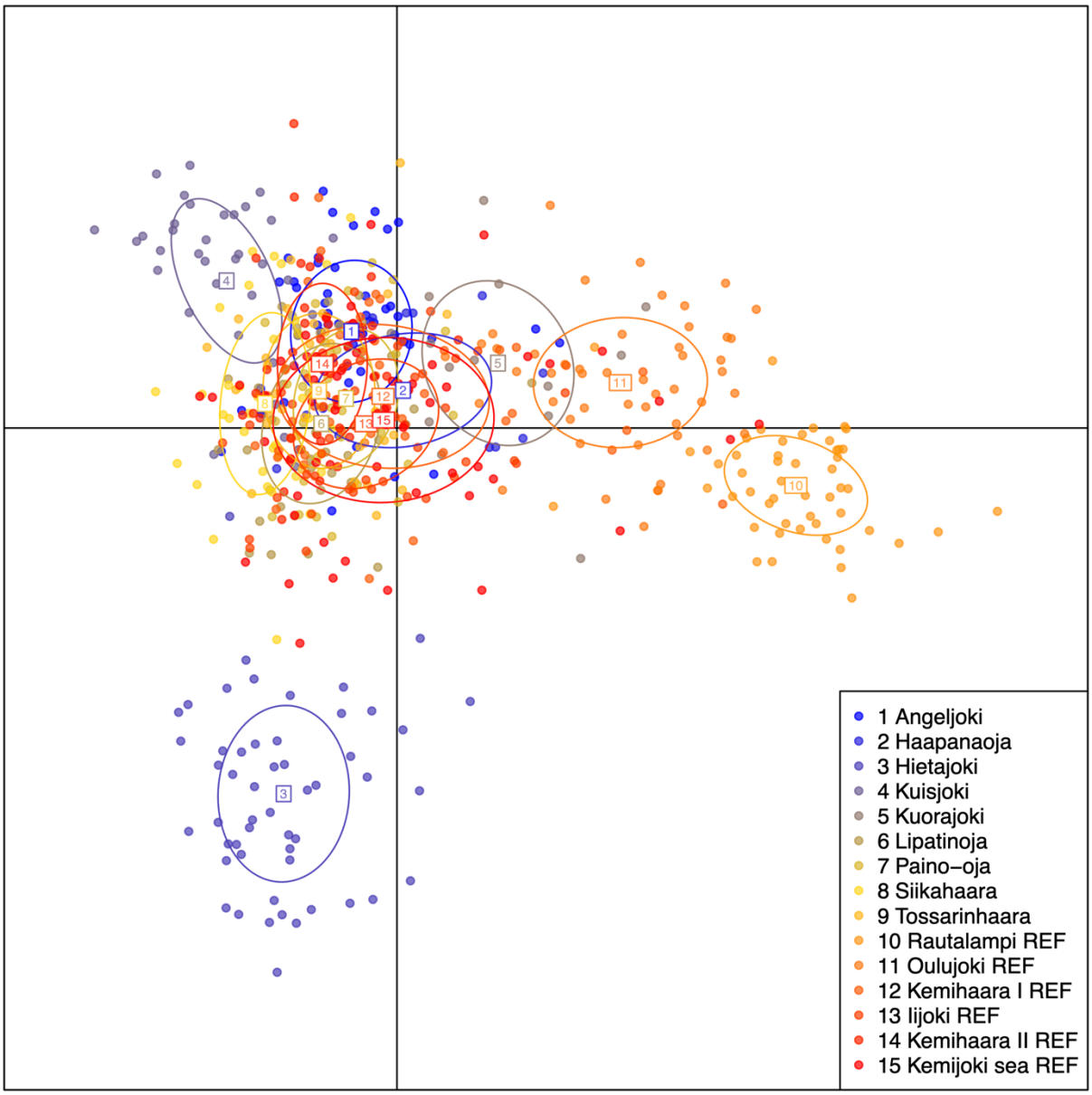


Figure 12A. Luiro-Kitinen basin DAPC scatterplot.

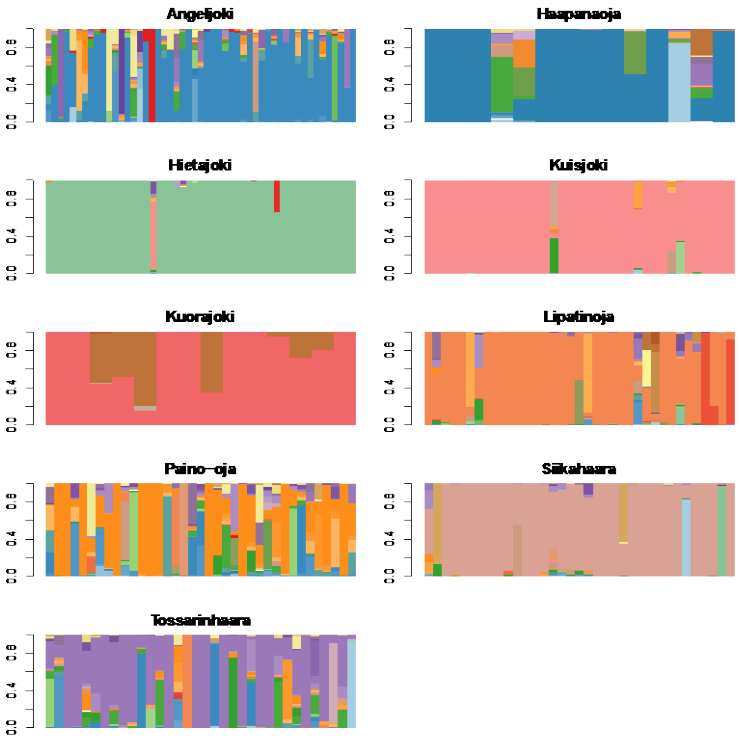


Figure 12B. Luiro-Kitinen basin reassignment plot.

A scatterplot showing genetic differences between the trout samples taken from the different catchments.

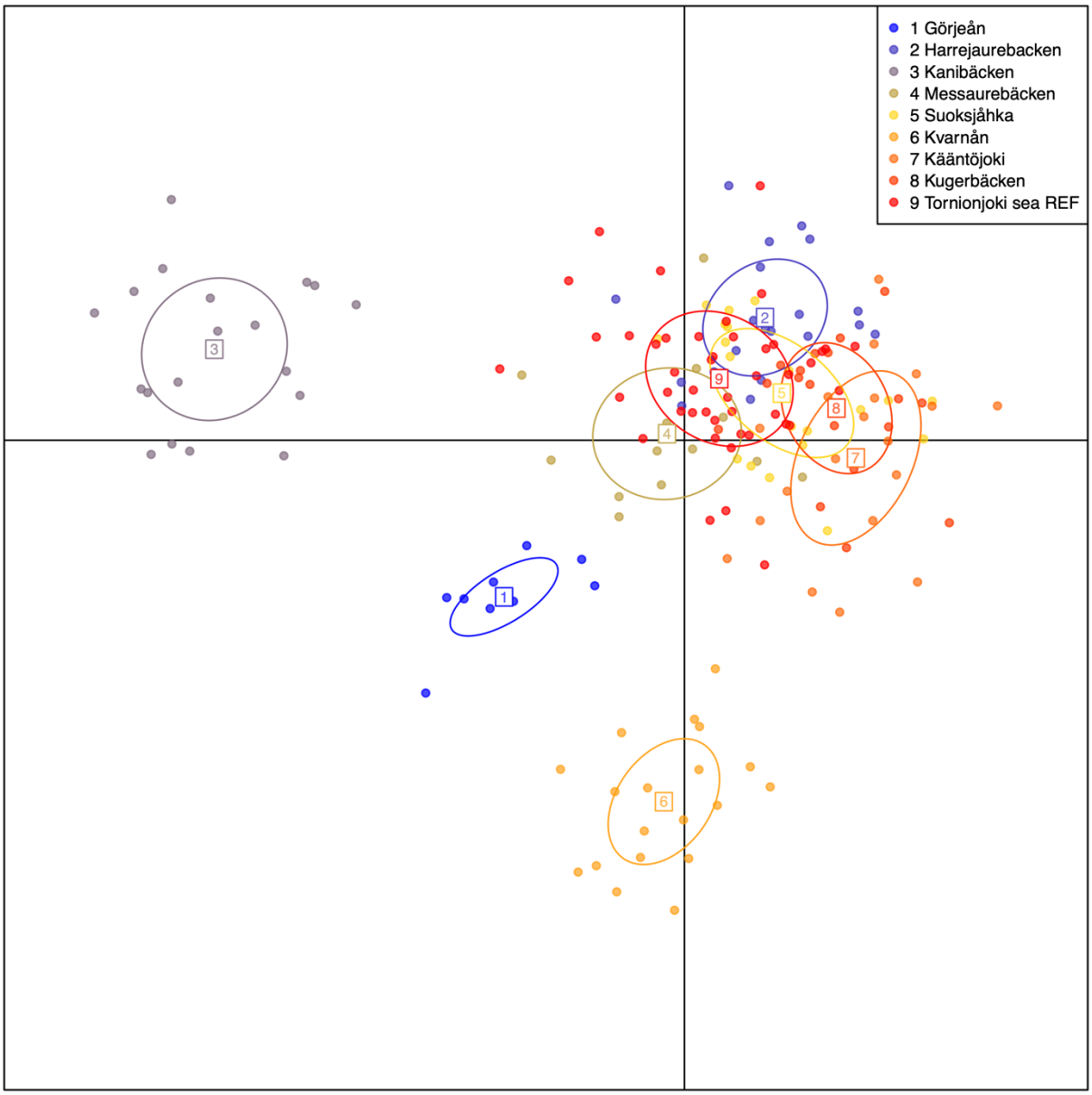


Figure 13A. River Kalix and Luleå basins DAPC scatterplot.

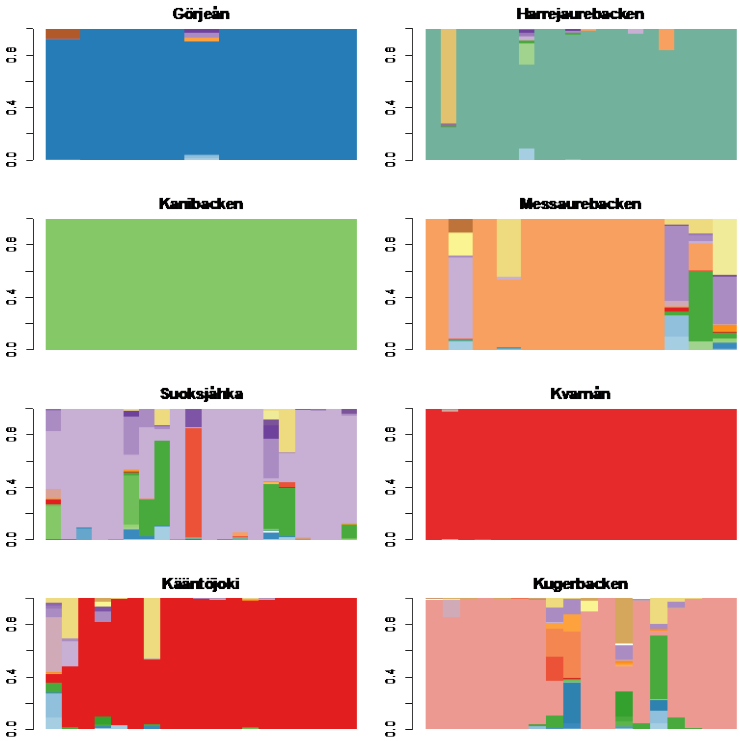


Figure 13B. River Kalix and Luleå basins reassignment plot.

Box 1. Glossary of indices of genetic diversity and differentiation (see Table 3).

Number of alleles (A): genetic diversity measured in the number of alleles in a population. Higher the number of alleles, higher the population's potential to adapt and persist

Expected heterozygosity (H_E): genetic diversity measured as the number of genes in a locus. The expected heterozygosity is the level of heterozygosity in the population without any influences to genetic variation (e.g. bottlenecks, natural selection)

Observed heterozygosity (H_O): the fraction of heterozygotes in a population. This is compared to the expected heterozygosity to make inferences on influences to genetic variation. If HO is lower than HE, it means that there is inbreeding. HO > HE can mean that the population is not following the assumptions of the population genetic models are based: for example, large population, random breeding, no substructure.

Allelic richness (A_R): genetic diveristy of a population measured as the number of alleles in a population and taking into account differences in the sample size among different populations.

Effective population size (Ne): the number of individuals reproducing in a population. A rule-of-thumb value for a population viable in short term is Ne > 50, although according more recent estimates it should be Ne > 100 (Frankham et al. 2014)

Relatedness (r_{wang}): the proportion of genes shared by individuals. Full siblings: 0.5, half sibs: 0.25, first cousins: 0.125.

F_{ST}: Fixation index measuring genetic differentiation. Here calculated as the proportion of genetic variance in a population (sampling site) to the total genetic variance (among all samples).

D_A: Measure of genetic distance, here based on differences in microsatellite genotypes.

6.1. References

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: *re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data*. Molecular Ecology Resources 14, 209–214. https://doi.org/10.1111/1755-0998.12157

Frankham, R., Bradshaw, C.J.A., Brook, B.W., 2014. *Genetics in conservation management: Revised recommendations for the 50/500 rules*, Red List criteria and population viability analyses.Biological Conservation, 170, 56-63.

Jombart, T., 2008. adegenet: *a R package for the multivariate analysis of genetic markers*. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., Balloux, F., 2010. *Discriminant analysis of principal components: a new method for the analysis of genetically structured populations*. BMC Genetics 11, 94. https://doi.org/10.1186/1471-2156-11-94

Jombart, T., & Collins, C. (2017). *A tutorial for Discriminant Analysis of Principal Components* (DAPC) using adegenet 2.1. 0. Imperial College, London, United Kingdom.

Jones, O.R., Wang, J., 2010. *COLONY: a program for parentage and sibship inference from multilocus genotype data*. Molecular Ecology Resources 10, 551–555. https://doi.org/10.1111/j.1755-0998.2009.02787.x

Keenan, K., McGinnity, P., Cross, T.F., Crozier, W.W., Prodöhl, P.A., 2013. diveRsity: *An R package for the estimation and exploration of population genetics parameters and their associated errors.* Methods in Ecology and Evolution 4, 782–788. https://doi.org/10.1111/2041-210x.12067

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: *a program for identifying clustering modes and packaging population structure inferences across K.* Molecular Ecology Resources 15, 1179–1191. https://doi.org/10.1111/1755-0998.12387

Kraemer, P., Gerlach, G., 2017. *Demerelate: calculating interindividual relatedness for kinship analysis based on codominant diploid genetic markers using R*. Molecular Ecology Resources 17, 1371–1377. https://doi.org/10.1111/1755-0998.12666

Langella, O., 2000. populations 1.2.31: *Population genetic software.* Department of populations, genetics and evolution, CNRS, France.

Nei, M., Tajima, F., Tateno, Y., 1983. *Accuracy of estimated phylogenetic trees from molecular data*. Journal of Molecular Evolution 19, 153–170. https://doi.org/10.1007/bf02300753

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. *Inference of Population Structure Using Multilocus Genotype Data*. Genetics 155, 945–959.

R Core Team, 2019. R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Waples, R.S., Do, C., 2008. *ldne: a program for estimating effective population size from data on linkage disequilibrium.* Molecular Ecology Resources 8, 753–756. https://doi.org/10.1111/j.1755-0998.2007.02061.x

Wright, S., 1949. *The Genetical Structure of Populations.* Annals of Eugenics 15, 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x