

Ecography

ECOG-05301

Mathieu-Bégné, E., Loot, G., Mazé-Guilmo, E., Mullet, V., Genthon, C. and Blanchet, S. 2020. Combining species distribution models and population genomics underlines the determinants of range limitation in an emerging parasite. – Ecography doi: 10.1111/ecog.05301

Supplementary material

Appendix 1

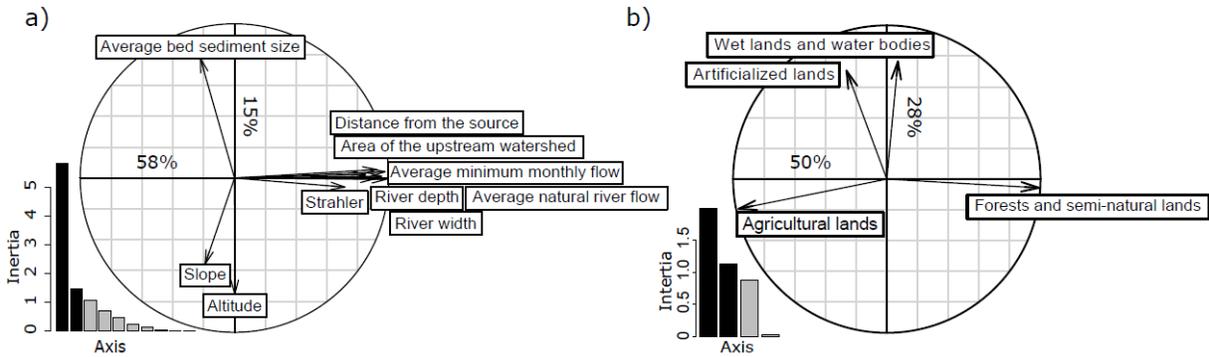


Figure A1: Principal Component Analyses for hydrographic variables (a) and landscape variables (b) respectively. Bar plots indicate how much of the variance is explained by each PCA axis and black bars indicate the number of PCA axes retained as synthetic variables in subsequent analysis. Percentages correspond to the percentage of variance explained by each PCA axis.

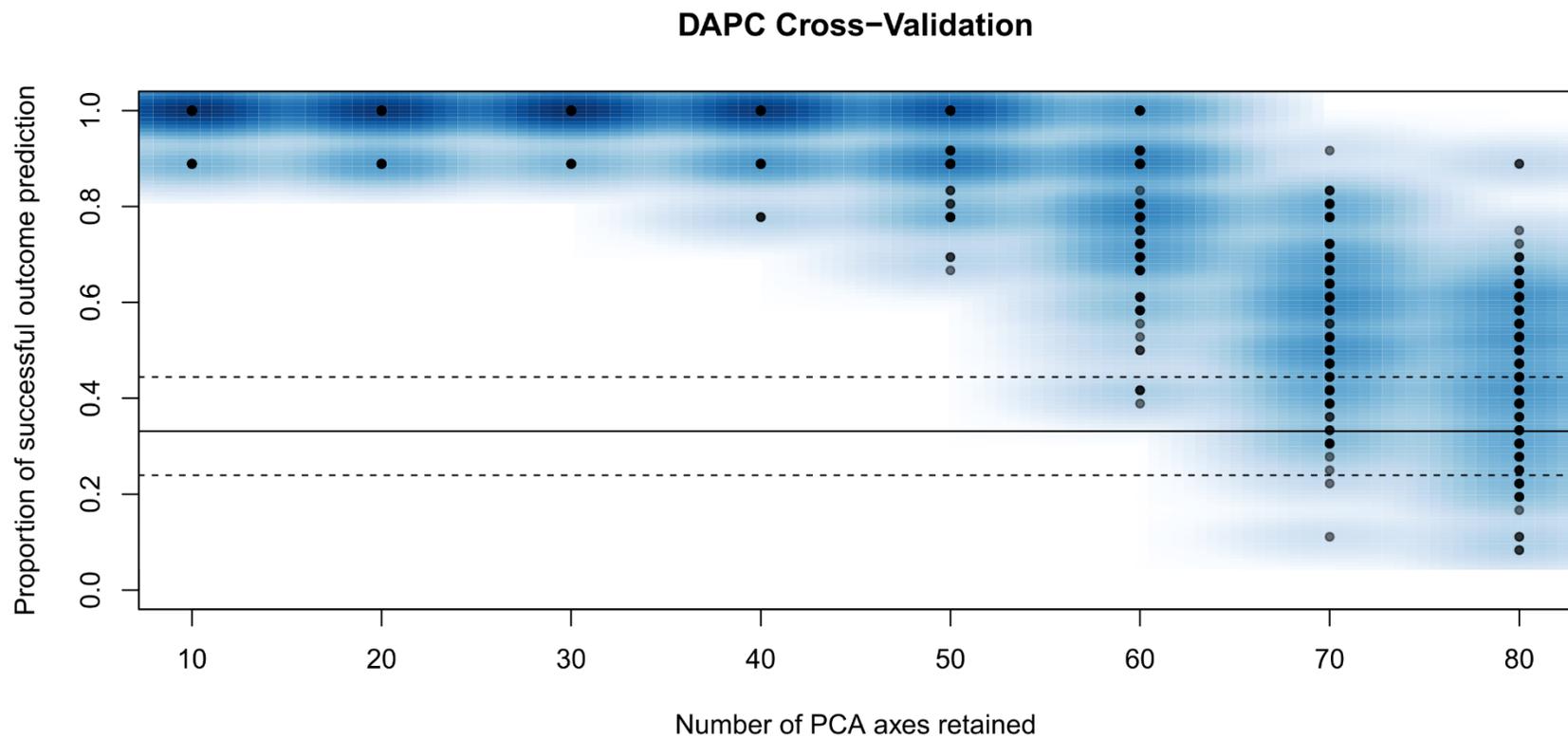


Figure A2: Graphic showing the result of the cross-validation procedure that aimed to identify the number of PCA components that should be retained for the DAPC (i.e., the number of components that maximize successful individual assignment to the k clusters based on 150 iterations). The cross-validation here suggests that 30 PCA components should be retained.

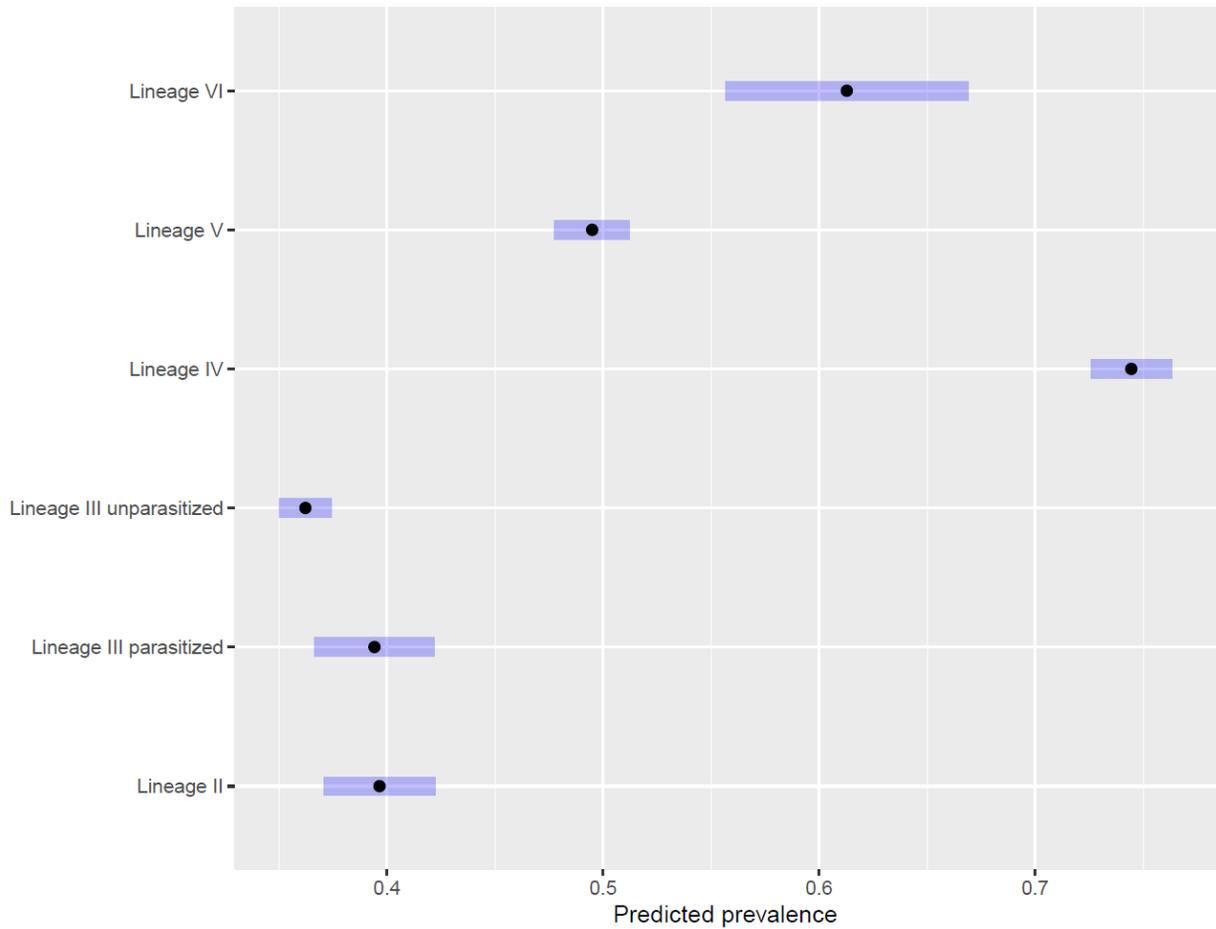


Figure A3: Forest plot showing the results of the contrast tests between predicted prevalence in each *Leuciscus* lineage. Dots refer to mean predicted prevalence and bar shows confidence intervals. Non-overlapping confidence intervals means that the difference between predicted prevalence in different lineages is significant.

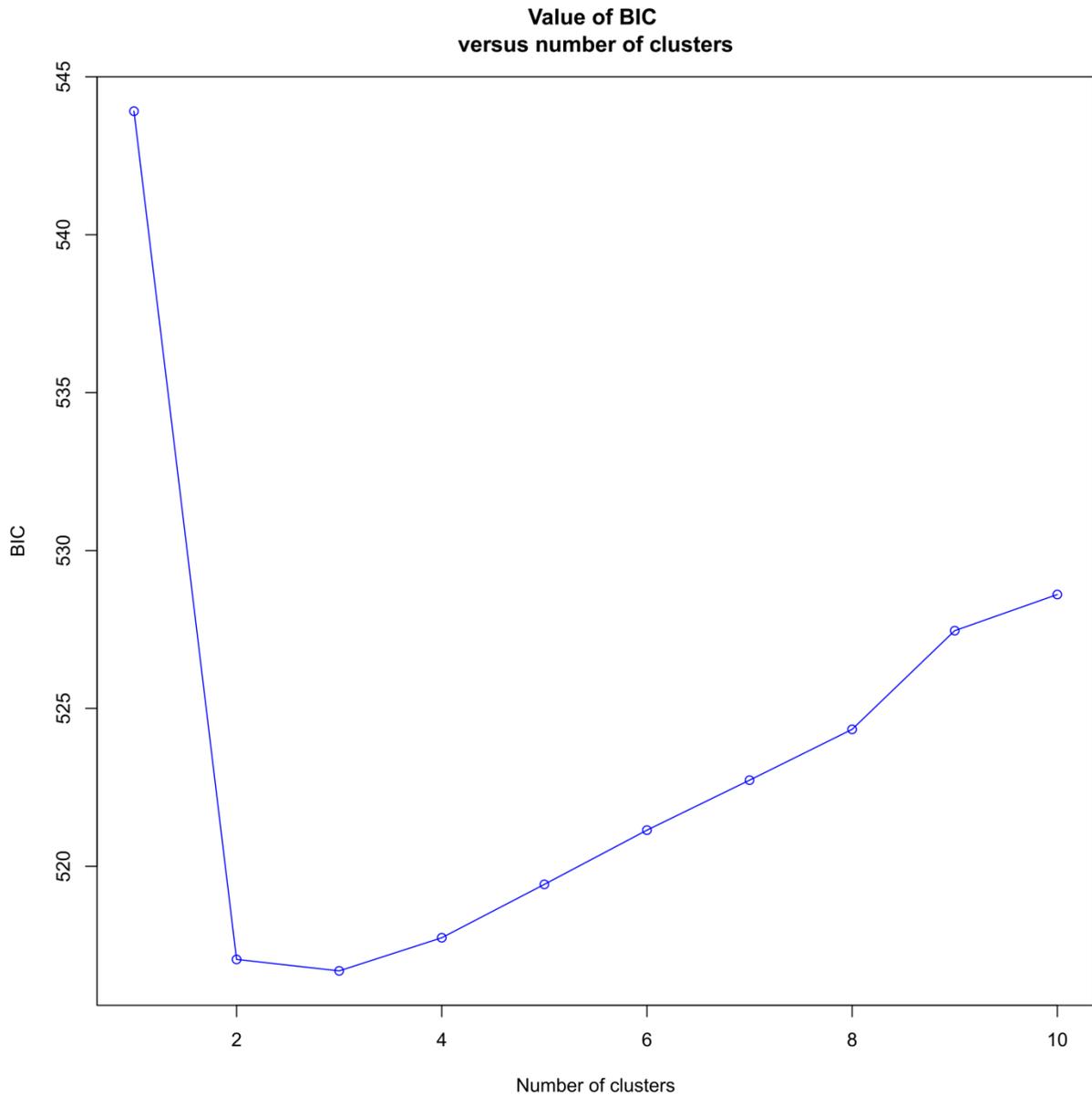


Figure A4: Evolution of the Bayesian Information Criterion (BIC) according to the number of clusters retained in the Discriminant Analysis of Principal Component (DAPC). The lowest BIC (here $k=3$) indicates the number of clusters that is the more likely to describe the input dataset (here allelic frequencies at 2 543 Single Nucleotide Polymorphisms).

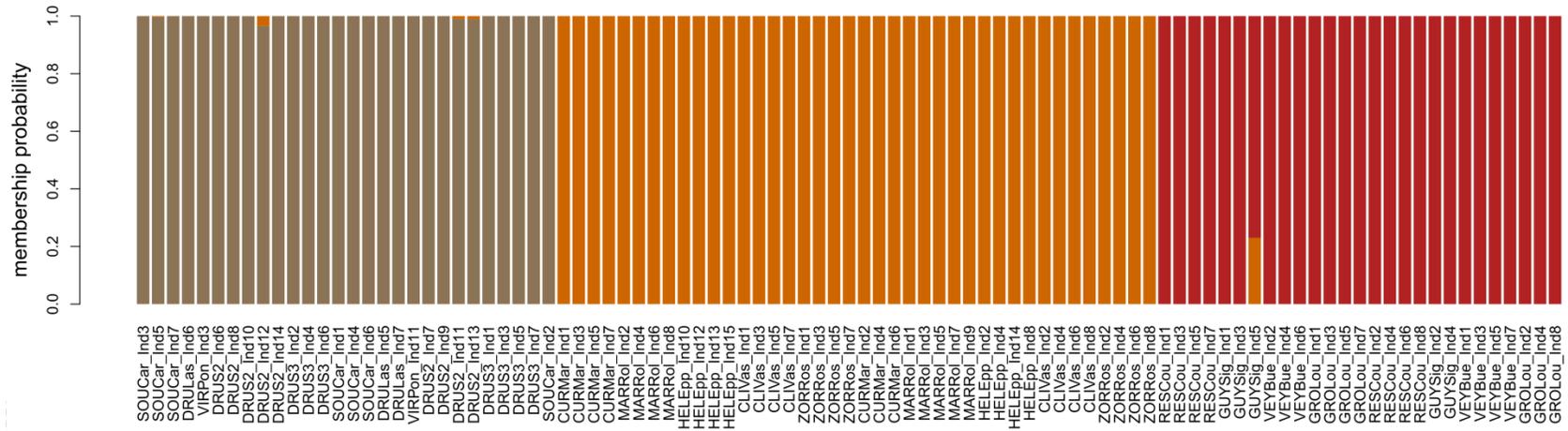


Figure A5: Graphic showing the individual assignment to the three clusters revealed by the DAPC. Each individual is represented by a vertical line that is partitioned into three colors that represent its estimated membership fractions in $k=3$ clusters. X-labels indicate for each of the individual its sampling site (eg, SOUCar) and its ID (eg., Ind3).

Table A1: Summary of the two Principal Component Analyses realized on hydrographic variables and landscape variables. Correlation of each raw variable with the two first components is displayed (PC1 and PC2) as well as the range of each raw variable (Range). Strongest correlation between each variable and either PC1 or PC2 are bolded.

Category	Variables	PC1	PC2	Range
Hydrographic	Altitude	0.002	-0.752	3-571 m
	Slope	-0.191	-0.557	0-48 ‰
	Distance from the source	0.920	0.055	3-394 km
	Area of the upstream watershed	0.975	0.041	7-20 740 km ²
	Strahler index	0.715	-0.059	1-6
	River width	0.982	-0.008	1-153 m
	River depth	0.908	-0.005	0-2 m
	Average size of the bed sediments	-0.221	0.774	-7-3
	Minimum average monthly river flow	0.848	0.023	3 436 m ³ /s
Average natural river flow	0.953	0.011	0-511 m ³ /s	
Landscape	Artificialized lands	-0.262	0.706	0%-25%
	Agricultural lands	-0.964	-0.193	0%-96%
	Forest and semi-natural lands	0.991	-0.057	0%-90%
	Wet lands and water bodies	0.072	0.765	0%-17%

Table A2: Table summarizing coefficient estimates, standard errors (SE) and associated t-value for the SDM model fitted on the dataset covering the current range of *T. polycolpus* (N=128 sites).

	Estimate	SE	t-value
(Intercept)	4.205	3.7E+00	1.139
Expected Heterozygosity	0.607	1.4E+00	0.444
Dace body length	0.053	1.9E-02	2.815
(Dace body length) ²	-1.1063E-04	4.900E-05	-2.190
Species identity (<i>L. leuciscus</i>)	-2.118	6.1E-01	-3.493
Number of hosts species	-0.247	2.7E-01	-0.920
Fish species Richness	-0.063	6.9E-02	-0.919
Water temperature	-0.548	2.5E-01	-2.177
Hydrology	-0.284	1.6E-01	-1.829
Topology	0.264	2.3E-01	1.150
Landscape 1	-0.551	1.7E-01	-3.289
Landscape 2	-0.302	2.4E-01	-1.268
Dams density	-0.011	1.7E-02	-0.686
Weirs density	0.008	1.6E-02	0.518
Weir position (none)	-0.812	6.6E-01	-1.227
Weirs position (upstream)	-1.559	6.7E-01	-2.336
Weirs position (upstream and downstream)	-1.680	6.9E-01	-2.420

Table A3: Table synthesizing information for SNPs showing a significant association ($\alpha = 0.05$) with parasite prevalence measured at the individual level (GWAS approach). We displayed for each SNP, its name (number of the loci_position of the SNP) and the outputs from an analysis of deviance: F statistic, degree of freedom, raw p-value and adjusted p-value using a false discovery rate approach (FDR).

Loci	F value	Df	pvalue	FDR
30365_13	18.07E8	61,2	<0.001	<0.001
31247_45	39.792	61,2	<0.001	<0.001
20735_106	24.223	59,2	<0.001	<0.001
6830_17	33.182	60,1	<0.001	<0.001
9706_73	17.988	63,2	<0.001	<0.001
9473_86	30.338	60,1	<0.001	<0.001
14042_124	16.998	60,2	<0.001	<0.001
15789_21	16.215	61,2	<0.001	0.001
23160_22	15.267	59,2	<0.001	0.001
14397_60	15.123	59,2	<0.001	0.001
3163_102	14.005	63,2	<0.001	0.002
31547_125	13.688	66,2	<0.001	0.002
18121_75	13.444	68,2	<0.001	0.002
12378_47	13.293	65,2	<0.001	0.002
3881_89	13.296	59,2	<0.001	0.002
18955_50	13.033	65,2	<0.001	0.002
451_130	12.841	59,2	<0.001	0.003
32965_13	12.119	60,2	<0.001	0.004
19792_54	11.949	65,2	<0.001	0.004
6084_110	11.866	60,2	<0.001	0.004
22427_103	11.781	62,2	<0.001	0.004
15517_29	11.734	60,2	<0.001	0.005
2126_86	11.676	59,2	<0.001	0.005
20777_89	11.505	59,2	<0.001	0.005
9391_96	10.832	71,2	<0.001	0.006
26600_107	17.674	67,1	<0.001	0.006
26536_32	10.694	62,2	<0.001	0.007
23585_10	10.423	63,2	<0.001	0.008
14791_89	16.833	61,1	<0.001	0.008
6225_83	10.435	60,2	<0.001	0.008
8863_98	10.217	65,2	<0.001	0.009
567_75	10.281	60,2	<0.001	0.009

32434_98	10.196	59,2	<0.001	0.009
13572_22	10.083	62,2	<0.001	0.01
25567_94	10.047	63,2	<0.001	0.01
4249_123	9.975	64,2	<0.001	0.01
33337_20	9.830	66,2	<0.001	0.01
41022_86	9.595	65,2	<0.001	0.012
31207_90	9.617	60,2	<0.001	0.012
18497_117	9.490	64,2	<0.001	0.012
1055_61	9.440	63,2	<0.001	0.013
10039_82	9.232	67,2	<0.001	0.014
25585_44	9.324	62,2	<0.001	0.014
17689_63	9.328	61,2	<0.001	0.014
16758_77	9.109	65,2	<0.001	0.015
41153_8	8.982	63,2	<0.001	0.016
28491_38	8.500	66,2	0.001	0.021
28298_104	8.503	64,2	0.001	0.021
2909_66	8.524	61,2	0.001	0.021
13326_67	13.195	60,1	0.001	0.022
12878_52	8.421	60,2	0.001	0.023
17321_81	8.388	60,2	0.001	0.023
644_69	8.332	62,2	0.001	0.023
25467_120	13.003	61,1	0.001	0.024
32865_92	8.332	59,2	0.001	0.024
11384_48	8.242	60,2	0.001	0.025
13842_7	12.781	61,1	0.001	0.025
20915_81	8.190	61,2	0.001	0.025
13913_80	12.550	61,1	0.001	0.026
6311_24	7.976	63,2	0.001	0.027
11820_144	7.907	63,2	0.001	0.028
19016_53	12.279	61,1	0.001	0.028
7497_72	7.903	61,2	0.001	0.029
34614_12	12.074	66,1	0.001	0.029
33082_120	7.768	62,2	0.001	0.03
16630_129	7.777	61,2	0.001	0.03
1972_142	7.726	63,2	0.001	0.03
19731_124	7.768	60,2	0.001	0.03
33374_128	7.625	67,2	0.001	0.031
725_116	7.667	59,2	0.001	0.032
9118_144	7.647	59,2	0.001	0.033
25353_20	7.579	59,2	0.001	0.034
9271_55	7.524	62,2	0.001	0.034
4594_43	7.367	60,2	0.001	0.037
28712_59	7.350	59,2	0.001	0.038

25438_128	7.238	63,2	0.001	0.039
15286_61	7.236	63,2	0.001	0.039
29363_50	7.188	63,2	0.002	0.04
1276_70	7.216	59,2	0.002	0.04
5853_56	7.123	65,2	0.002	0.041
12705_36	7.116	65,2	0.002	0.041
2752_120	7.138	60,2	0.002	0.042
3737_143	6.991	71,2	0.002	0.042
10393_72	7.044	59,2	0.002	0.044
33235_100	6.979	64,2	0.002	0.044
29939_90	10.594	64,1	0.002	0.044
16343_28	6.943	61,2	0.002	0.045
11938_12	6.748	65,2	0.002	0.049
27398_144	10.239	61,1	0.002	0.049
26335_72	6.748	64,2	0.002	0.05
