

Supporting Material for “Bayesian clustering algorithms ascertaining spatial population structure: A new computer program and a comparison study” by Chen et al. 2007.

All results for the 5 island data sets, including EM runs and several values of the interaction parameter  $\psi$  for TESS and GENECLUST are available from the following url:

<http://www-timc.imag.fr/Olivier.Francois/all.5islands.txt>

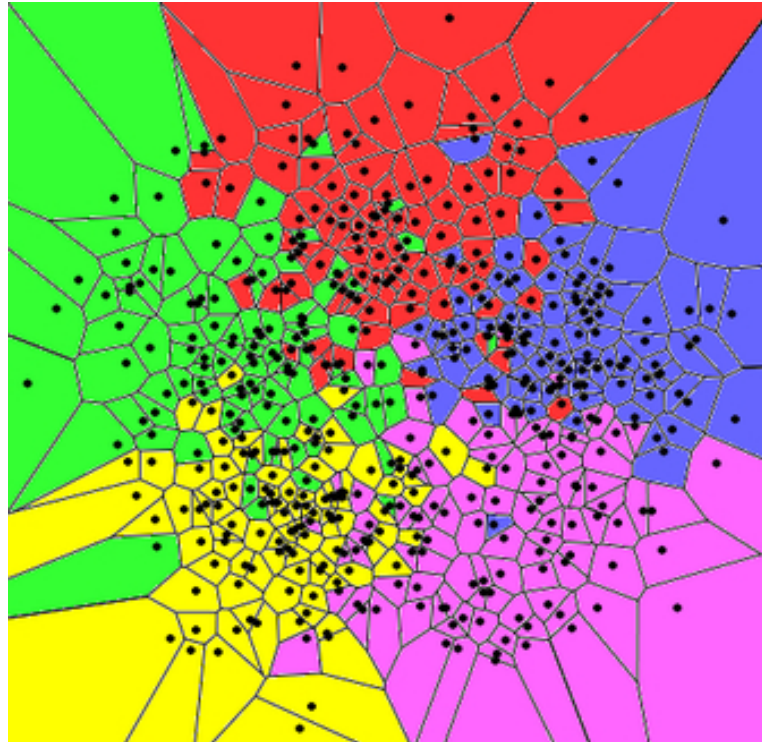


Figure SM1. Ground truth for one 5 island data set. (500 individuals, 20 % spatial overlap). Spatial coordinates for each genetic cluster were simulated from standard Gaussian distributions. Each color represents a population of origin in the simulation.

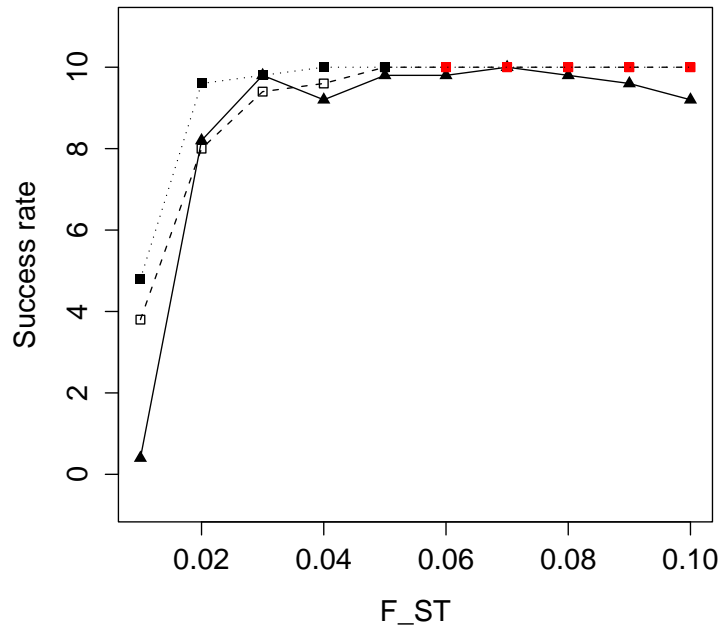


Figure SM2. Five island simulation results (500 individuals, 20 % spatial overlap, 10 unlinked codominant loci,  $F_{ST}$  ranging from 0.01 to 0.10). Number of success over ten replicated runs for GENELAND: filled triangles, TESS: filled squares, GENECLUST: empty squares (hidden by the filled squares). The results are for  $K_{max} = K = 5$ , i.e. the number of populations is assumed to be correct.

**Table SM1.** Five island model. (500 individuals, 20 % spatial overlap, 10 unlinked codominant loci,  $F_{ST}$  ranging from 0.02 to 0.05). Individual assignment data for GENELAND and GENECLUST. The results are for  $K_{\max} = K = 5$ , i.e. the number of populations is assumed to be correct.

Software	$F_{ST}$	Data set	Total correct member. probab.	Avg % misassigned	Number of successful runs
GENELAND $\lambda = 100$	0.02	1	0.844	14.0	9
		2	0.857	12.8	8
		3	0.876	10.2	9
		4	0.855	12.4	8
		5	0.855	12.4	7
		Avg	0.857	12.4	8.2
	0.03	1	0.898	9.4	10
		2	0.910	8.2	10
		3	0.895	9.4	10
		4	0.902	8.8	10
		5	0.887	10.6	9
		Avg	0.898	9.3	9.8
	0.04	1	0.903	9.0	10
		2	0.906	8.6	7
		3	0.902	9.8	10
		4	0.899	9.2	9
		5	0.934	6.0	10
		Avg	0.909	8.5	9.2
	0.05	1	0.929	6.8	10
		2	0.921	7.4	9
		3	0.935	5.8	10
		4	0.915	8.4	10
		5	0.933	6.2	10
		Avg	0.927	6.9	9.8
	GENECLUST $\psi = 0.6$	0.02	1	0.845	9.8
2			0.838	9.2	7
3			0.870	6.0	10
4			0.830	8.0	9
5			0.844	9.4	8
Avg			0.845	8.5	8.0
0.03		1	0.937	3.8	9
		2	0.925	4.0	9
		3	0.940	3.2	9
		4	0.937	3.0	10
		5	0.919	3.0	10
		Avg	0.932	3.4	9.4
0.04		1	0.980	0.4	10
		2	0.959	2.2	8
		3	0.983	0.6	10
		4	0.967	1.8	10
		5	0.974	1.2	10
		Avg	0.973	1.2	9.6
0.05		1	0.995	0.2	10
		2	0.992	0.2	10
		3	0.965	0.4	10
		4	0.990	0.8	10
		5	0.989	0.4	10
		Avg	0.986	0.4	10.0

**Table SM2.** Five island model. (500 individuals, 20 % spatial overlap, 10 unlinked codominant loci,  $F_{ST}$  ranging from 0.02 to 0.05). Individual assignment data for TESS. The results are for  $K_{\max} = K = 5$ , i.e. the number of populations is assumed to be correct.

Software	$F_{ST}$	Data set	Total correct member. probab.	Avg % misassigned	Number of successful runs
TESS (MCMC) $\psi = 0.6$	0.02	1	0.848	10.4	10
		2	0.806	10.2	10
		3	0.892	5.8	10
		4	0.873	7.4	8
		5	0.860	8.8	10
		Avg	0.856	8.5	9.6
	0.03	1	0.938	4.0	10
		2	0.930	4.2	10
		3	0.945	3.2	10
		4	0.940	3.0	9
		5	0.950	3.2	10
		Avg	0.941	3.5	9.8
	0.04	1	0.983	0.4	10
		2	0.966	1.8	10
		3	0.985	0.6	10
		4	0.976	1.4	10
		5	0.977	1.2	10
		Avg	0.977	1.1	10.0
	0.05	1	0.997	0.2	10
		2	0.994	0.2	10
		3	0.994	0.4	10
		4	0.991	0.8	10
		5	0.993	0.4	10
		Avg	0.994	0.4	10.0
	TESS (EM) $\psi = 0.6$	0.02	1	0.876	10.8
2			0.885	9.8	7
3			0.919	6.6	9
4			0.902	7.4	8
5			0.887	9.4	7
Avg			0.894	8.8	7.8
0.03		1	0.954	4.0	8
		2	0.950	3.6	7
		3	0.963	3.0	7
		4	0.962	3.0	8
		5	0.962	3.4	10
		Avg	0.958	3.4	8.0
0.04		1	0.993	0.4	8
		2	0.978	2.0	5
		3	0.993	0.4	8
		4	0.987	1.2	7
		5	0.987	1.2	7
		Avg	0.988	1.0	7.0
0.05		1	0.998	0.2	10
		2	0.998	0.0	6
		3	0.996	0.4	7
		4	0.993	0.8	8
		5	0.995	0.4	10
		Avg	0.996	0.4	8.2

**Table SM3.** Five island model – Detecting the correct number of populations (500 individuals, 20 % spatial overlap, 10 unlinked codominant loci,  $F_{ST}$  ranging from 0.02 to 0.05). Individual assignment data for GENELAND and GENECLUST. The results are for  $K_{\max} = 6$ , i.e. the number of populations is assumed to be unknown.



Software	$F_{ST}$	Data set	Total correct member. probab.	Avg % misassigned	Number of successful runs	
GENELAND $\lambda = 100$	0.02	1	0.427	14.0	4	
		2	0.283	16.4	4	
		3	0.294	13.8	3	
		4	0.598	15.2	1	
		5	0.762	13.4	3	
		Avg	0.473	14.6	3.0	
	0.03	1	-	-	-	
		2	0.410	10.6	1	
		3	0.374	10.0	4	
		4	0.571	9.8	5	
		5	0.423	8.4	4	
		Avg	0.445	9.7	3.5	
	0.04	1	0.308	9.0	1	
		2	0.450	11.4	3	
		3	0.649	11.4	4	
		4	0.787	7.0	2	
		5	0.796	10.4	3	
		Avg	0.598	9.8	2.6	
	GENECLUST $\psi = 0.6$	0.02	1	0.802	12.0	10
			2	0.812	10.6	10
3			0.854	7.0	10	
4			0.825	9.2	10	
5			0.816	11.0	10	
Avg			0.822	10.0	10.0	
0.03		1	0.915	4.6	10	
		2	0.901	5.8	10	
		3	0.914	4.6	10	
		4	0.915	3.6	10	
		5	0.922	3.6	10	
		Avg	0.913	4.4	10.0	
0.04		1	0.968	0.6	10	
		2	0.954	2.2	10	
	3	0.974	0.6	10		
	4	0.962	2.0	10		
	5	0.967	1.6	10		
	Avg	0.965	1.4	10		

**Table SM4.** Five island model – Detecting the correct number of populations (500 individuals, 20 % spatial overlap, 10 unlinked codominant loci,  $F_{ST}$  ranging from 0.02 to 0.05). Individual assignment data for TESS. The results are for  $K_{\max} = 6$ , i.e. the number of populations is assumed to be unknown.

Software	$F_{ST}$	Data set	Total correct member. probab.	Avg % misassigned	Number of successful runs	
TESS (MCMC) $\psi = 0.6$	0.02	1	0.813	13.4	10	
		2	0.814	11.6	10	
		3	0.857	7.6	10	
		4	0.838	9.8	10	
		5	0.816	12.6	10	
		Avg	0.828	11.0	10.0	
	0.03	1	0.918	4.6	10	
		2	0.905	5.6	10	
		3	0.920	5.2	10	
		4	0.918	3.4	10	
		5	0.923	5.2	10	
		Avg	0.917	4.8	10.0	
	0.04	1	0.969	0.4	10	
		2	0.953	1.8	10	
		3	0.975	0.6	10	
		4	0.962	2.0	10	
		5	0.965	2.0	10	
		Avg	0.965	1.4	10.0	
	TESS (EM) $\psi = 0.6$	0.02	1	0.839	14.6	10
			2	0.849	13.4	9
3			0.874	11.2	10	
4			0.858	12.4	10	
5			0.851	13.4	10	
Avg			0.854	13.0	9.8	
0.03		1	0.915	8.0	10	
		2	0.921	6.8	9	
		3	0.929	6.4	10	
		4	0.922	7.2	10	
		5	0.928	7.0	10	
		Avg	0.923	7.1	9.8	
0.04		1	0.962	3.4	9	
		2	0.949	4.8	9	
		3	0.963	3.4	10	
		4	0.962	3.8	10	
		5	0.966	3.0	10	
		Avg	0.960	3.7	10	

**Table SM5.** Two island model. Individual assignment data for STRUCTURE obtained for various levels of geographical distance (200 individuals, 20 unlinked codominant loci,  $F_{ST}$  around 0.02). The geographical connectivity between the two subpopulations varies from 2% to 40%. The results are for  $K_{\max} = 3$ , i.e. the number of populations is unknown.

Software	$D$	$F_{ST}$	membership probab.	Avg % misassigned	Number of success (/5)
STRUCTURE	0.5	0.015	0.885	7.5	5
		0.018	0.926	5.5	5
		0.019	0.907	3.0	4
		0.020	0.950	2.5	5
		0.023	0.940	4.0	5
		Avg	0.922	4.5	4.8
	1.0	0.015	0.880	10.0	5
		0.016	0.905	5.5	5
		0.019	0.923	5.0	5
		0.020	0.901	6.5	5
		0.022	0.922	4.5	5
	Avg	0.906	6.3	5.0	
	1.5	0.019	0.897	6.5	5
		0.020	0.921	4.5	5
		0.021	0.936	4.0	5
		0.022	0.924	4.5	5
		0.023	0.923	4.5	5
	Avg	0.920	4.8	5.0	
	2.0	0.018	0.942	2.5	5
		0.020	0.927	6.5	5
		0.021	0.940	3.5	5
		0.022	0.934	4.5	5
		0.027	0.973	0.5	5
	Avg	0.943	3.5	5.0	
	3.0	0.015	0.849	11.5	5
		0.016	0.899	4.5	5
		0.018	0.925	5.0	5
		0.019	0.941	4.0	5
0.020		0.922	4.5	5	
Avg	0.907	5.9	5.0		
4.0	0.016	0.882	7.0	5	
	0.017	0.915	6.0	5	
	0.018	0.891	8.5	5	
	0.019	0.849	13.5	5	
	0.021	0.935	4.0	5	
Avg	0.894	7.8	5.0		

**Table SM6.** Two island model. Individual assignment data for GENELAND obtained for various levels of geographical distance (200 individuals, 20 unlinked codominant loci,  $F_{ST}$  around 0.02). The geographical connectivity between the two subpopulations varies from 2% to 40%. The results are for  $K_{\max} = 3$ , i.e. the number of populations is unknown.

Software	$D$	$F_{ST}$	membership probab.	Avg % misassigned	Number of success (/5)
GENELAND	0.5	0.015	–	–	0
		0.018	–	–	0
		0.019	–	–	0
		0.020	–	–	0
		0.023	–	–	0
		Avg	–	–	0
	1.0	0.015	–	–	0
		0.016	–	–	0
		0.019	–	–	0
		0.020	–	–	0
		0.022	–	–	0
		Avg	–	–	0
	1.5	0.019	–	–	0
		0.020	0.910	8.0	2
		0.021	0.870	10.0	1
		0.022	0.910	8.0	2
		0.023	0.880	9.0	2
		Avg	0.893	8.8	1.4
	2.0	0.018	0.900	9.0	2
		0.020	0.950	4.0	4
		0.021	0.900	9.0	3
		0.022	0.910	6.0	3
		0.027	0.950	4.0	3
		Avg	0.922	6.4	3.0
	3.0	0.015	0.920	7.0	1
		0.016	0.460	6.0	1
		0.018	0.390	2.0	4
		0.019	0.960	3.0	4
0.020		0.960	2.0	4	
Avg		0.738	4.0	2.8	
4.0	0.016	0.380	1.0	4	
	0.017	0.510	2.0	4	
	0.018	0.460	0.0	3	
	0.019	0.510	1.0	3	
	0.021	0.370	0.0	3	
	Avg	0.446	0.8	3.4	

**Table SM7.** Two island model. Individual assignment data for TESS obtained for various levels of geographical distance (200 individuals, 20 unlinked codominant loci,  $F_{ST}$  around 0.02). The geographical connectivity between the two subpopulations varies from 2% to 40%. The results are for  $K_{\max} = 3$ , i.e. the number of populations is unknown.



Software	$D$	$F_{ST}$	membership probab.	Avg % misassigned	Number of success (/5)
TESS	0.5	0.015	–	–	0
		0.018	–	–	0
		0.019	–	–	0
		0.020	0.927	5.0	4
		0.023	0.926	7.5	2
		Avg	0.927	6.3	1.2
	1.0	0.015	–	–	0
		0.016	0.912	6.5	3
		0.019	0.932	5.0	1
		0.020	0.896	8.0	3
		0.022	0.926	5.0	4
		Avg	0.917	6.1	2.2
	1.5	0.019	0.922	6.5	2
		0.020	0.922	5.0	4
		0.021	0.947	3.0	4
		0.022	0.955	2.5	5
		0.023	0.919	4.5	2
		Avg	0.933	4.3	3.4
	2.0	0.018	0.968	2.0	4
		0.020	0.960	2.5	4
		0.021	0.973	1.0	3
		0.022	0.968	0.5	3
		0.027	0.985	1.0	5
		Avg	0.971	1.4	3.8
	3.0	0.015	0.944	2.5	3
		0.016	0.973	1.5	3
		0.018	0.978	1.5	4
		0.019	0.969	2.5	3
0.020		0.967	3.0	4	
Avg		0.966	2.2	3.4	
4.0	0.016	0.969	1.5	3	
	0.017	0.979	0.0	4	
	0.018	0.978	0.5	4	
	0.019	0.977	0.5	4	
	0.021	0.985	0.5	5	
	Avg	0.978	0.6	4.0	

**Table SM8.** Two island model. Summary. Individual assignment data for STRUCTURE, GENELAND and TESS, averages obtained using 5 data sets for various levels of geographical distance (200 individuals, 20 unlinked codominant loci,  $F_{ST}$  around 0.02). The geographical connectivity between the two subpopulations varies from 2% to 40%. The results are for  $K_{\max} = 3$ , i.e. the number of populations is unknown.

Software	$D$	Membership probab.	Avg % misassigned	Number of successful runs
STRUCTURE	0.5	0.922	4.5	4.8
	1.0	0.906	6.3	5.0
	1.5	0.920	4.8	5.0
	2.0	0.943	3.5	5.0
	3.0	0.907	5.9	5.0
	4.0	0.894	7.8	5.0
GENELAND	0.5	-	-	0
	1.0	-	-	0
	1.5	0.893	8.8	1.4
	2.0	0.922	6.4	3.0
	3.0	0.738	4.0	2.8
	4.0	0.446	0.8	3.4
TESS	0.5	0.927	6.3	1.2
	1.0	0.917	6.1	2.2
	2.0	0.971	1.4	3.8
	3.0	0.966	2.2	3.4
	4.0	0.978	0.6	4.0