

生物資源モデリング

北門 利英（海洋生物資源学科）

Lecture 7 遺伝的混合と遺伝的汚染のモデリング

Contents

1. 水産資源集団の混合率推定における典型的な問題
 2. 外来集団の混合率の推定 (Mixture model)
 3. 外来集団の遺伝的混合の推定 (Admixture model)
 4. 有明海アサリへの適用例
- 補足： 個体のアサインメント

1. 水産資源集団の混合率推定における典型的な問題

水産資源の持続的利用と管理

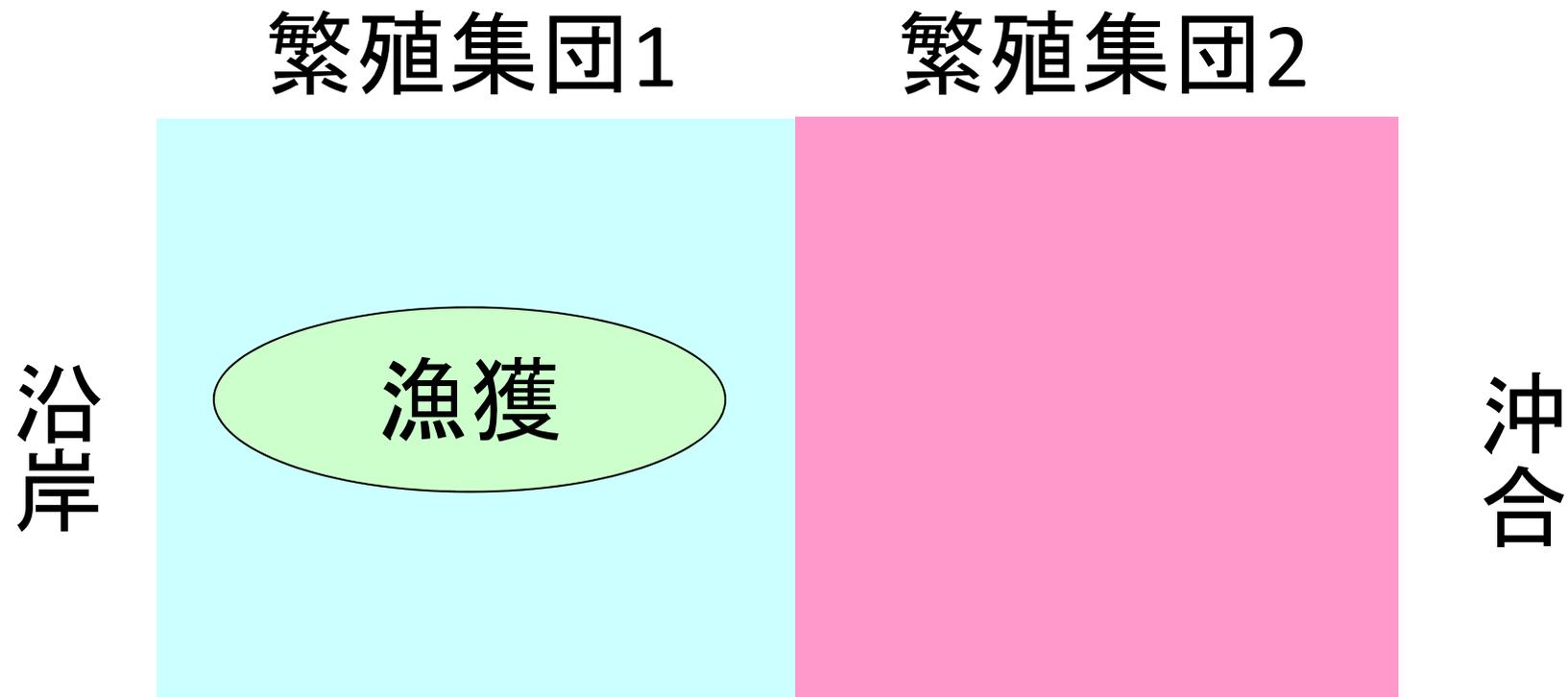
過剰漁獲や生息環境の悪化

⇒水産資源の資源量や漁獲量の減少

持続的な水産資源の利用のために

- 資源評価(資源量や資源変動メカニズムの推測)
- 漁業に対する規制等による資源管理
(漁獲量限度, 禁漁区, 禁漁期などの設定)
- 放流事業(水産増殖)
 - 人工的に栽培した稚魚の放流し, 漁獲量増加を図る
 - 産卵親魚として次世代への再生産も期待

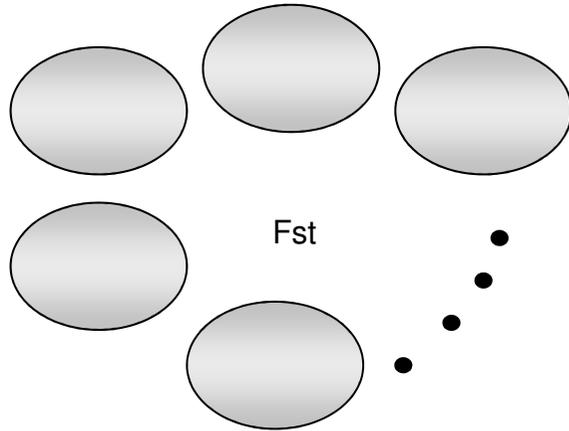
集団構造と資源管理



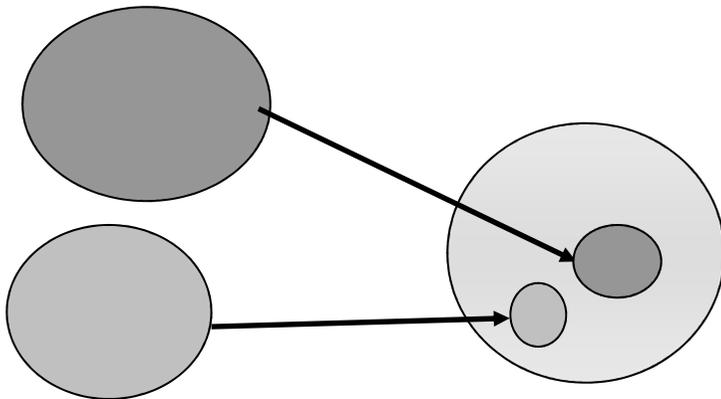
資源量や資源動体に加えて，
資源の集団構造の把握が重要

様々な集団構造

集団間の遺伝的差異

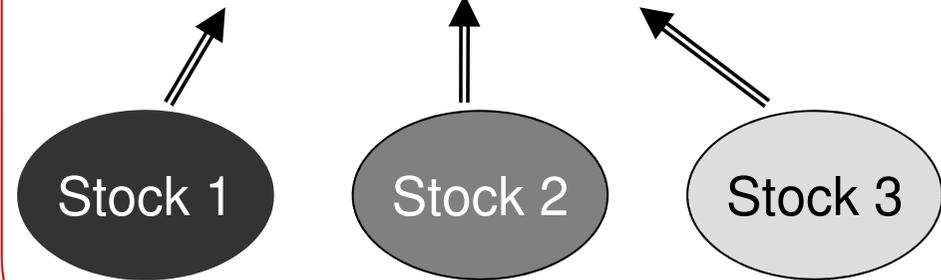
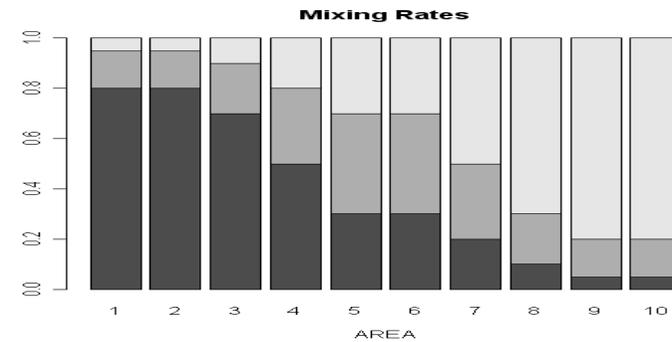


混入割合の推定



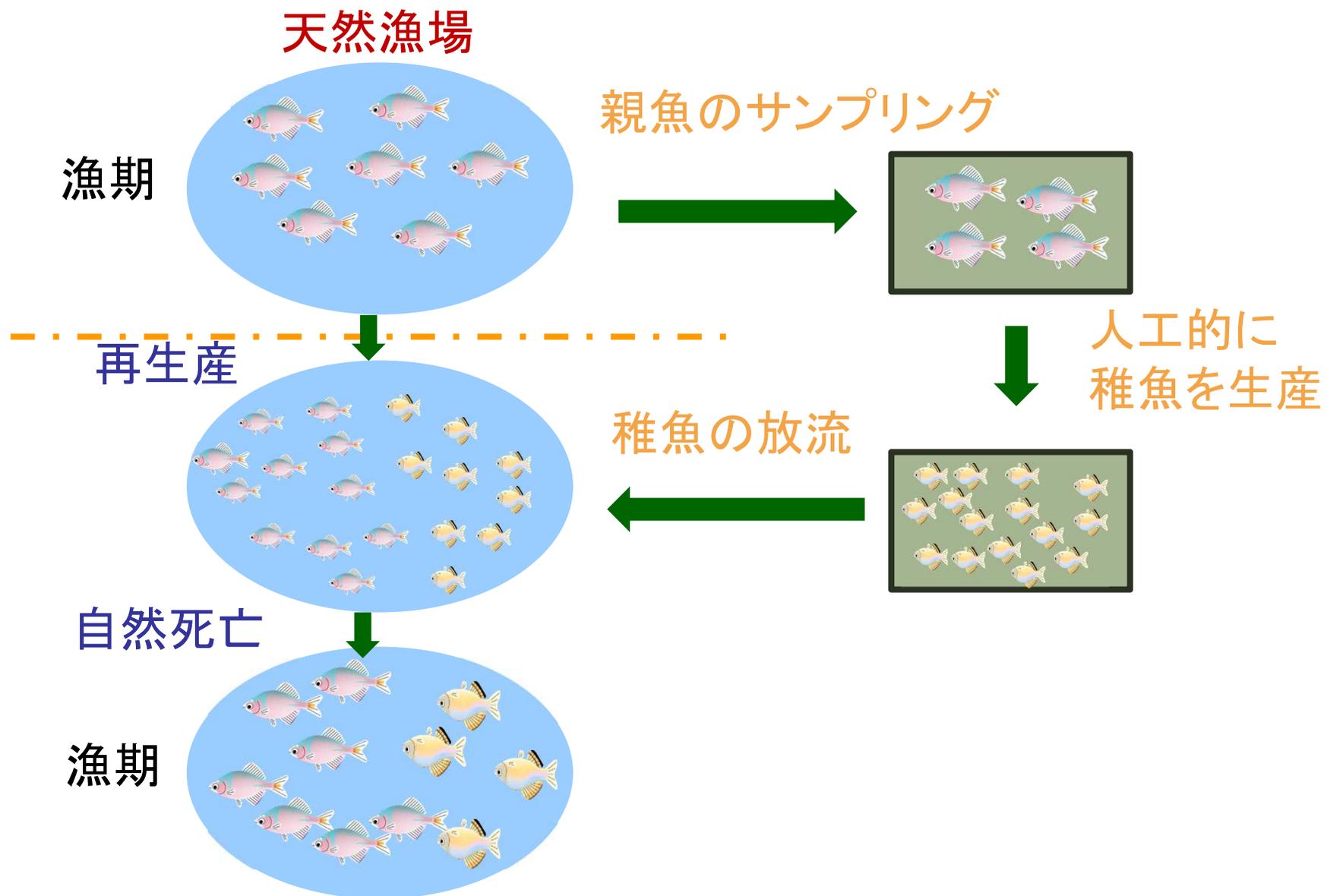
集団の混合率の推定

摂餌域あるいは回遊経路での混合率



繁殖域 (未知)

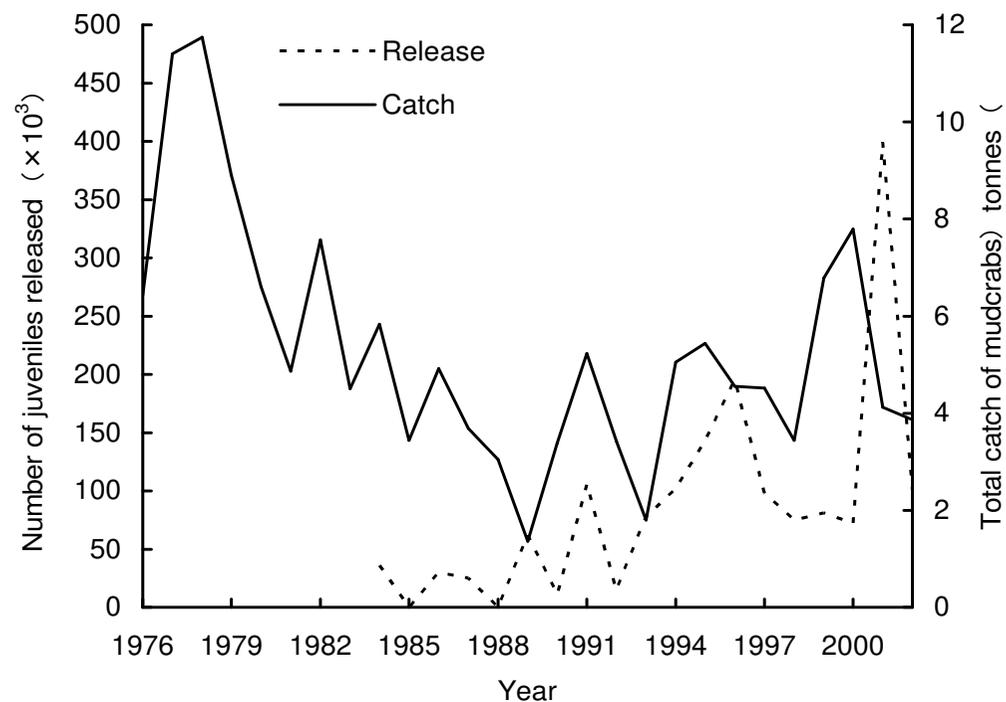
水産増殖における放流集団の混合



例：ノコギリガザミの放流事業



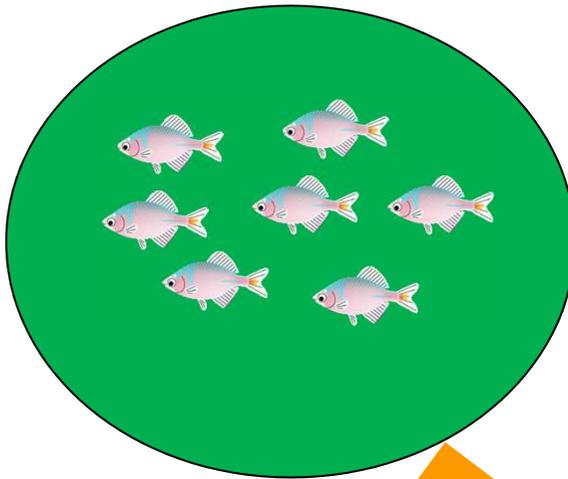
浦戸湾漁獲量と放流数の推移



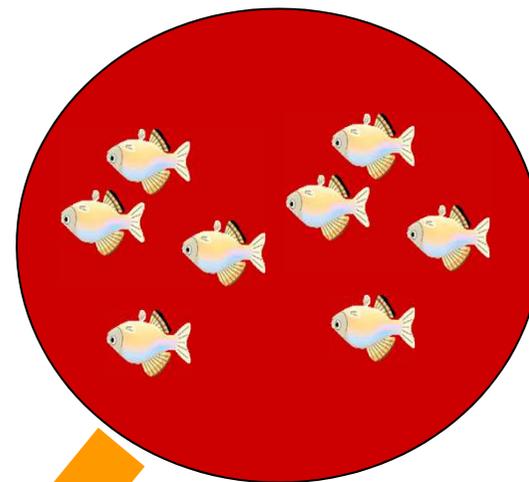
- 放流が漁獲量に貢献しているか？
- 放流個体の混合率はどの程度か？

混合率の推定の典型的な例

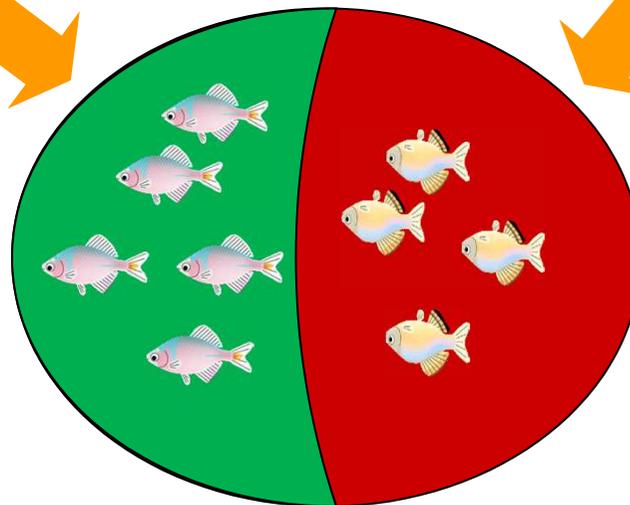
基準群 1



基準群 2



混合群

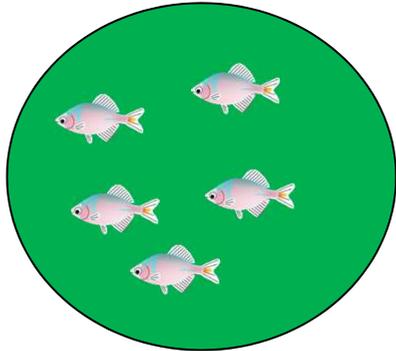


しばしば人工的な
標識の利用が難しい

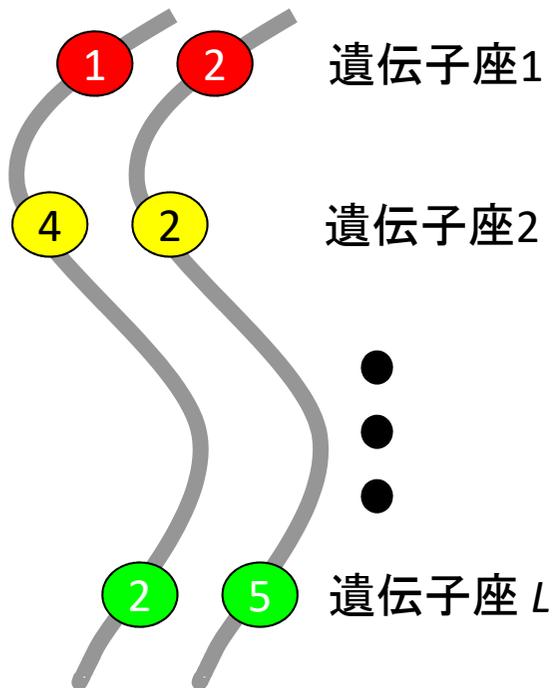


遺伝標識の利用

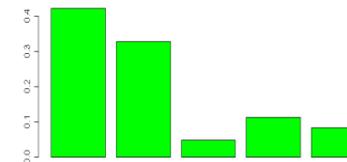
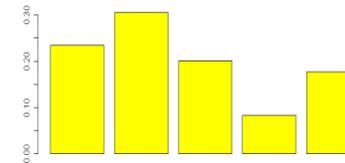
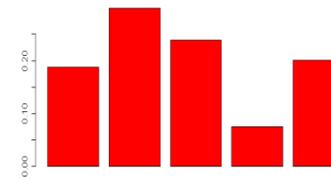
遺伝情報の観測



個体の遺伝情報



集団の対立遺伝子頻度

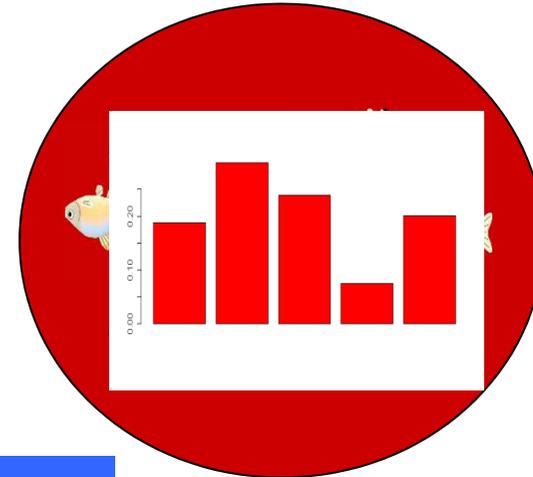
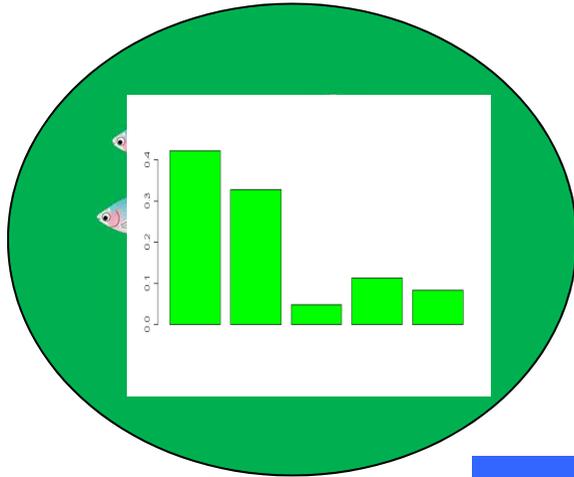


異なる集団間では
対立遺伝子頻度が異なる

混合率の推定

基準群 1 の対立遺伝子頻度

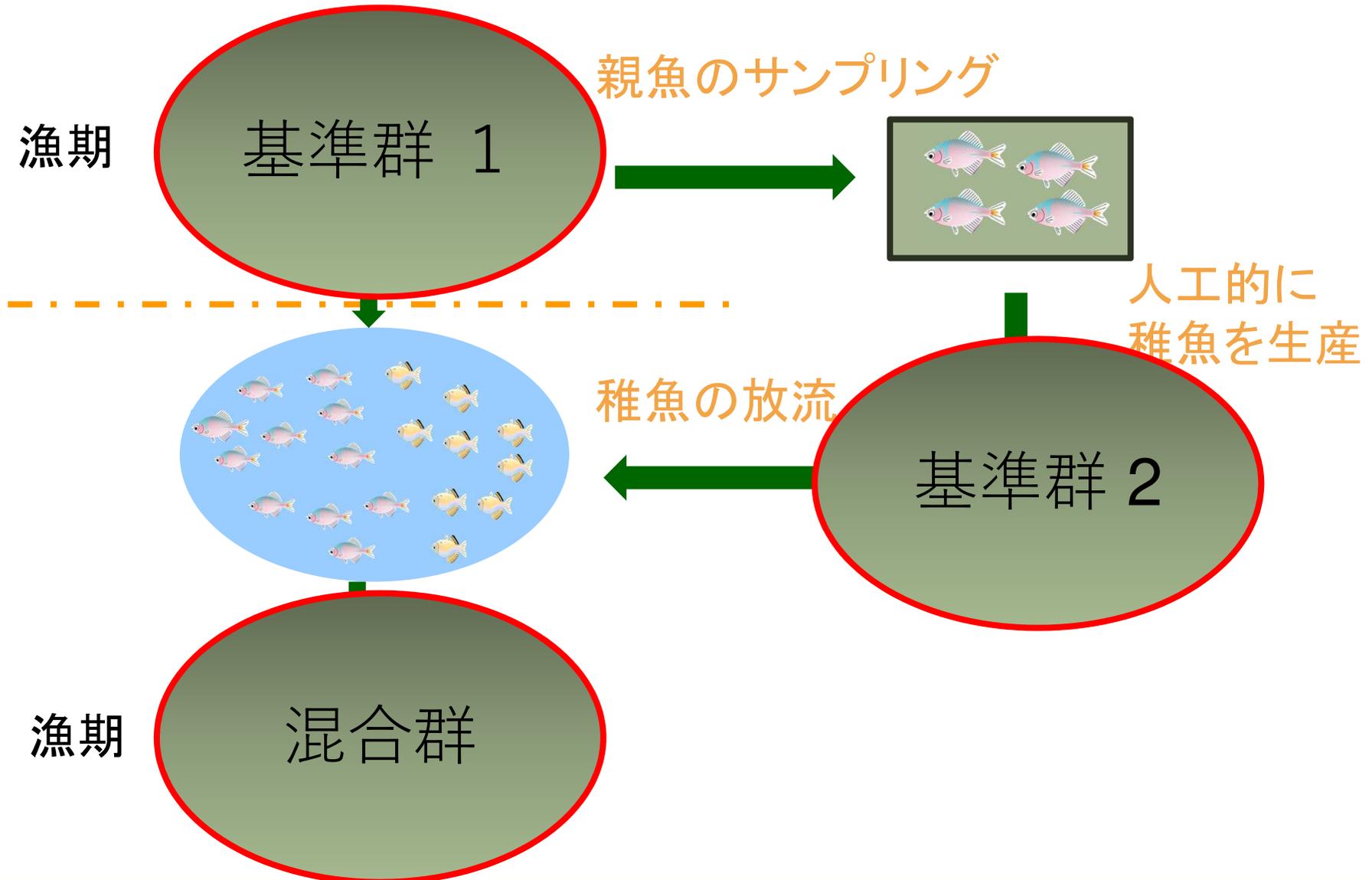
基準群 2 の対立遺伝子頻度



混合率 $\omega : 1 - \omega$



水産増殖における放流集団の混合

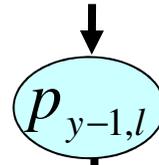


継時サンプリングによる混合率追跡

Y-1年

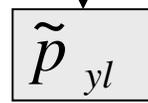
天然群

漁期サンプリング ←



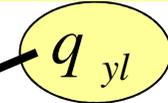
Y年

$$\tilde{p}_{yl} \sim \text{Dirich}(\theta p_{y-1,l})$$

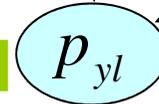


$q_{yl} : \text{known}$

放流群



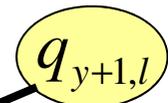
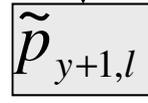
漁期サンプリング ←



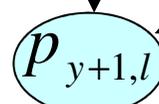
$$p_{1y} = (1 - \omega_y) \tilde{p}_{1,y-1} + \omega_y p_{2y}$$

Y+1年

$$\tilde{p}_{y+1,l} \sim \text{Dirich}(\theta p_{yl})$$



漁期サンプリング ←



Kitakado et al. (2006, 2008)

ノコギリガザミの混合率(mtDNA ハプロタイプ)

天然群のmtDNAハプロタイプ頻度

Year	Haplotype									
	1	2	3	4	5	6	7	8	9	Others
1996	79	23	13	16	10	3	1	3	0	1
1997	90	65	16	13	11	2	5	2	0	9
1998	57	37	22	11	7	0	1	2	0	6
1999	159	124	39	15	11	3	5	5	5	37
2000	225	147	47	32	21	2	4	5	4	34
2001	71	32	8	7	5	0	0	1	11	23
2002	105	50	30	17	0	5	0	0	5	35

放流群のmtDNAハプロタイプ頻度

Year	Haplotype									
	1	2	3	4	5	6	7	6	7	Others
1996	0	0	0	0	0	0	0	0	0	0
1997	0	98,300	0	0	0	0	0	0	0	0
1998	0	0	64,000	0	0	0	11,000	0	0	0
1999	0	0	0	81,000	0	0	0	0	0	0
2000	0	0	0	0	72,000	0	0	0	0	0
2001	0	0	0	0	0	0	0	0	149,000	0
2002	0	0	0	0	0	0	0	0	0	0

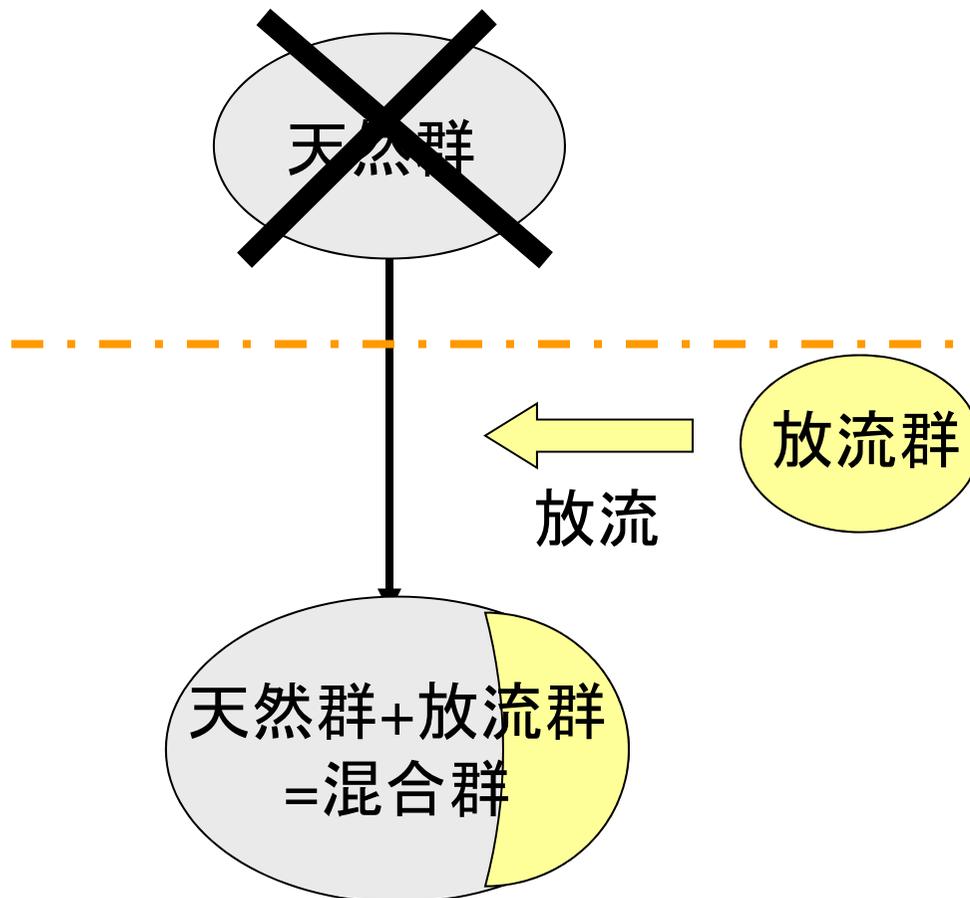
混合率の推定値

Year	Mixing Proportion	
1997	0.152	(0.023)
1998	0.025	(0.021)
1999	0.000	(0.000)
2000	0.015	(0.007)
2001	0.039	(0.009)



「増殖における放流」と「外来種or外来集団の放流」

混合する前の情報が
得られない場合もある



- 増殖における「放流群」は同一種、同一集団の天然群を親として生産されるため、過剰な放流をしない限り、遺伝的影響は小さい

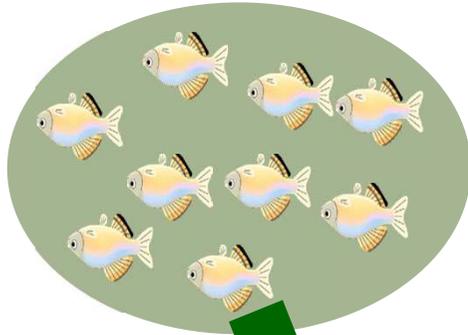
- 一方で、外来種や他の集団の放流は、生息環境の占領や、遺伝的汚染が懸念材料（しばしば望まれない放流）

2. 外来集団の混合率の推定

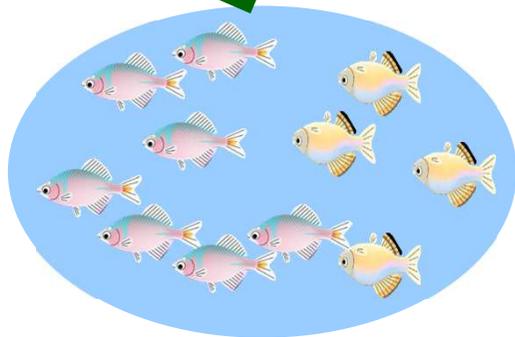
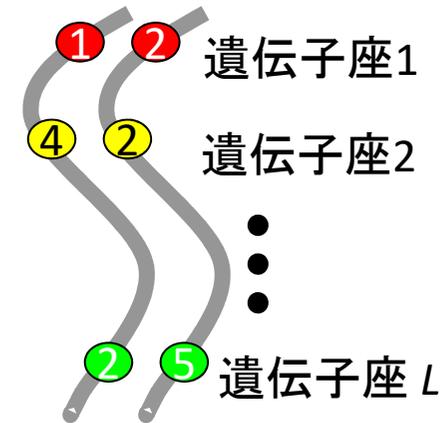
Mixture model

外来集団の混合率推定の設定

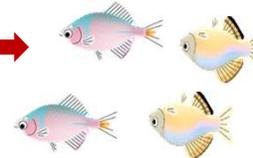
外来集団



N_1 個体の遺伝子型観測 \Rightarrow 尤度1



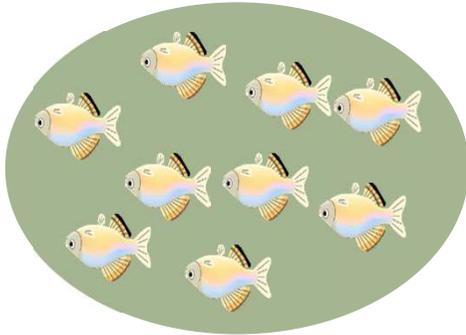
N_2 個体の遺伝子型観測 \Rightarrow 尤度2



混合が疑われている天然漁場集団
(放流される前の情報なし)

外来集団からの観測情報

外来集団



N_1 個体の遺伝子型観測 \Rightarrow 尤度 1



第 l 遺伝子座の“対立遺伝子”頻度ベクトル

$$p_{1l} = (p_{1l1}, \dots, p_{1lJ_l}), l = 1, 2, \dots, L$$

第 l 遺伝子座の“遺伝子型” (j, j') の頻度

$$G_{1l}(j, j') \quad \text{HW平衡からの逸脱度}$$

$$= \begin{cases} (1 - F_{IS,1})p_{1lj}^2 + F_{IS,1}p_{1lj} & \text{if } j = j', \\ 2(1 - F_{IS,1})p_{1lj}p_{1lj'} & \text{if } j \neq j' \end{cases}$$

第 i 番目の個体の第 l 遺伝子座の
“遺伝子型”観測値の確率分布

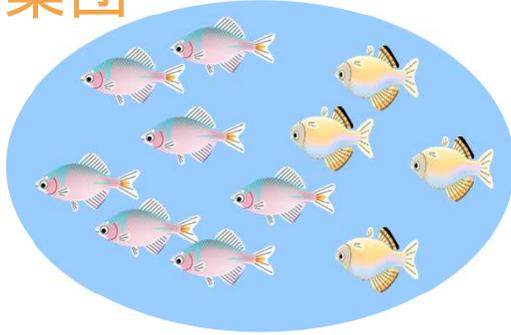
$$Pr((Y_{1il}^{(1)}, Y_{1il}^{(2)}) = (y_{1il}^{(1)}, y_{1il}^{(2)})) \\ = G_{1l}(y_{1il}^{(1)}, y_{1il}^{(2)})$$

全標本の“遺伝子型”観測値の
確率分布(連鎖平衡を仮定)

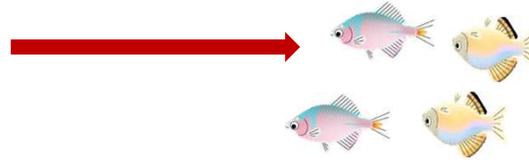
$$L_1 = \prod_{i=1}^{N_1} \prod_{l=1}^L G_{1l}(y_{1il}^{(1)}, y_{1il}^{(2)})$$

混合集団からの観測情報

混合集団



N_2 個体の遺伝子型観測 \Rightarrow 尤度2



集団1 (外来集団)における
第 l 遺伝子座の“遺伝子型” 頻度

$$G_{1l}(j, j') = \begin{cases} (1 - F_{IS,1})p_{1lj}^2 + F_{IS,1}p_{1lj} & \text{if } j = j', \\ 2(1 - F_{IS,1})p_{1lj}p_{1lj'} & \text{if } j \neq j' \end{cases}$$

集団2 (在来集団)における
第 l 遺伝子座の“遺伝子型” 頻度

$$G_{2l}(j, j') = \begin{cases} (1 - F_{IS,2})p_{2lj}^2 + F_{IS,2}p_{2lj} & \text{if } j = j', \\ 2(1 - F_{IS,2})p_{2lj}p_{2lj'} & \text{if } j \neq j' \end{cases}$$

第 i 番目の個体の第 l 遺伝子座の
“遺伝子型”観測値の確率分布

$$\omega \prod_{l=1}^L G_{1l}(y_{2il}^{(1)}, y_{2il}^{(2)}) + (1 - \omega) \prod_{l=1}^L G_{2l}(y_{2il}^{(1)}, y_{2il}^{(2)})$$

全標本の“遺伝子型”観測値の
確率分布

$$L_2 = \prod_{i=1}^{N_2} \left[\omega \prod_{l=1}^L G_{1l}(y_{2il}^{(1)}, y_{2il}^{(2)}) + (1 - \omega) \prod_{l=1}^L G_{2l}(y_{2il}^{(1)}, y_{2il}^{(2)}) \right]$$

パラメータ推定

外来集団 N_1 個体の遺伝子型観測

$$L_1 = \prod_{i=1}^{N_1} \prod_{l=1}^L G_{1l}(y_{1il}^{(1)}, y_{1il}^{(2)})$$

混合集団(外来+在来) N_2 個体の遺伝子型観測

$$L_2 = \prod_{i=1}^{N_2} \left[\omega \prod_{l=1}^L G_{1l}(y_{2il}^{(1)}, y_{2il}^{(2)}) + (1 - \omega) \prod_{l=1}^L G_{2l}(y_{2il}^{(1)}, y_{2il}^{(2)}) \right]$$

未知パラメータ：

- 外来集団個体の混合率 ω
- 各集団におけるHW平衡からの逸脱度 $F_{IS,1}$, $F_{IS,2}$
- 各集団の各遺伝子座における対立遺伝子頻度

$$p_{1l} = (p_{1l1}, \dots, p_{1lJ_l}) \quad p_{2l} = (p_{2l1}, \dots, p_{2lJ_l})$$

シミュレーションテスト (Mixture)

外来集団と在来集団の対立遺伝子頻度 (bi-allelic)

$$(p_{1l1}, p_{1l2}) = (0.5(1 + \sqrt{F_{ST}}), 1 - 0.5(1 + \sqrt{F_{ST}}))$$

$$(p_{2l1}, p_{2l2}) = (1 - 0.5(1 + \sqrt{F_{ST}}), 0.5(1 + \sqrt{F_{ST}}))$$

遺伝的差異 F_{ST}

遺伝子座数 = 10

サンプル数: 各集団100個体

「混合率」, 「HW平衡からの逸脱度」, 「 F_{ST} 」の幾つかの組み合わせで推定値の平均, およびMSEを評価

1) HW不平衡(HWD)を仮定したモデル

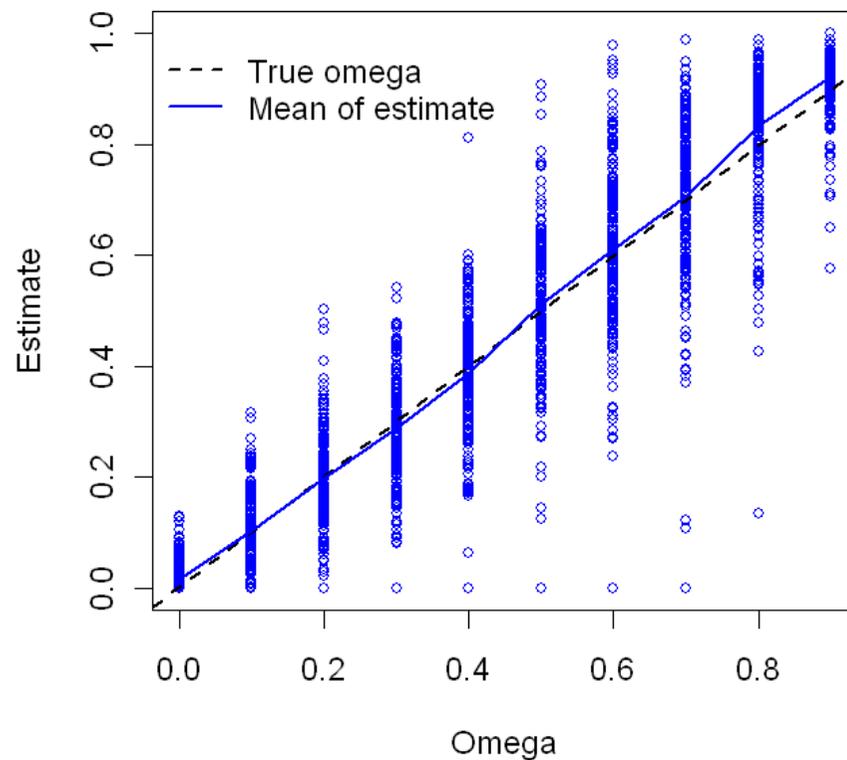
2) HW平衡(HWE)を仮定したモデル(HWからのずれを無視)

HWEは混合を問うモデルでしばしば仮定される

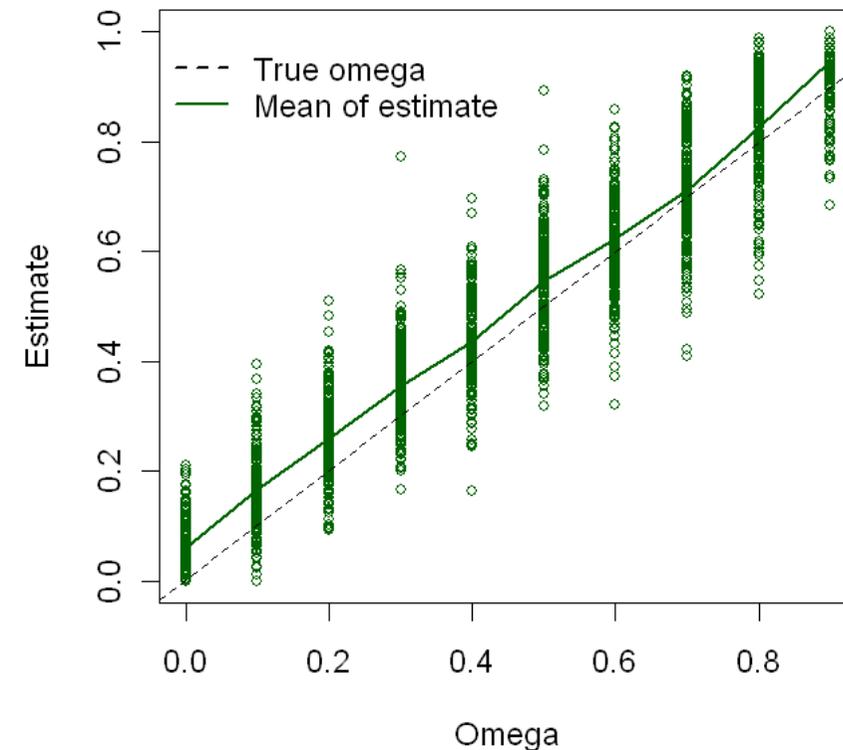
混合率を変化させたときの性能評価(1)

$F_{IS}=0.3, F_{ST}=0.05$ で固定

Based on HWD model



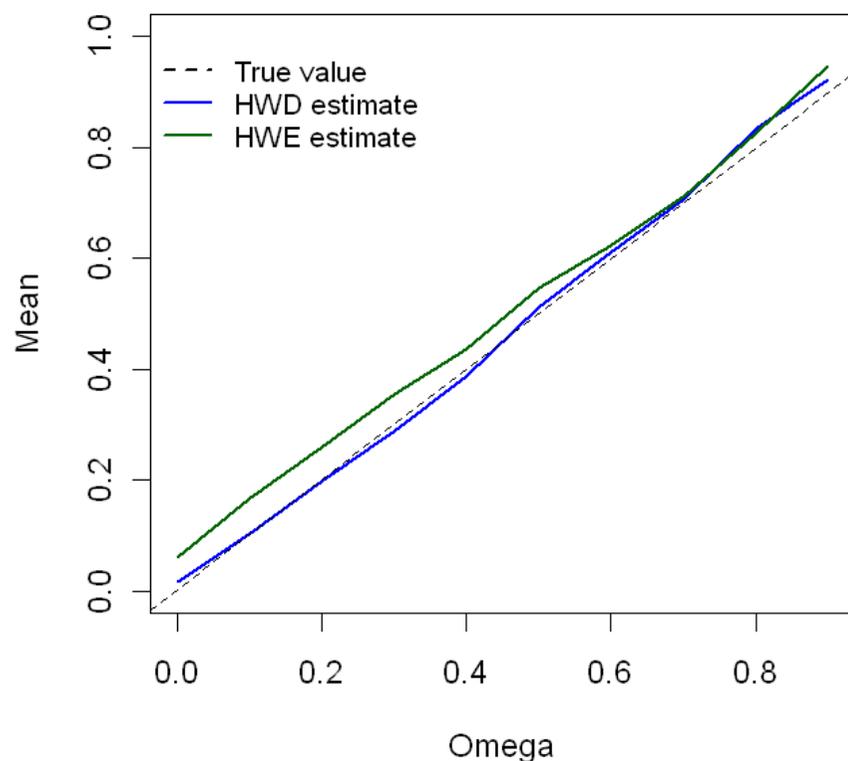
Based on HWE model



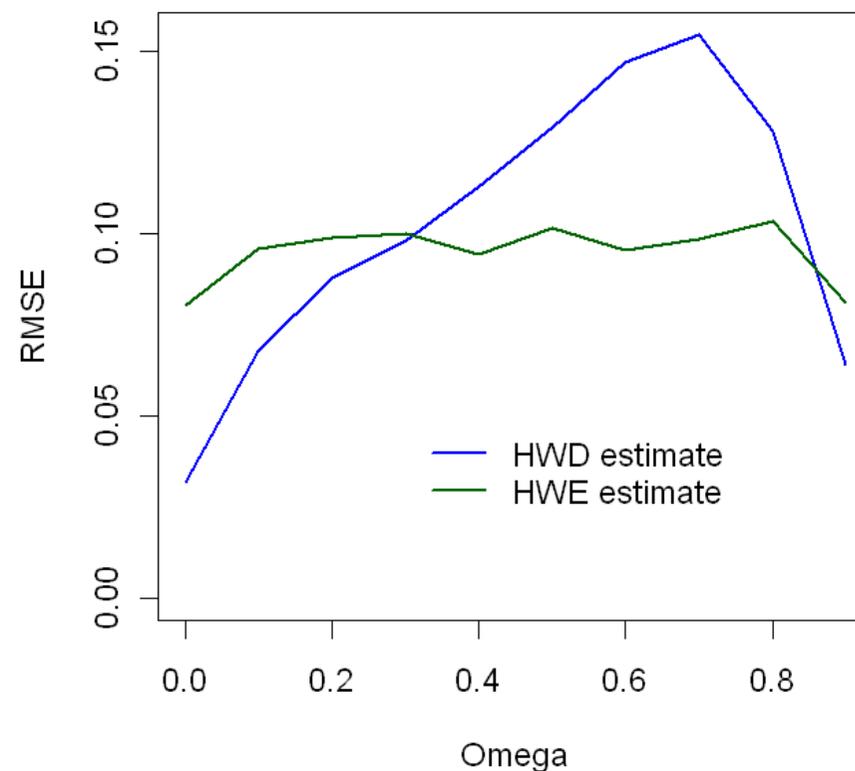
混合率を変化させたときの性能評価(2)

$F_{IS}=0.3, F_{ST}=0.05$ で固定

Mean of estimate



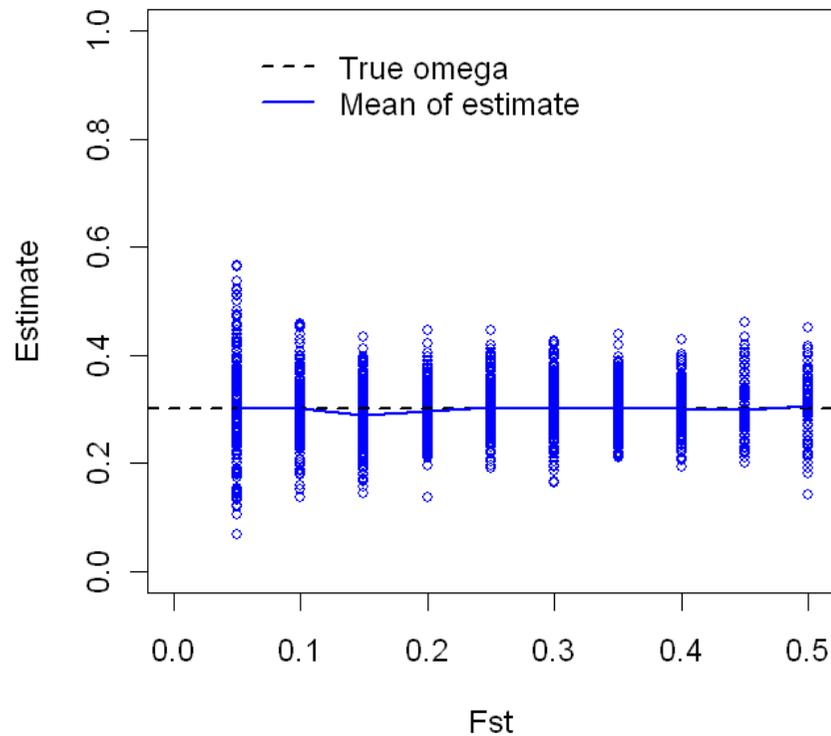
RMSE of estimate



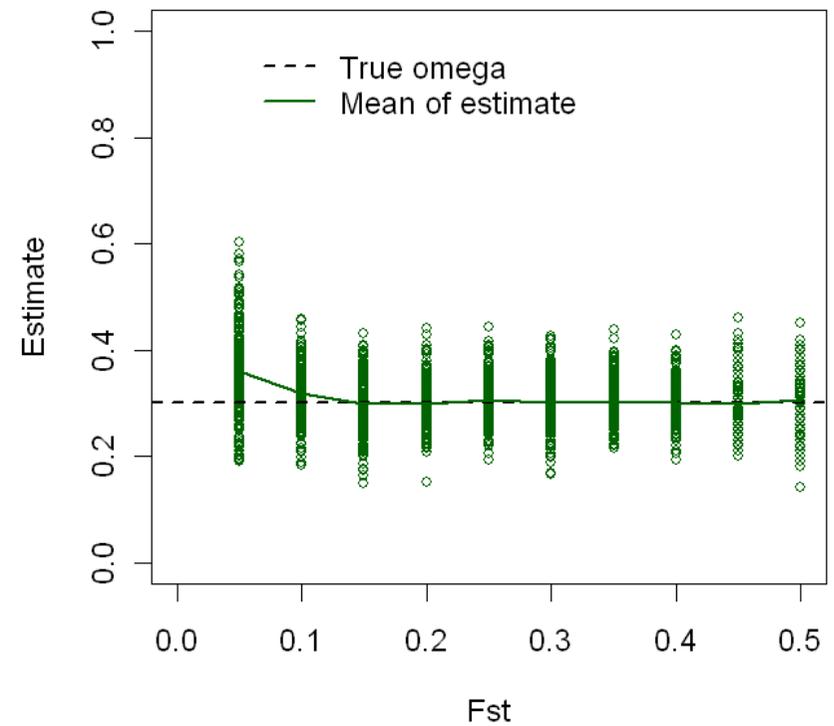
F_{ST} を変化させたときの性能評価

$\omega=0.3, F_{IS}=0.3$ で固定

Based on HWD model



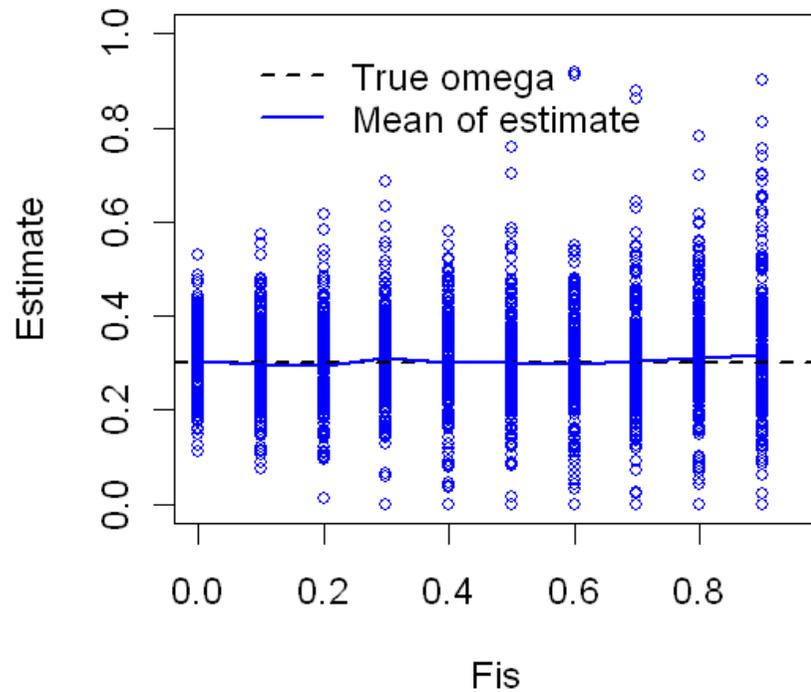
Based on HWE model



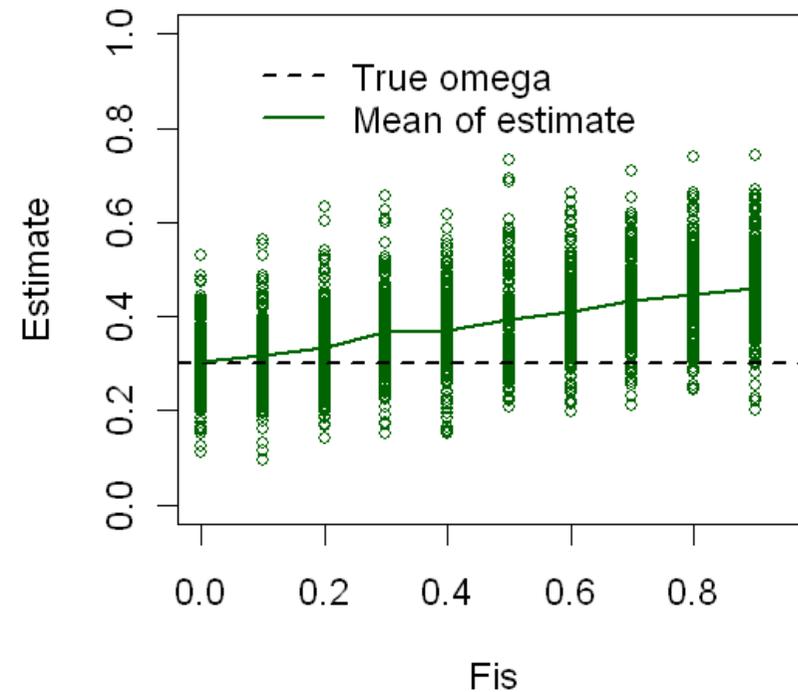
F_{IS} を変化させたときの性能評価

$\omega=0.3, F_{ST}=0.05$ で固定

Based on HWD model



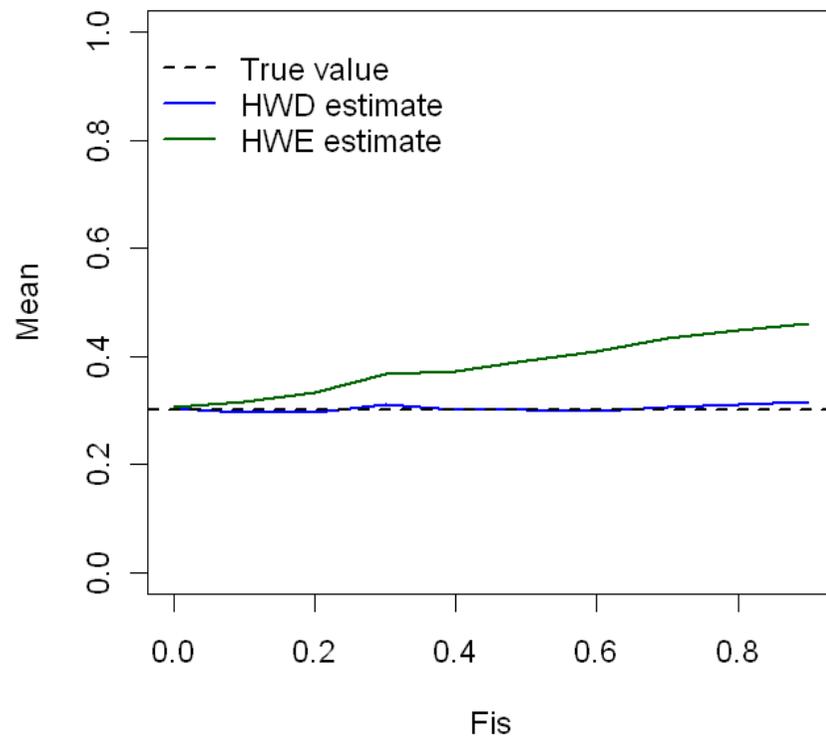
Based on HWE model



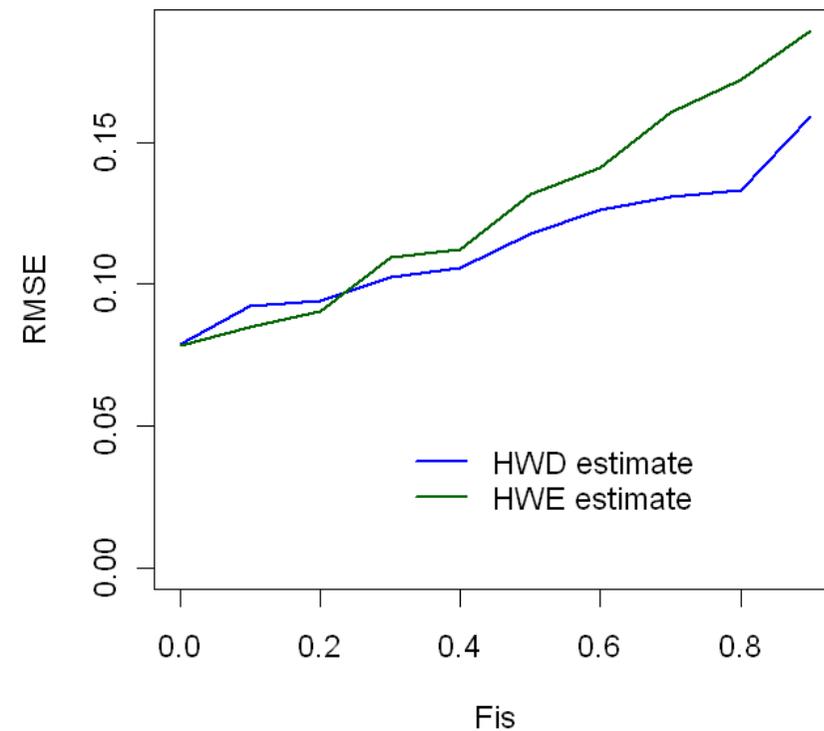
F_{IS} を変化させたときの性能評価

$\omega=0.3, F_{ST}=0.05$ で固定

Mean of estimate



RMSE of estimate

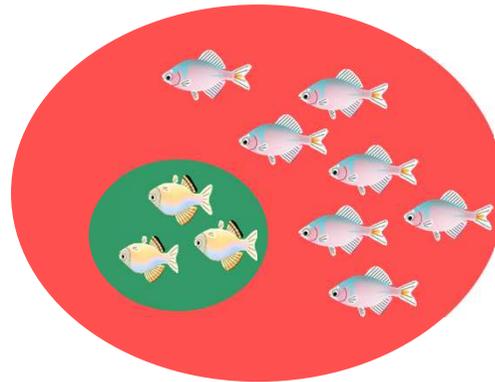


3. 外来集団による遺伝的汚染

Admixture model

Admixture Model

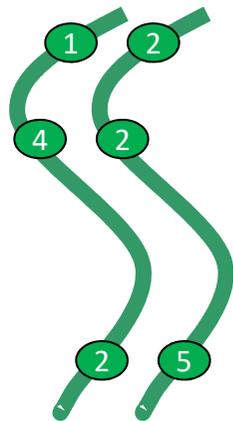
混合率 ω



- 再生産力の違い
- 外来集団の環境への適応
- 遭遇率の違い
- 外来・在来の交雑の有無

$\omega^2 ?$

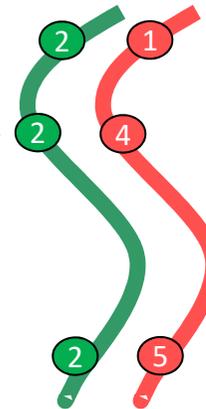
外来 × 外来



再生産

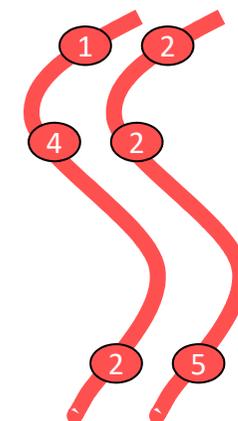
$2\omega(1-\omega) ?$

外来 × 在来



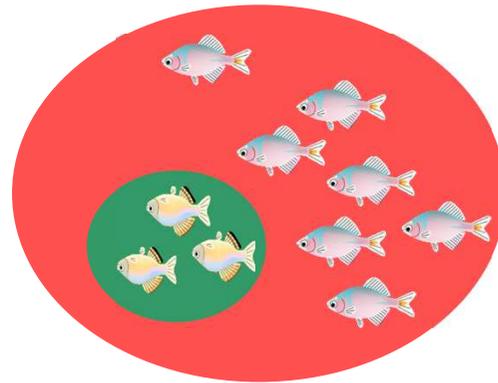
$(1-\omega)^2 ?$

在来 × 在来



Admixture Model

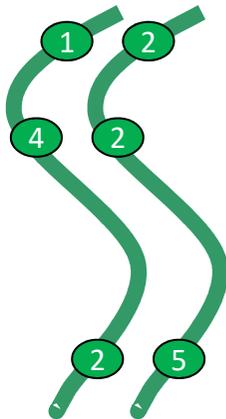
混合率 ω



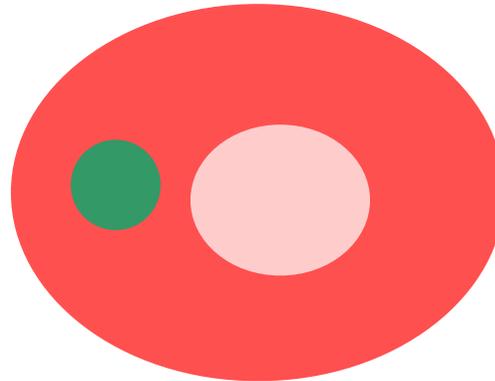
遺伝的汚染の有無を
表すパラメータ

δ_{11}

外来 × 外来

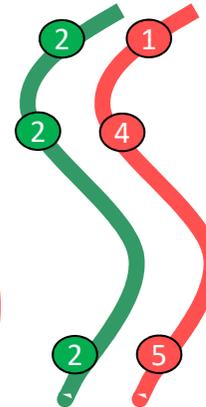


再生産



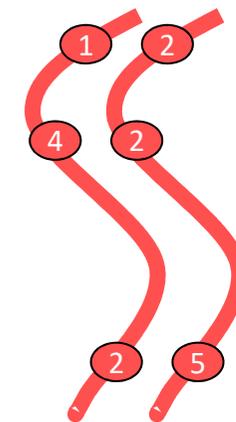
δ_{12}

外来 × 在来



δ_{22}

在来 × 在来



Admixture Model (Likelihood)

外来 × 外来

$$\tilde{G}_{11,l}(j, j') = G_{1l}(j, j') = \begin{cases} (1 - F_{IS,1})p_{1lj}^2 + F_{IS,1}p_{1lj} & \text{if } j = j', \\ 2(1 - F_{IS,1})p_{1lj}p_{1lj'} & \text{if } j \neq j' \end{cases}$$

外来 × 在来

$$\tilde{G}_{12,l}(j, j') = \begin{cases} p_{1lj}p_{2lj'} & \text{if } j = j', \\ p_{1lj}p_{2lj'} + p_{1lj'}p_{2lj} & \text{if } j \neq j' \end{cases}$$

在来 × 在来

$$\tilde{G}_{22,l}(j, j') = G_{2l}(j, j') = \begin{cases} (1 - F_{IS,2})p_{2lj}^2 + F_{IS,2}p_{2lj} & \text{if } j = j', \\ 2(1 - F_{IS,2})p_{2lj}p_{2lj'} & \text{if } j \neq j' \end{cases}$$

$$\tilde{L}_2 = \prod_{i=1}^{N_2} \left[\delta_{11} \prod_{l=1}^L \tilde{G}_{11,l}(y_{2il}^{(1)}, y_{2il}^{(2)}) + \delta_{12} \prod_{l=1}^L \tilde{G}_{12,l}(y_{2il}^{(1)}, y_{2il}^{(2)}) + \delta_{22} \prod_{l=1}^L \tilde{G}_{22,l}(y_{2il}^{(1)}, y_{2il}^{(2)}) \right]$$

シミュレーション結果 (admixture)

F_{IS}	Based on HWD model					
	δ_{11}		δ_{12}		δ_{22}	
0.0	0.115	(0.068)	0.261	(0.210)	0.625	(0.180)
0.1	0.112	(0.067)	0.259	(0.148)	0.629	(0.134)
0.2	0.105	(0.055)	0.289	(0.113)	0.605	(0.111)
0.3	0.102	(0.060)	0.298	(0.081)	0.600	(0.087)
0.4	0.099	(0.059)	0.280	(0.064)	0.621	(0.080)
0.5	0.103	(0.059)	0.296	(0.051)	0.601	(0.067)

F_{IS}	Based on HWE model					
	δ_{11}		δ_{12}		δ_{22}	
0.0	0.118	(0.066)	0.281	(0.215)	0.607	(0.193)
0.1	0.167	(0.069)	0.070	(0.104)	0.763	(0.098)
0.2	0.200	(0.068)	0.017	(0.047)	0.782	(0.073)
0.3	0.225	(0.069)	0.001	(0.012)	0.774	(0.070)
0.4	0.248	(0.076)	0.000	(0.000)	0.752	(0.076)
0.5	0.273	(0.081)	0.000	(0.000)	0.727	(0.081)

固定したパラメータ真値

$$\delta_{11} = 0.1$$

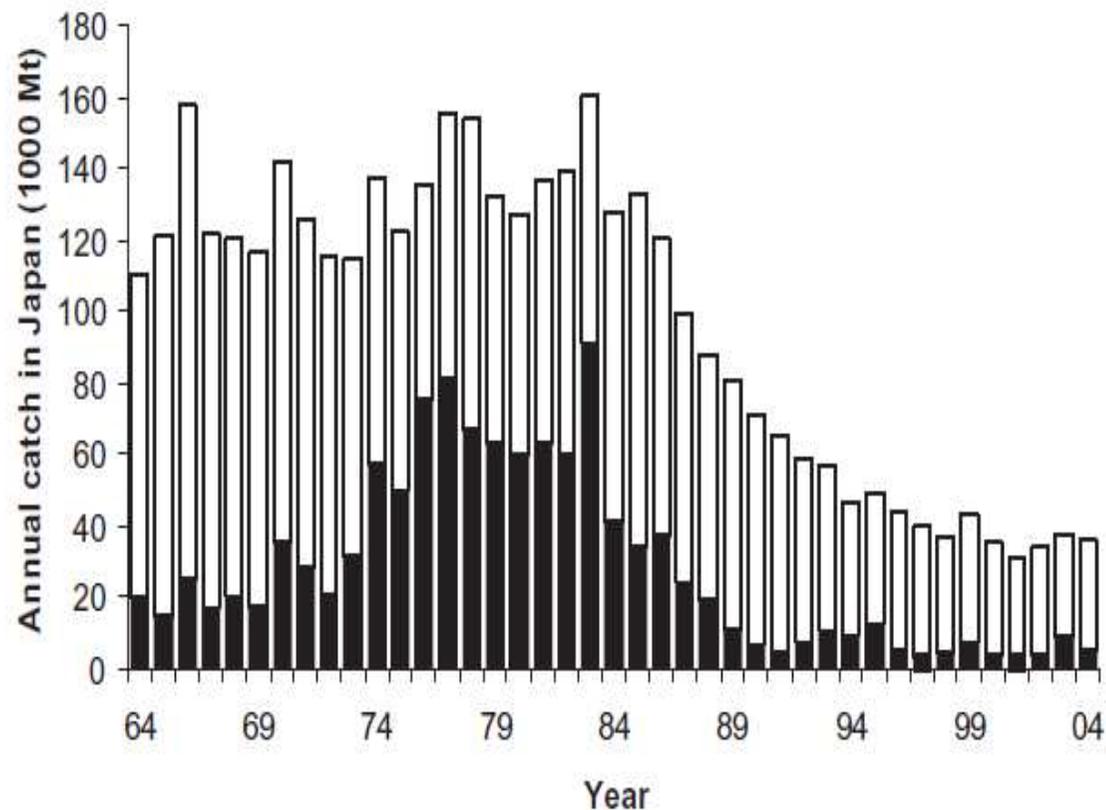
$$\delta_{12} = 0.3$$

$$\delta_{22} = 0.6$$

$$F_{ST} = 0.05$$

4. 有明海アサリへの適用

国内および有明アサリの漁獲量変動



有明海

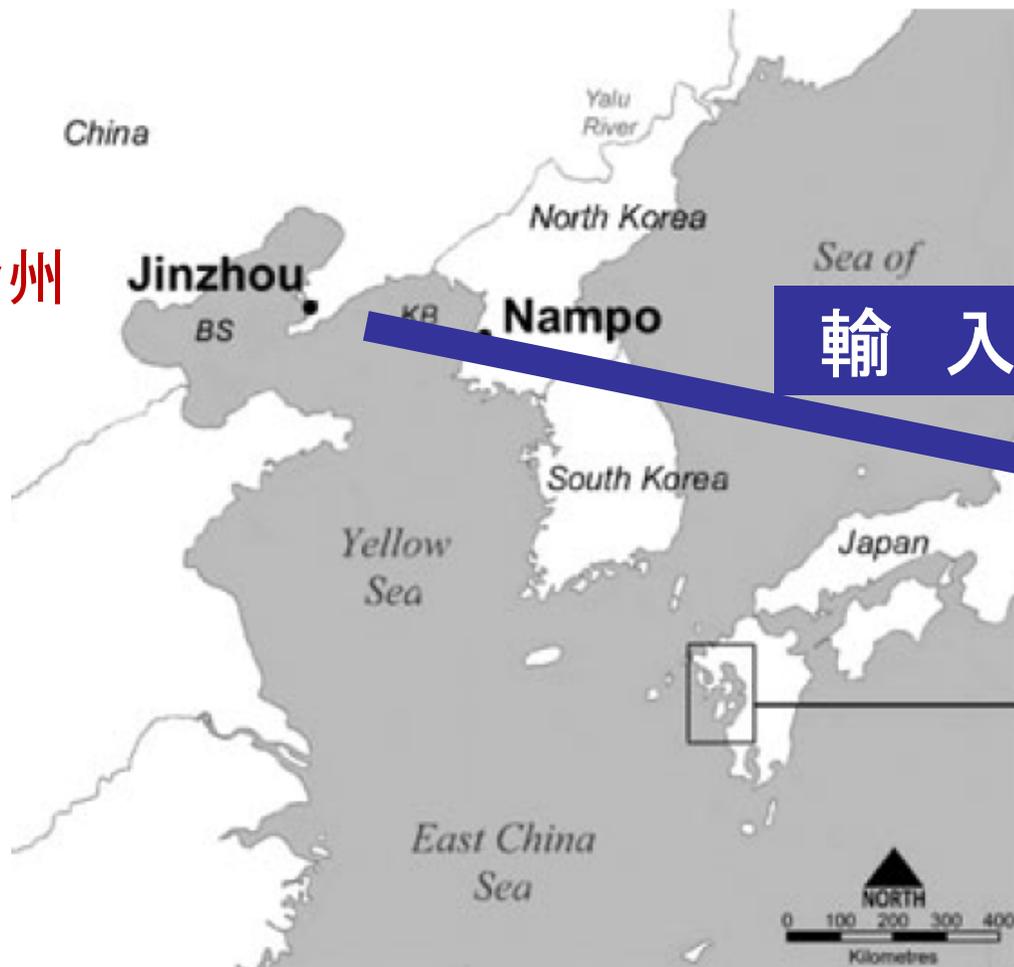


Fig. 1 Annual catch of the littleneck clam from 1964 to 2004 in all Japan (□) and the percentage of catches in the Ariake Sea and Shiranui Sea combined (■).

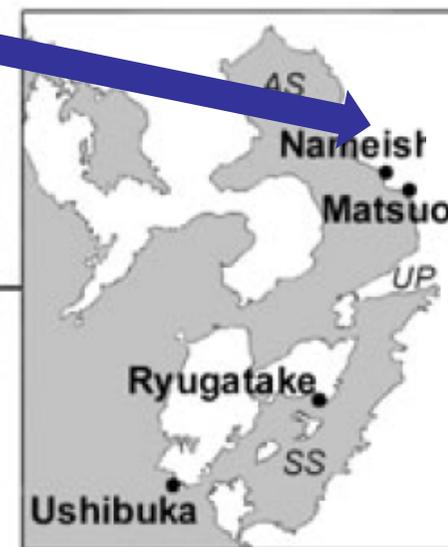
Vargas et al. (2008)

有明海アサリと外国からの輸入アサリ

金州



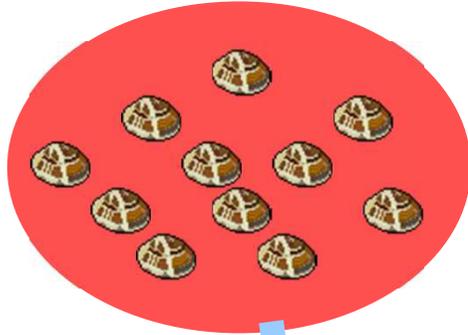
輸入



滑石
(蓄養場)
松尾

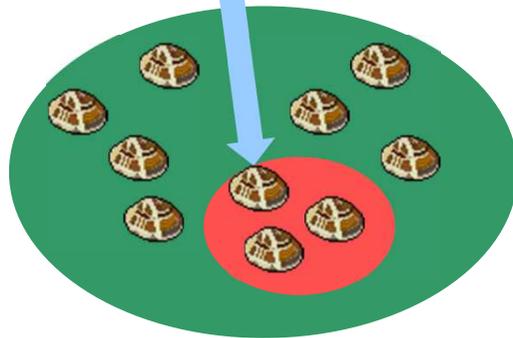
個体のサンプリング

外来集団 (中国渤海湾金州)



金州: 91個体をサンプリング

有明集団 + 中国金州集団?



滑石の“天然漁場”: 103個体をサンプリング
(蓄養場とは別の場所)

松尾の“天然漁場”: 105個体をサンプリング

- それぞれ8遺伝子座の遺伝子型を観測
- Mixture, Admixture それぞれのモデルで推定

Mixture model における結果

Locality	ω	$F_{IS,1}$	$F_{IS,2}$	log-likelihood
Under HWD				
滑石	0.320 (0.072)	0.111 (0.028)	0.169 (0.040)	-1999.72
松尾	0.000 (0.000)	0.115 (0.032)	0.075 (0.028)	-1891.85
滑石+松尾	0.180 (0.078)	0.111 (0.030)	0.116 (0.024)	-2990.72
Under HWE				
滑石	0.416 (0.067)	—	—	-1954.56
松尾	0.000 (0.000)	—	—	-1908.40
滑石+松尾	0.400 (0.066)	—	—	-3026.50

☆蓄養場のある滑石の天然漁場では、中国輸入アサリの高い混合率が示唆された

☆一方で、放流実績なしとされている松尾の混合率はゼロと推定

☆HW平衡を無視すると混合率が高めに推定される

Admixture model における結果

Locality	δ_{11}	δ_{12}	δ_{22}
Under HW disequilibrium			
滑石	0.320 (0.072)	0.000 (0.000)	0.680 (0.032)
松尾	0.000 (0.000)	0.000 (0.000)	1.000 (0.032)
滑石+松尾	0.180 (0.078)	0.000 (0.000)	0.820 (0.032)
Under HW equilibrium			
滑石	0.416 (0.067)	0.000 (0.000)	0.584 (0.067)
松尾	0.000 (0.000)	0.000 (0.000)	1.000 (0.000)
滑石+松尾	0.414 (0.066)	0.000 (0.000)	0.586 (0.064)

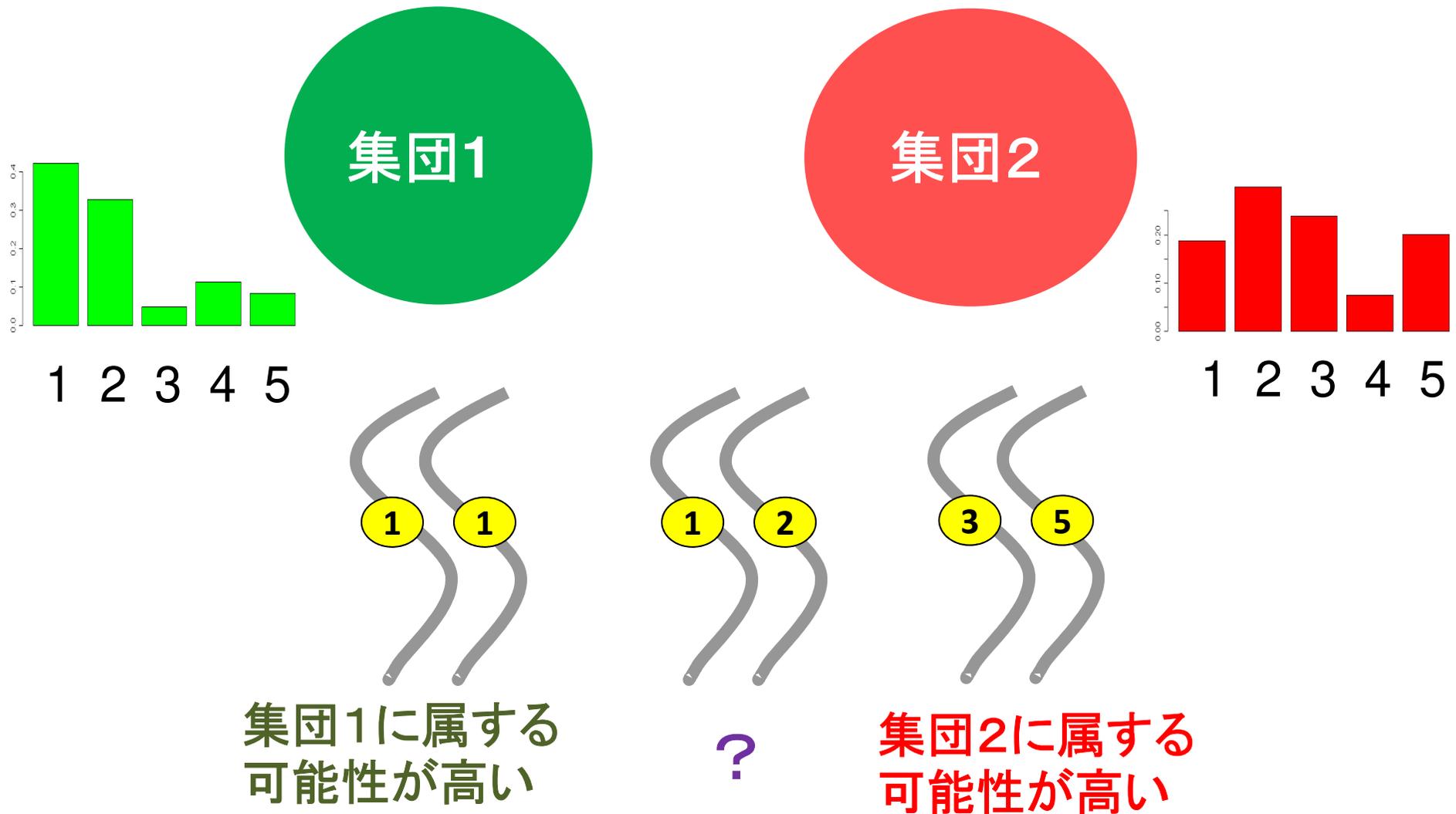
- ☆ 有明アサリと中国輸入アサリの交雑を表すパラメータはいずれも0と推定され、遺伝的汚染は生じていないことが示唆された
- ☆ 有明の水温が、中国輸入アサリに適していないため？

補足：個体のアサンメント

個体のアサインメント

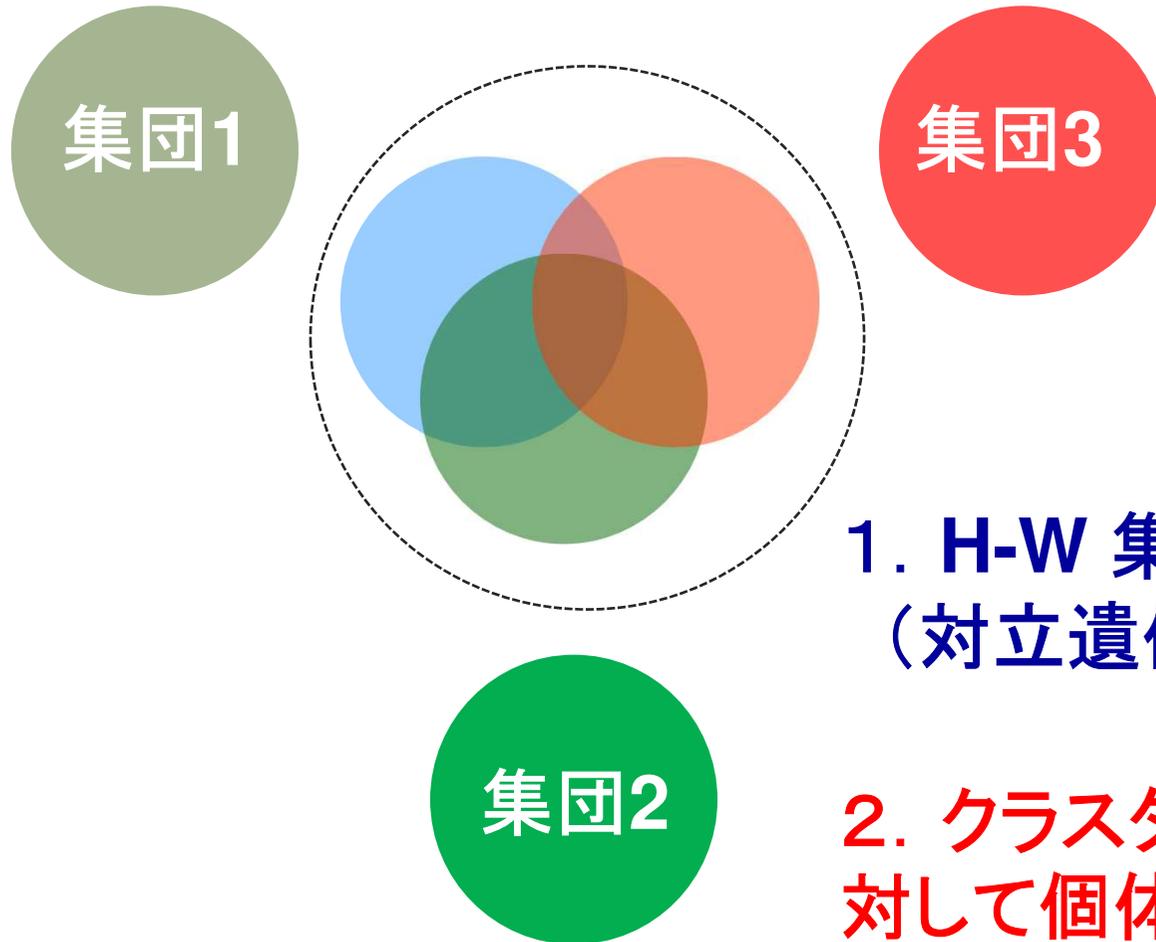
- 個体ごとの遺伝子型情報を利用し，それぞれの個体が，どの集団に属しているかを（確率的に）割り当てる
- 基準となる **Source Population** の情報を必要とする方法と，しない方法がある

アサインメントの考え方



Source Population がないとき

STRUCTURE (Pritchard et al. 2000, Falush et al. 2003)



1. H-W 集団をクラスタリング
(対立遺伝子頻度の推測)

2. クラスタリングされた集団に
対して個体をアサインメント

アサインメント確率のベイズ推定

Z: ある個体の集団への帰属を表す潜在変数

事前確率

$$P(Z=1) = P(Z=2) = P(Z=3) = \frac{1}{3}$$



事後確率

$$\Pr(Z | Data) = \frac{\int \int \Pr(Data | Z, P) \Pr(Z, P) dP dZ}{\int \int \Pr(Data | Z, P) \Pr(Z, P) dP dZ}$$

多重積分！計算困難

マルコフ連鎖モンテカルロ法(MCMC)

Gibbsサンプリング(モンテカルロシミュレーション)

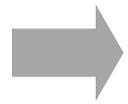
$$1. \Pr(P | Y, Z^{(i)}) \Rightarrow P^{(i+1)}$$

$$2. \Pr(Z | Y, P^{(i+1)}) \Rightarrow Z^{(i+1)}$$

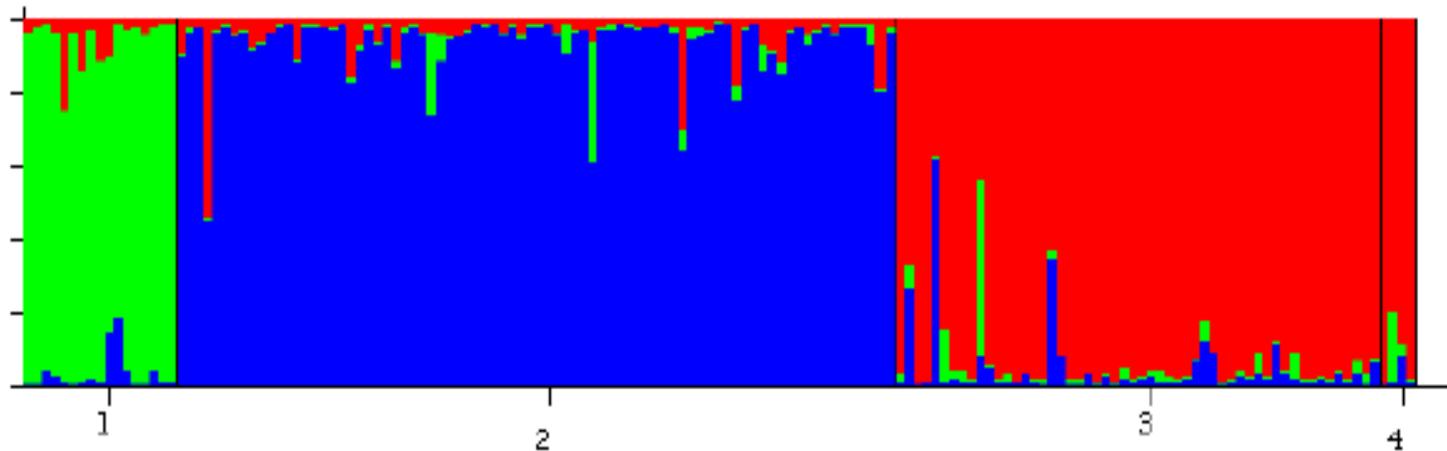
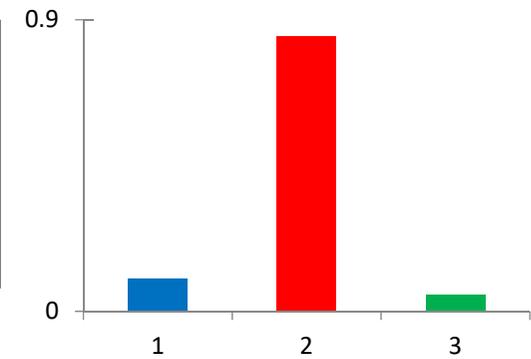
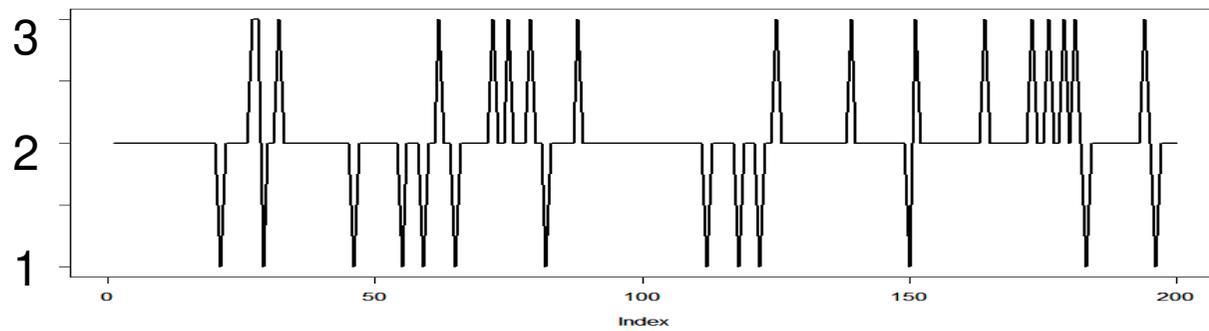
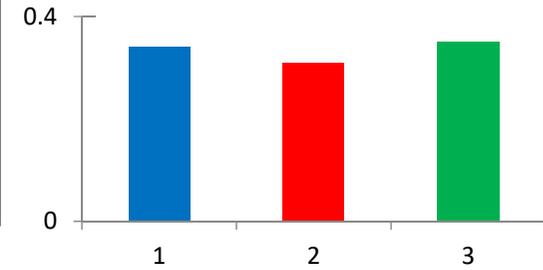
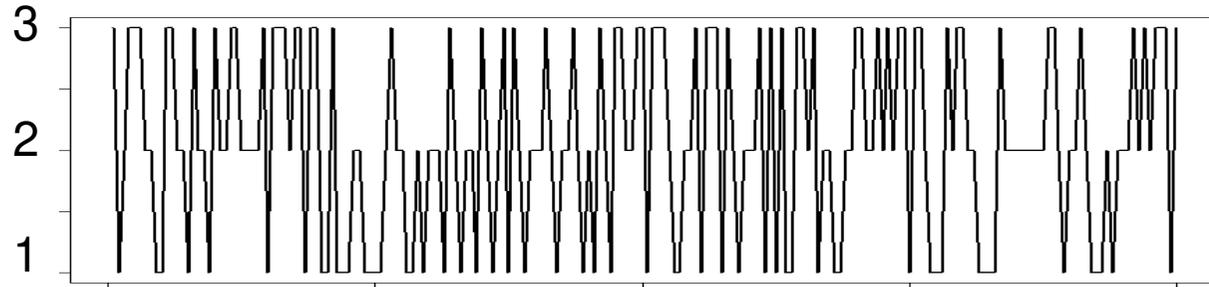
1. H-W集団をクラスタリングにより構成し, それぞれの対立遺伝子頻度を計算
2. 暫定的にクラスタリングされた集団をSource集団として, 個体をそれぞれの集団にアサインメント

$$\cdots \Rightarrow (P^{(i)}, Z^{(i)}) \Rightarrow (P^{(i+1)}, Z^{(i+1)}) \Rightarrow \cdots$$

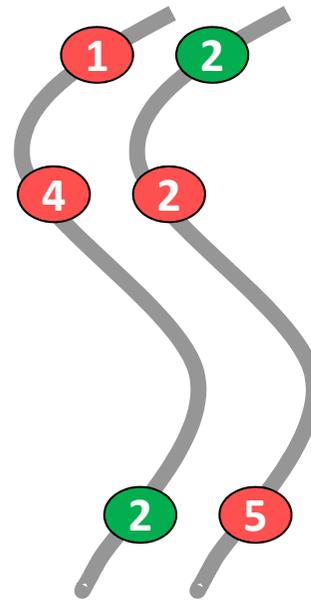
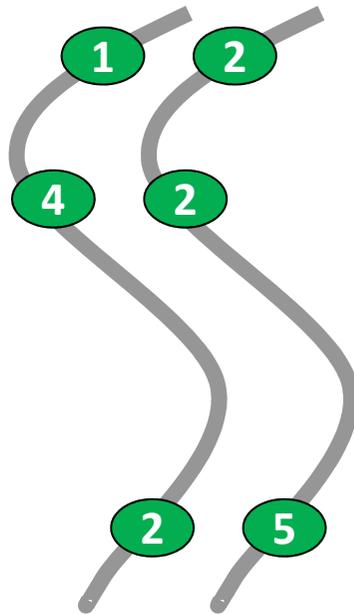
$$P(Z = k) = \frac{1}{3}$$



$$P(Z = k \mid \text{Data}) = ?$$



Mixture と Admixture



Mixture

個体自身の由来を問う。
アサインメントの確率は、
＝個体が各集団に属する確率

Admixture

個体内の遺伝子の由来を問う。
アサインメントの確率
＝個体内の遺伝子の由来の割合