褐色菜鴨高飼效品系

申請登記審查資料

審查單位：行政院農業委員會

申請單位：行政院農業委員會畜產試驗所

中華民國 106 年 10 月

申請登記審查資料

種畜禽或種原登記申請書

<table>
<tr><td>申請種禽種類</td><td colspan="4">褐色菜鴨</td></tr>
<tr><td>申請登記名稱</td><td colspan="4">褐色菜鴨高飼效品系 (Better Feed Efficiency Brown Tsaiya)</td></tr>
<tr><td>申請登記代號</td><td colspan="4">BT-BFE</td></tr>
<tr><td rowspan="4">登記申請人
(代表人)</td><td>姓名</td><td colspan="3">黃振芳</td></tr>
<tr><td>機關名稱</td><td colspan="3">行政院農業委員會畜產試驗所</td></tr>
<tr><td rowspan="2">電子信箱</td><td rowspan="2">admin@mail.tlri.gov.tw</td><td>電話</td><td>(06)5911211</td></tr>
<tr><td>傳真</td><td>(06)5911210</td></tr>
<tr><td>聯絡地址</td><td colspan="3">71246 臺南市新化區牧場 112 號</td></tr>
<tr><td>品系來源</td><td colspan="4">於 2009 年選取行政院農業委員會畜產試驗所褐色菜鴨畜試一號第 16 代公鴨 157 隻與母鴨 195 隻，藉考量能量利用與體組成變化之殘差飼料採食量選拔，改進鴨隻飼料利用效率。第零代與第一代皆根據殘差飼料採食量表型值選留種鴨繁殖，自第二代開始，依殘差飼料採食量性狀的無偏差育種價估測值選留最佳育種價估測值鴨隻，繁殖試驗用鴨群，至第六代選拔之種公鴨與種母鴨分別為 78 隻與 148 隻。</td></tr>
<tr><td>育成機關</td><td colspan="4">行政院農業委員會畜產試驗所宜蘭分所。</td></tr>
<tr><td>品系特性</td><td colspan="4">1. 外觀特徵：母鴨全身為淡至深褐色，喙及腳脛橙黃，隨產蛋週齡增加而顏色漸褪；公鴨頭頸部暗褐色，頸中部或有白色頸圈，背部灰褐色，前胸呈葡萄栗色，腹部為灰色或灰褐色，尾部有性捲羽，喙黃綠色、黃色或灰黑色不一，腳橙黃。
2. 產蛋性能：第六代初產日齡為 116 ± 10 天；40 週齡產蛋數為 145 ± 18 枚、52 週齡產蛋數為 203 ± 28 枚。
3. 依 34~37 週殘差飼料採食量育種價估測值估算，選拔品系每日殘差飼料採食量可減少 8.0 g，以產蛋期 10 個月估算，每隻產蛋母鴨約可節省 2.4 kg 飼料成本支出。</td></tr>
<tr><td>飼養管理及防疫計畫</td><td colspan="4">1. 飼養分期：鴨隻生長階段分成育雛期(0~4 週齡)、育成期(4~16 週齡)及產蛋期(16 週齡以後)。
2. 飼料營養及管理：育雛期(0~4 週齡)及育成前期(4~8 週齡)均餵飼含代謝能 2,909 kcal/kg 及粗蛋白質 19.5%之粉狀料，育成後期(8 週齡~初產)餵飼含代謝能 2,660 kcal/kg 及粗蛋白質 13.5% 之粒狀料，產蛋期餵飼含代謝能 2,712 kcal/kg 及粗蛋白質 20 %之粒狀蛋鴨料，各飼養階段皆任飼，水自由飲用。
3. 防疫計畫：鴨隻 4 週齡與 8 週齡進行家禽霍亂疫苗注射各乙次。</td></tr>
<tr><td>種禽主要用途</td><td colspan="4">1. 可作為高飼效純系育種。
2. 可作為與民間褐色菜鴨雜交生產商用蛋鴨之種原。
3. 可直接作為高飼效之鴨蛋供應品系。</td></tr>
</table>

申請機關：行政院農業委員會畜產試驗所

申請人：黃振芳　　申請日期：中華民國 106 年 10 月 2 日

目　　錄

表　　次

圖　　次

一、擬申請種畜禽或種原新品系登記名稱

褐色菜鴨高飼效品系

二、新品系代號

BT-BFE

三、育種目標

褐色菜鴨為臺灣優良之蛋鴨品種，體型小，產蛋多，蛋重大，且蛋殼堅固，為我國加工蛋 (皮蛋、鹹蛋) 之主要來源。以畜產試驗所宜蘭分所育成之褐色菜鴨畜試一號為種原，依個體殘差飼料採食量育種價進行選留，改進鴨隻飼料利用效率，以因應飼料成本日益高漲，增加產業之競爭力。

四、育成經過

(一)種原來源

於 2009 年選取行政院農業委員會畜產試驗所褐色菜鴨畜試一號第 16 代公鴨 157 隻與母鴨 195 隻作為試驗用鴨群，藉殘差飼料採食量性狀選拔，改進鴨隻飼料利用效率。第零代與第一代皆根據殘差飼料採食量表型值選留種母鴨及其全同胞兄弟作為種公鴨，避開親屬關係進行繁殖；自第二代開始，依殘差飼料採食量性狀的無偏差育種價估測值選留最佳育種價估測值公、母鴨 (公鴨之育種價依其系譜關係估算)，繁殖試驗用鴨群，每一代並逢機繁殖對照品系。

(二)選育流程與配種設計

本分所自 2009 年藉由遺傳育種理論與混合模式之應用，設計以系譜選育的方式進行殘差飼料採食量性狀選拔，藉以改進鴨隻飼料利用效率。殘差飼料採食量性狀選拔係運用體重、蛋重、體重變化等參數於線性迴歸估測個體採食量，實際採食量減去預估採食量即為殘差飼料採食量 (Koch *et al*., 1963; Bordas and Mérat, 1981)，低殘差飼料採食量個體即具有較佳飼料利用

效率。由於殘差飼料採食量加入能量利用與體組成變化影響之考量，且在諸多研究皆顯示其屬中度至高度遺傳變異率，故較僅考量產蛋量及採食量的飼料換蛋率具有更佳之飼料利用選拔效率。主要試驗及選育流程如下：

1. 鴨隻 0~4 週齡配合保溫燈採高床育雛，4 週齡後移入育成舍 (採高床育成) ，自 12 週齡移入鋼構鴨舍 (有水簾、風扇)，採個別籠飼方式飼養，每隻活動面積為 990 cm^2(長 30 cm * 寬 33 cm * 高 45 cm)，母鴨約於 16 週進入產蛋期。

2. 34~37 週齡為殘差飼料採食量檢定期間，每 3~4 天供給 700 g 蛋鴨料，並以面寬 15 公分之特製壓克力飼料槽餵給，避免隔壁鴨隻盜食 (圖 1a)。以乳頭式飲水器供應飲水，兩隻共用一個飲水乳頭。每 3 或 4 天定時測定鴨隻飼料消耗量，每天收集產蛋並秤其蛋重，檢定開始及結束各秤取鴨重 1 次。根據個體在產蛋期 (34~37 週齡) 之飼料採食量、總蛋重、體重變化及檢定結束體重計算個體之殘差飼料採食量；褐色菜鴨高飼效品系與對照品系之殘差飼料採食量皆依下列方程式計算：

$$RFC = FC - pFC = FC - [a(BW)^{0.5} + b(\Delta BW) + c(EM) + d]$$

其中 FC 為 34~37 週齡飼料採食量、pFC 為預測飼料採食量、BW 為檢定結束體重、ΔBW為體重變化、EM 為 34~37 週齡蛋產量、RFC 為殘差飼料採食量，a~d 為將 BW、ΔBW及 EM 三項性狀進行複回歸後獲得之係數。

3. 收集褐色菜鴨高飼效品系每代資料後，第零、一代根據殘差飼料採食量表型值選留較佳種鴨繁殖；第二代開始，則使用 SAS® 統計軟體進行資料處理及統計分析，利用系譜之親屬關係資料經重編碼後，進行最佳線性無偏差估測值 (BLUP) 之統計分析，依個體殘差飼料採食量育種價估測值，選留最佳估測值種鴨進行繁殖；對照品系則採逢機繁殖。兩品系各依前述條件挑選公鴨 10~12 隻及母鴨

40~48 隻，於 44~46 週齡進行純系配種，50~52 週齡孵化產生下一代，供繼續試驗。截至目前，選育過程共進行 6 個世代(圖 1b)。

(a)

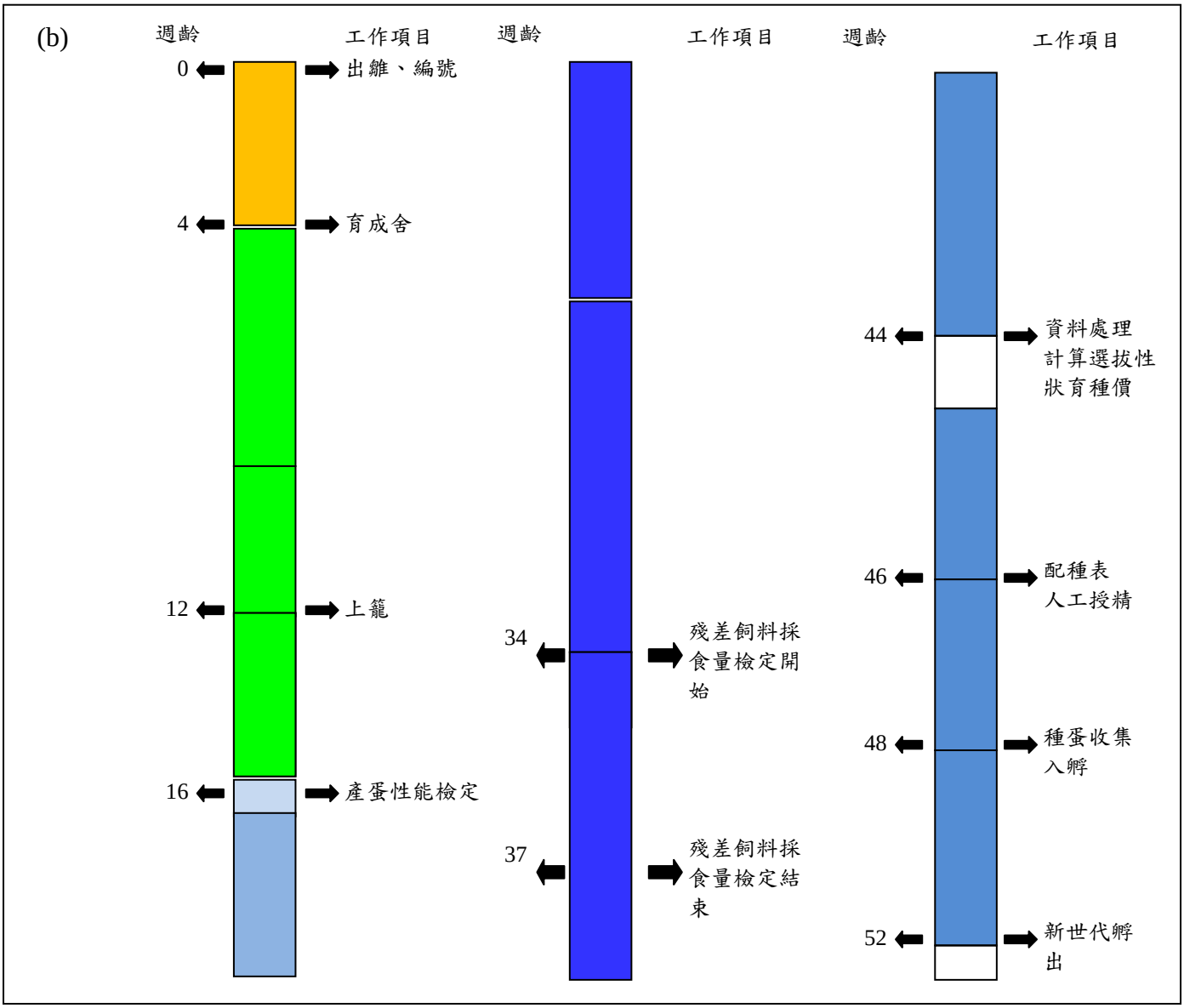


圖 1. 褐色菜鴨高飼效品系(a)殘差飼料採食量檢定圖示(b)選育試驗流程與操作項目。

五、選拔試驗報告

(一)選拔試驗鴨隻及選拔率

試驗設計以小族群系譜選育的方式，於 2009 年選取行政院農業委員會畜產試驗所褐色菜鴨畜試一號第 16 代公鴨 157 隻與母鴨 195隻，後續則每代預計選取母鴨 120 隻作為試驗用種鴨群，進行殘差飼料採食量相關性狀檢定。檢定後，第零代與第一代皆根據殘差飼料採食量表型值選留種母鴨及其全同胞兄弟作為種公鴨，並避開親屬關係進行繁殖；自第二代開始，依殘差飼料採食量性狀的無偏差育種價估測值選留最佳育種價估測值公、母鴨繁殖試驗用鴨群(公鴨之育種價依其系譜關係估算)，每一代並逢機繁殖對照品系 (如表1)。

褐色菜鴨高飼效及對照品系每代各挑選公鴨 10~12 隻、母鴨 40~48 隻用於繁殖，於 44~46 週齡時進行純系配種，於達 50~52 週齡孵化產生下一代，供繼續選育試驗使用。

從第零代至第六代計 7 年，平均世代間距為 1 年，共有公鴨 1,021 隻與母鴨 1,561 隻參與性能檢定，其中公鴨 161 隻與母鴨 600 隻參與配種繁殖。本品系外觀如圖 2 與圖 3 所示。自 2009 年建立本選育鴨群第零代，計算褐色菜鴨高飼效品系計算選拔強度，除第三代因供應商提供之飼料品質問題，造成該世代族群偏小，其餘世代母鴨之選拔率介於 20.8~42.1% 之間，平均為 33.1%，公鴨則介於 12.7~17.1% 之間，平均為 15.7% (如表1)。

圖 2. 褐色菜鴨高飼效品系種公鴨(畜產試驗所楊振豐先生拍攝)。

圖 3. 褐色菜鴨高飼效品系種母鴨(畜產試驗所楊振豐先生拍攝)。

表 1. 褐色菜鴨高飼效品系之族群結構

世代	孵化日期	族群	檢定隻數	留種數	選拔率(%)
第零代	11/17/2008~ 12/1/2008		M = 157 F = 195	M = 20 F = 82	12.7 42.1
第一代	4/19/2010~ 5/3/2010	選拔	M = 66 F = 119	M = 10 F = 42	15.2 35.3
		對照	M = 78 F = 120	M = 11 F = 46	14.1 38.3
第二代	6/29/2011~ 7/13/2011	選拔	M = 70 F = 125	M = 12 F = 26	17.1 20.8
		對照	M = 83 F = 123	M = 12 F = 61	14.5 49.6
第三代	6/25/2012~ 7/9/2012	選拔	M = 42 F = 46	M = 12 F = 38	28.6 82.6
		對照	M = 84 F = 89	M = 12 F = 46	14.3 51.7
第四代	7/1/2013/~ 7/15/2013	選拔	M = 65 F = 104	M = 11 F = 42	16.9 40.4
		對照	M = 81 F = 118	M = 13 F = 46	16.0 39.0
第五代	7/4/2014~ 7/18/2014	選拔	M = 71 F = 136	M = 12 F = 40	16.9 29.4
		對照	M = 82 F = 144	M = 12 F = 40	14.6 27.8
第六代	6/22/2015~ 7/6/2015	選拔	M = 78 F = 148	M = 12 F = 45	15.4 30.4
		對照	M = 64 F = 94	M = 12 F = 46	18.8 48.9
合計			M= 1,021 F = 1,561	M = 161 F = 600	

(二)近親係數

本品系經 6 代選拔之後，計算褐色菜鴨高飼效品系每世代之近親係數平均值與標準偏差，公鴨與母鴨之近親係數平均值結果相近，假設種原來源無親屬關係，則第零代、第一代與第二代之近親係數平均值均為 0，褐色菜鴨高飼效品系親代之選拔，係依高遺傳估測值之排序避開全同胞與半同胞之配種方式來進行，其近親係數適度緩慢的逐漸增加，至第六代時褐色菜鴨高飼效品系公、母鴨近親係數同為 0.083 ± 0.017 (如表 2)。

表 2. 褐色菜鴨高飼效品系各代公鴨與母鴨平均近親係數

世代	公鴨		母鴨	
第零代	0	(N = 10)	0	(N = 121)
第一代	0	(N = 66)	0	(N = 119)
第二代	0	(N = 70)	0	(N = 125)
第三代	0.045 ± 0.015	(N = 42)	0.045 ± 0.012	(N = 46)
第四代	0.056 ± 0.033	(N = 65)	0.063 ± 0.036	(N = 104)
第五代	0.067 ± 0.027	(N = 71)	0.067 ± 0.025	(N = 136)
第六代	0.083 ± 0.017	(N = 78)	0.083 ± 0.017	(N = 148)

(三)調查項目性能與表型相關

為探討褐色菜鴨殘差飼料採食量之選育效果，根據個體在34~37週齡之飼料採食量 (FC)、34~37週齡蛋產量 (EM)、體重變化(ΔBW)及檢定結束體重 (BW) 計算個體之預測飼料採食量 (pFC) 及殘差飼料採食量 (RFC)；將每代資料累積後利用系譜之親屬關係資料，進行選拔性狀最佳線性無偏差預測值 (BLUP) 之統計分析後，供評估比較殘差飼料採食量各性狀之差異及遺傳改進使用。

第六代初產日齡為 116 ± 10 日；40 週齡累積產蛋數為 145 ± 18 個；52 週齡累積產蛋數為 203 ± 28 個。第零代至第六代統計各項性能之平均值 ± 標準偏差結果如表 3 與表 4 所示，褐色菜鴨高飼效、對照品系第六代之 FC 分別為 3859.6 ± 480.2 g、4236.7 ± 508.3 g；pFC 分別為 4027.7 ± 262.5 g、3965.6 ± 215.6 g；RFC 分別為 -168.1 ± 388.5 g、271.1 ± 450.1 g；FE 分別為 2.3 ± 0.4、2.5 ± 0.4；EM 分別為 1699.2 ± 206.3 g、1706.0 ± 166.0 g。經計算褐色菜鴨高飼效品系各項性狀表型相關，FC 與 RFC 及 EM 性狀呈極顯著正相關，與 FE 無顯著相關；而 FE 亦與 EM 呈極顯著負相關。其中 FC 與 RFC 之相關係數達 0.70 ，而 FE 與 EM 之相關係數則達 -0.66，其餘相關係數為低至中度相關 (如表 5)。

表 3. 褐色菜鴨高飼效品系第六代性狀表型值平均、標準偏差、範圍、歪斜度及峰度 (N = 148)

性狀	平均	標準偏差	範圍	歪斜度	峰度
FC[1]	3859.6	480.2	2398.0~4993.0	-0.22	0.09
pFC[2]	4027.7	262.5	2681.6~4601.6	-1.42	5.15
RFC[3]	-168.1	388.5	-1241.8~1144.1	0.58	0.45
FE[4]	2.3	0.4	1.8~4.4	2.76	10.90
EM[5]	1699.2	206.3	672.0~2337.0	-1.51	5.70
AFE[6]	116.1	10.0	100~144	0.63	-0.11
EN40[7]	145.3	18.1	53~173	-1.61	4.57
EN52[8]	202.5	28.4	86~255	-0.85	1.26

[1]: 34~37 週齡飼料採食量 (g)

[2]:預測飼料採食量 (g)

[3]:殘差飼料採食量 (g)

[4]:飼料換蛋率

[5]: 34~37 週齡蛋產量 (g)

[6]:初產日齡

[7]: 40 週齡累積產蛋數

[8]: 52 週齡累積產蛋數

表 4. 褐色菜鴨高飼效及對照品系第零代到第六代五項性狀表型值平均及標準偏差

性狀	第零代	第一代		第二代		第三代		第四代		第五代		第六代	
	(N=195)	選拔 (N=119)	對照 (N=120)	選拔 (N=125)	對照 (N=123)	選拔 (N=46)	對照 (N=89)	選拔 (N=104)	對照 (N=118)	選拔 (N=136)	對照 (N=144)	選拔 (N=148)	對照 (N=94)
FC[1]	3502.7	3897.4	3865.5	3635.2	3756.6	3270.1	3074.0	3811.0	3751.9	3759.1	3892.5	3859.6	4236.7
	± 674.2	± 529.2	± 529.1	± 471.8	± 449.2	± 562.6	± 571.0	± 399.7	± 390.5	± 494.5	± 466.0	± 480.2	± 508.3
pFC[2]	3507.2	3971.1	3772.0	3749.6	3645.3	3306.3	2988.0	3877.0	3699.4	3852.9	3813.0	4027.7	3965.6
	± 494.4	± 382.0	± 397.6	± 334.2	± 313.2	± 437.2	± 538.0	± 231.6	± 211.6	± 220.9	± 232.0	± 262.5	± 215.6
RFC[3]	-4.5	-73.6	93.6	-114.7	111.0	-36.2	85.9	-59.6	52.6	-93.8	78.9	-168.1	271.1
	± 473.3	± 336.6	± 333.9	± 302.5	± 316.0	± 329.4	± 331.0	± 313.6	± 309.4	± 445.2	± 396.0	± 388.5	± 450.1
FE[4]	3.1	2.9	3.3	2.7	2.8	3.2	3.2	2.3	2.4	2.4	2.6	2.3	2.5
	± 3.3	± 1.0	± 1.4	± 0.8	± 0.8	± 0.8	± 0.9	± 0.5	± 0.6	± 0.5	± 0.7	± 0.4	± 0.4
EM[5]	1513.7	1343.0	1228.1	1422.3	1430.6	1042.8	866.0	1720.1	1629.5	1609.0	1595.0	1699.2	1706.0
	± 462.3	± 440.7	± 489.1	± 299.5	± 306.2	± 327.7	± 448.0	± 245.0	± 257.3	± 257.4	± 334.0	± 206.3	± 166.0

[1]: 34~37 週齡飼料採食量 (g)

[2]:預測飼料採食量 (g)

[3]:殘差飼料採食量 (g)

[4]:飼料換蛋率

[5]: 34~37 週齡蛋產量 (g)

(四) 遺傳參數與遺傳相關

依據劉等 (2012)，以四週為檢定期之殘差飼料採食量與檢定全期 (22~52 週齡) 之表型相關達 0.90，遺傳相關則介於 0.93 至 1 之間，分析殘差飼料採食量遺傳率介於 0.3 至 0.43 之間，顯示褐色菜鴨的殘差飼料採食量性狀係可選拔者。

劉 (2015) 後續進一步以 34 週齡開始為期 4 週的飼料採食量、總蛋重、鴨隻檢定結束體重及體重變化進行檢定，持續 6 代後，結果顯示褐色菜鴨高飼效品系與對照品系之殘差飼料採食量育種價估測值在第一至第六代平均分別為 -22.1、-62.1、-111.0、-107.8、-172.5 及 -246.0 g 與 23.0、14.6、15.1、-30.4、-58.1 及 -22.5 g，品系間育種價差異 (S-C) 則為 -45.1、-76.7、-126.1、-77.4、-114.4 及 -223.5 g。褐色菜鴨高飼效品系之殘差飼料採食量不論自表型值或是育種價估測值，皆較對照品系為佳。

截至目前，共已收集 6 代資料，據以進行相關選拔效率評估。結果顯示褐色菜鴨高飼效品系之 FC、RFC、FE 及 EM 之遺傳率分別為 0.33、0.12、0.13 及 0.32。而遺傳相關之分析結果顯示，EM 則與 FC 呈高度正相關，而 RFC 與 FC 及 EM 呈中度正相關，與 FE 則呈中度負相關，顯示 RFC 選拔可間接改善 FE。FC 與 FE 呈中度負相關；而 FE 與 EM 則呈輕度負相關 (如表 5)。

表 5. 褐色菜鴨高飼效品系四項性狀之遺傳率(對角線)、遺傳相關±標準機差 (對角線上方)、表型相關(對角線下方)

	FC	RFC	FE	EM
FC	0.33±0.04	0.59±0.08	-0.32±0.15	0.83±0.05
RFC	0.70**	0.12±0.05	-0.30±0.24	0.37±0.12
FE	-0.03	0.29**	0.13±0.04	-0.11±0.17
EM	0.49**	-0.04	-0.66**	0.32±0.05

** 表具極顯著差異 ($P < 0.0001$)

(五) 褐色菜鴨高飼效品系殘差飼料採食量

殘差飼料採食量 (residual feed consumption; RFC) 的測定工作早在 1941 年由馬里蘭大學的 Dr. Byerly 所提出 (Bordas and Minvielle, 1999)。在預測工具中，體重、產蛋重量、體重變化是最常運用於線性迴歸預測採食量的因子，而諸多研究顯示殘差飼料採食量係屬高遺傳變異率者(0.4~0.5, Lutting and Urff, 1991)。

宜蘭分所之褐色菜鴨高飼效族群自 2009 年建立，分為褐色菜鴨高飼效品系 (S) 與對照品系 (C)，迄今已完成第六代檢定與選拔，比較第一代至第六代之褐色菜鴨高飼效與對照品系，其平均殘差飼料採食量差距 (S-C) 分別為 -167.2、-225.7、-122.1、-112.2、-172.7 及 -439.2 g；第一代至第六代品系間殘差飼料採食量育種價差異(S-C)則為 -45.11、-76.7、-126.1、-77.4、-114.4 及 -223.5 g (圖 4)，已有極大差距，接近洛島紅雞隻殘差飼料採食量雙向選拔9年之成果。依第六代34~37週 (共28日) 殘差飼料採食量育種價估測值估算，相較於對照品系，褐色菜鴨高飼效品系每日殘差飼料採食量少達 8.0 g (223.5/28=8.0)，以產蛋期10個月估算，每隻產蛋母鴨約可節省 2.4 kg (8.0 g × 300 日=2.4 kg) 飼料成本支出。以臺灣平均在產蛋鴨150萬隻為計算基礎，目前選拔成果預估每年約可為農民省下新臺幣 4,680 萬元 (2.4 kg × 150萬隻 × 蛋鴨料 13 元/kg) 之飼料成本。

與同以殘差飼料採食量選拔蛋禽研究相較，洛島紅雞隻在經歷長期之殘差飼料採食量雙向選拔後，高飼料採食品系 (R+) 與低飼料採食品系 (R-) 在 32~36 週齡之殘差飼料採食量及飼料總採食量等性狀呈現顯著性差異。殘差飼料採食量與總蛋重 (egg mass)、產蛋數、蛋重及體重均無顯著表型相關，而與飼料採食量之遺傳相關則高達 0.5 (Bordas et al., 1992; Bordas et al., 1996)，與本品系之 0.59 相近。又如 Basso et al. (2012) 建立北京鴨殘差飼料採食量檢定方法並估測其遺傳參數，其中採食量、蛋產量及殘差飼料採食量的遺傳率分別為 0.34、0.06 及 0.24，與本品系相較，採食量遺傳率幾乎相同，蛋產量低於本品系，而殘差飼料採食量則稍高於本品系。而殘差飼料採食量與採食量之遺傳相關更高達 0.89，

高於本品系甚多。

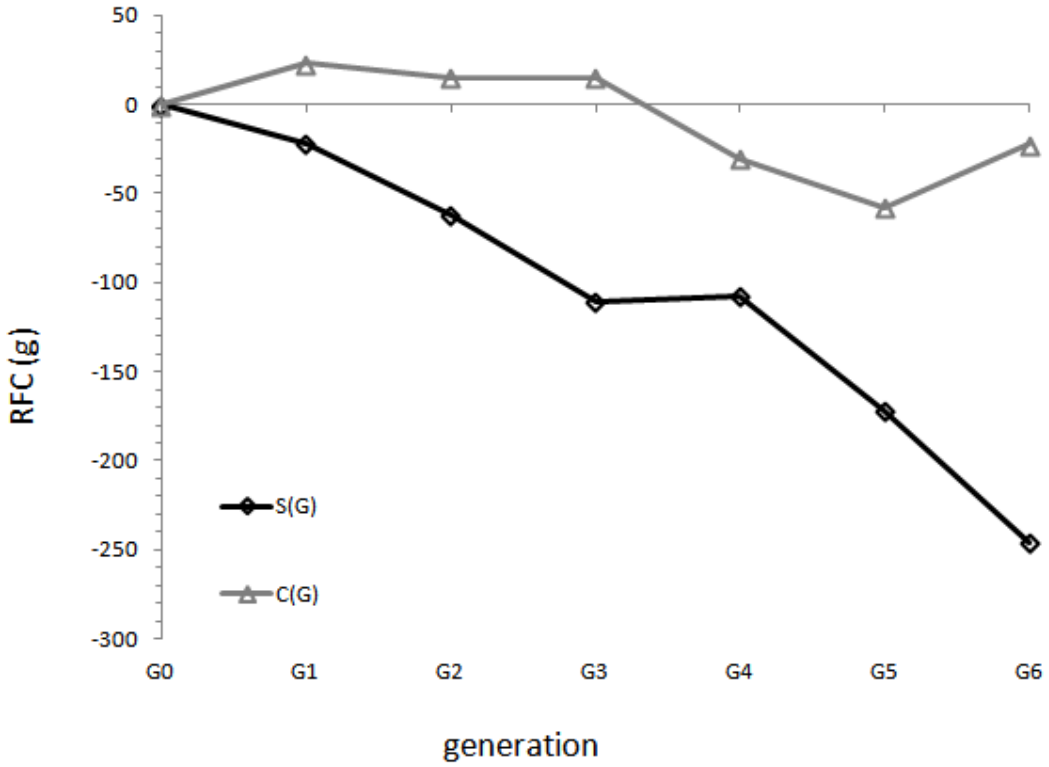


圖 4. 褐色菜鴨高飼效(S) 與對照品系 (C)殘差飼料採食量育種價估測值 (G)。

(六) 遺傳形質

為評估褐色菜鴨高飼效品系於選拔過程中之遺傳變異變化，本試驗分別自第二與第四代之留種母鴨挑選後代 1 公 1 母，共 90 隻，利用 11 組自菜鴨基因組 DNA 篩選之微衛星標記 (Hsiao et al., 2008) 進行分析，結果分別如表 6 與表 7 所示。11 組標記中共有 5 組標記在雞 (*Gallus gallus*) 染色體上具有直系同源基因座；於鴨 (*Anas platyrhynchos*) 基因組 DNA 方面，11 組標記分屬不同基因組架構 (scaffold)。遺傳分析結果，褐色菜鴨對照品系與高飼效品系分別觀測到 47 與 45 個交替基因，所有微衛星標記皆具中度 (0.25 ≦ PIC < 0.5) 至高度多態性 (PIC ≧ 0.5)，顯示此 11 組微衛星標記應可用於褐色菜鴨對照及高飼效品系之遺傳分析。分別混合兩個世代之褐色菜鴨對照品系與高飼效品系來看，兩者平均每個基因座各具有 4.3 個 (2~8個) 與 4.1 個 (2~7個) 交替基因；褐色菜鴨對照品系之觀測異質度 (observed heterozygosity) 介於 0.191 到 0.851，平均為 0.571，高飼效品系則介於 0.233 到0.837，平均為 0.522；而褐色菜鴨對照品系之期望異質度 (expected heterozygosity) 介於 0.397 到 0.771，

平均為 0.642，高飼效品系則介於 0.410 到 0.705，平均為 0.588。比較褐色菜鴨對照與高飼效品系間與第二、第四代間之交替基因數、有效交替基因數、觀測異質度、期望異質度，結果顯示各項參數在兩組間或世代間皆無顯著差異。

族群結構分析方面，11 組微衛星標記在褐色菜鴨對照品系與高飼效品系之第二與第四代共4個次族群分析結果皆符合哈溫平衡。再依 Wright (1965) 方法計算懷特氏固定指數 (Wright's fixation index) 近交指數 F_{IS}，褐色菜鴨對照品系與高飼效品系之平均分別為 0.118 與 0.117，且各標記之 F_{IS} 值差異大，經雙尾t檢定後，已知各次族群之平均族群近交指數與 0 皆無顯著差異 ($p > 0.05$)，顯示褐色菜鴨對照與高飼效品系均尚無嚴重近親衰退情事。另以 STRUCTURE 2.3 軟體 (Pritchard et al., 2000) 進行群數分布之模擬分析，設 K 為 2 (圖 5)，結果顯示第二代時，褐色菜鴨對照與高飼效品系之遺傳組成大致相同；然至第四代，其分群與樣本組別來源幾乎一致，顯示褐色菜鴨對照與高飼效品系之遺傳組成已有差異。表 8 則顯示，無論是在第二代或第四代，褐色菜鴨對照與高飼效品系間之遺傳分化指數 (F_{ST}) 皆指向褐色菜鴨對照與高飼效品系間具有顯著分化 ($P < 0.05$)；然在第二代之分化程度極低 (F_{ST} =0.0328)，至第四代時才有中度分化 (F_{ST} =0.0505)。若將對照與高飼效品系分別來看，第二代與第四代間之分化皆不顯著，惟高飼效品系之兩世代間遺傳分化指數略高於對照品系。以上結果皆說明褐色菜鴨高飼效品系與對照品系之族群分化有隨著殘差飼料採食量選拔代數增加而有加深之現象。

表 6. 應用 11 組菜鴨微衛星標記於對照品系第二代與第四代之遺傳變異

基因座	雞染色體編號[1]/鴨基因組架構編號[2]	第二代 (N = 23)							第四代 (N = 24)						
		片段大小	交替基因數	有效交替基因數	觀測異質度	期望異質度	多態性訊息含量	族群近交係數	片段大小	交替基因數	有效交替基因數	觀測異質度	期望異質度	多態性訊息含量	族群近交係數
APT001	1 / 1509	178~206	3	2.1	0.217	0.529	0.421	0.590	178~202	2	1.9	0.333	0.496	0.368	0.329
APT004	3 / 192	290~314	7	4.1	0.913	0.774	0.722	-0.180	290~314	7	3.5	0.500	0.727	0.670	0.312
APT008	NA / 358	184~196	4	3.5	0.870	0.727	0.662	-0.197	184~196	5	3.4	0.083	0.760	0.706	0.891
APT010	NA / 1199	192~212	4	3.1	0.696	0.689	0.610	-0.010	192~212	3	2.8	0.667	0.656	0.570	-0.017
APT012	2 / 5	185~205	5	3.4	0.739	0.724	0.656	-0.021	185~205	4	3.8	0.667	0.749	0.685	0.109
APT017	1 / 481	173~185	4	3.0	0.391	0.683	0.611	0.428	173~185	4	2.4	0.417	0.588	0.517	0.291
APT020	NA / 197	177~201	5	3.8	0.783	0.757	0.697	-0.034	177~201	5	4.4	0.917	0.791	0.737	-0.159
APT025	NA / 121	105~117	4	1.5	0.391	0.347	0.321	-0.127	105~117	4	1.8	0.542	0.448	0.406	-0.210
APT026	7 / 477	130~146	4	3.2	0.870	0.704	0.640	-0.236	130~146	4	4.0	0.833	0.765	0.702	-0.089
APT032	NA / 45	207~259	3	2.4	0.783	0.587	0.482	-0.334	207~259	3	2.2	0.583	0.547	0.430	-0.066
APT033	NA / 14	262~266	2	2.0	0.130	0.507	0.373	0.744	262~266	2	2.0	0.250	0.511	0.375	0.511
Average			4.1	2.9	0.617	0.639	0.563	0.057		3.9	2.9	0.527	0.640	0.561	0.173
SD			1.3	0.8	0.281	0.132	0.139	0.361		1.4	0.9	0.248	0.126	0.146	0.333

[1]: 直系同源基因位於雞染色體編號

[2]: 直系同源基因位於鴨基因組架構編號

[3]: 懷特氏固定指數之族群近交係數 F_{IS} =1-(觀測異質度/期望異質度)(Wright, 1965)

表 7. 應用 11 組菜鴨微衛星標記於褐色菜鴨高飼效品系第二代與第四代之遺傳變異

基因座	雞染色體編號[1]/鴨基因組架構編號[2]	第二代 (N = 17)							第四代 (N = 26)						
		片段大小	交替基因數	有效交替基因數	觀測異質度	期望異質度	多態性訊息含量	族群近交係數	片段大小	交替基因數	有效交替基因數	觀測異質度	期望異質度	多態性訊息含量	族群近交係數
APT001	1 / 1509	178~206	3	2.5	0.235	0.620	0.531	0.621	174~206	4	2.6	0.385	0.624	0.538	0.383
APT004	3 / 192	290~314	6	3.9	0.647	0.765	0.707	0.154	286~314	7	2.8	0.692	0.658	0.605	-0.052
APT008	NA / 358	184~196	4	2.5	0.471	0.619	0.518	0.239	184~196	5	1.8	0.154	0.489	0.439	0.685
APT010	NA / 1199	192~212	4	2.0	0.529	0.513	0.457	-0.031	192~212	4	2.4	0.577	0.597	0.541	0.034
APT012	2 / 5	185~205	4	2.7	0.588	0.651	0.561	0.097	185~205	4	2.5	0.615	0.611	0.518	-0.007
APT017	1 / 481	173~185	4	2.8	0.765	0.668	0.580	-0.145	173~185	4	2.5	0.577	0.608	0.535	0.051
APT020	NA / 197	185~201	5	2.7	0.647	0.651	0.578	0.006	185~201	5	2.8	0.615	0.655	0.579	0.061
APT025	NA / 121	105~117	4	1.4	0.353	0.316	0.285	-0.117	105~117	4	1.9	0.500	0.472	0.431	-0.059
APT026	7 / 477	130~146	4	3.1	0.824	0.693	0.620	-0.189	130~146	4	3.4	0.846	0.722	0.659	-0.172
APT032	NA / 45	207~259	2	1.9	0.647	0.487	0.361	-0.329	207~259	2	1.9	0.423	0.491	0.366	0.138
APT033	NA / 14	262~266	2	1.6	0.118	0.371	0.295	0.682	262~266	2	2.0	0.308	0.498	0.369	0.382
Average			3.8	2.5	0.529	0.578	0.499	0.090		4.1	2.4	0.517	0.584	0.507	0.131
SD			1.2	0.7	0.219	0.139	0.136	0.321		1.4	0.5	0.192	0.084	0.095	0.252

[1]: 直系同源基因位於雞染色體編號

[2]: 直系同源基因位於鴨基因組架構編號

[3]: 懷特氏固定指數之族群近交係數 F_{IS} =1-(觀測異質度/期望異質度)(Wright, 1965)

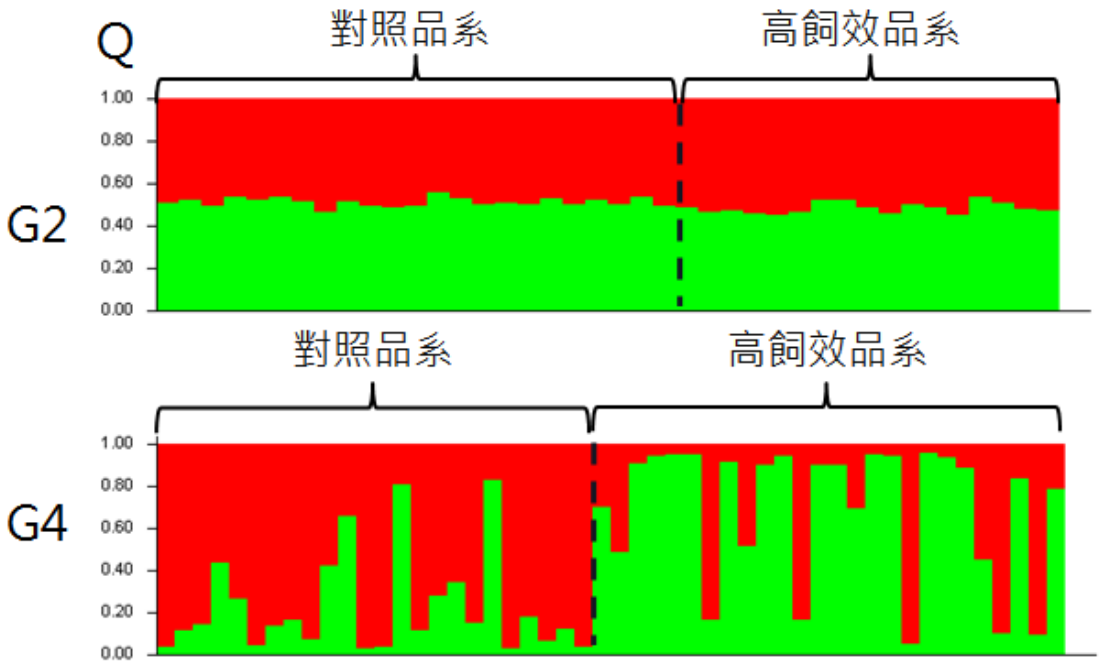


圖 5. 褐色菜鴨高飼效及對照品系第二代與第四代 STRUCTURE 分群分析之結果。此圖可能分群數(K)設為 2；Q：個體基因源自該群集之比例，不同群集以不同顏色表示，縱軸每圖條代表一個體。

表 8. 應用 11 組微衛星標記分析褐色菜鴨高飼效、對照品系第二代與第四代族群之遺傳分化指數 (F_{ST}，對角線右上)

	對照品系第四代	高飼效品系第二代	高飼效品系第四代
對照品系第二代	0.0060	0.0328*	0.0641*
對照品系第四代		0.0600*	0.0505*
高飼效品系第二代			0.0219

*具有顯著分化 ($P < 0.05$)

六、飼養管理及衛生防疫措施

(一) 飼養管理

1. 試驗用鴨群依菜鴨模式飼養管理，其主要流程如下：

   (1) 出生登記(系譜建立、蹼號)。
   (2) 育雛期(0 至 4 週齡)。
   (3) 育成期(4 至 12 週齡)。
   (4) 上籠(12 週齡)。
   (5) 初產 (16 週齡)。
   (6) 產蛋檢定開始(初產至 52 週齡)。
   (7) 殘差飼料採食量檢定(34 至 37 週齡)
   (8) 產蛋檢定結束 (52 週齡)。
   (9) 依配種表人工授精。
   (10) 種蛋收集。
   (11) 入孵。
   (12) 新世代孵出。

2. 育雛期 (0~4 週齡) 及育成前期 (4~8 週齡) 餵飼含粗蛋白 19.5%、代謝能 2,909 kcal/kg 飼糧；育成後期 (8 週齡~初產) 餵飼含粗蛋白 13.5 %、代謝能 2,660 kcal/kg 飼糧；產蛋期餵飼含粗蛋白 20%、代謝能 2,712 kcal/kg 飼糧 (表 9)，任飼、水自由飲用。

3. 鴨隻 0~4 週齡配合保溫燈採高床育雛，4 週齡後移入育成舍 (採高床育成)，12 週齡後公、母鴨隻移入鋼構鴨舍(有水簾、風扇裝置)，採個別籠飼方式飼養，每隻活動面積為 990 cm^2 (長 30 cm * 寬 33 cm * 高 45 cm)飼養。

4. 飼養過程中，按防疫計畫接種疫苗，產蛋期每日補充 15~20 lux 光照 14 小時，飼料及飲水採任食。

表 9. 各期飼料配方　　單位：kg/100kg

成分	9 號料 (0~8 週齡育雛料)	8 號料 (8 週齡~初產前育成料)	蛋鴨料 (產蛋期)
玉米粉	55.3	51.7	49.7
紅麩皮	-	10.0	6.5
大麥粉	10.3	20.0	-
大豆粕	25.3	10.0	27.0
魚粉	2.0	-	3.3
酵母粉	3.0	2.0	2.0
粗糠粉	-	2.4	-
甲硫胺酸	0.05	0.05	0.05
磷酸氫鈣	1.1	1.5	1.5
石灰石粉	1.1	1.6	6.6
粗鹽	0.3	0.3	0.4
大豆油	1.1	-	2.5
維生素、礦物質預混物	0.5	0.5	0.5
計算值			
粗蛋白質 (%)	19.5	13.5	20
代謝能 (kcal/kg)	2,909	2,660	2,712
鈣 (%)	0.81	0.94	3.05
有效磷 (%)	0.36	0.44	0.44
含硫胺基酸(%)	0.70	0.51	0.70

(二) 衛生防疫

1. 防疫措施

依據畜產試驗所宜蘭分所訂定之「畜禽飼養場所自衛防疫措施」之規定辦理：

(1) 場區於出入口設置消毒管制站，嚴禁民間淘汰鴨隻車輛進入場區；管制車輛(飼料車、運食蛋車、工程車)與一般洽公車輛需循本分所之大門口消毒標準作業程序進行消毒後始可放行。

(2) 工作人員確實更換工作衣物，經消毒槽進出場區。

(3) 嚴禁非場內車輛進入場區；載運飼料、運輸鴨隻及鴨舍維修等必須進入場區之車輛，應經由消毒池與消毒噴霧後進入，消毒池液面長度至少需為輪胎圓周長之 1.5 倍以上。

(4) 非場區工作人員進入鴨舍，應換妥乾淨衣物並洗手後，經消毒槽進出。

(5) 散裝飼料桶應盡量設於場區外圍，以降低非本場車輛進入場區頻率，並避免司機及隨同人員進入鴨舍內。

(6) 外購鴨隻應隔離飼養檢疫，並實施必要之免疫工作，確定無帶疫病後方可移入場內。

(7) 鴨舍保持通風、清潔及乾燥，排水溝保持暢通，以防止蚊蠅孳生。

(8) 鴨舍及設備之消毒，可分為定期消毒及不定期消毒兩種。當鴨舍有空欄時，隨時清洗並消毒。消毒劑之使用，應配合季節、天氣、場合及防治對象，至少準備 3 種以上不同種類之消毒劑輪流使用，俾降低場內病原濃度。

(9) 鴨舍空出時，必須徹底洗刷乾淨，並經消毒後應空置一週以上才能引進鴨隻飼養，飼養前再進行消毒一次。

(10) 注射用器械應保持清潔衛生，以避免機械性傳播病原。

(11) 每棟鴨舍出入口均應設置消毒槽，供員工進出消毒用。

(12) 設置防鳥網，加強野貓、狗及飛鳥驅離工作，定期施放毒餌滅鼠；手飼用之飼料車勿剩餘飼料於其內，或將飼料車加蓋，以杜絕鼠類及飛鳥等進入覓食。

(13) 每日巡視，及早發現罹病鴨隻，治療期間應採取隔離措施，以防止疫病傳染；治療無效者予以淘汰，以避免疾病擴散。

(14) 隨時注意國內外疫情報導，以採取必要之防疫措施。

(15) 每月檢視「畜牧場衛生管理工作紀錄簿」、「大門口出入人員紀錄表」及「大門口消毒水配製紀錄表」是否詳實紀錄。

(16) 每月依據「內部控制檢查表」查核飼養單位衛生防疫措施。

(17) 發現可疑疫情或鴨隻死亡數異常升高時，向特約獸醫師通報，立即限制發病鴨群移動，同時加強各項防疫措施，防止疾病擴散。

2. 疫苗注射計畫

依據行政院農業委員會畜產試驗所宜蘭分所生物安全防疫作業程序說明表辦理。鴨隻分別於 4 及 8 週齡實施家禽霍亂疫苗免疫注射。購買疫苗時注意檢驗合格封籤、包裝完整性及有效期限，並妥善保存；疫苗注射情形列冊記錄備查。

七、參考文獻

1. 劉秀洲。2015。種畜禽研究團隊—高飼效褐色菜鴨之選育與應用。行政院農業委員會畜產試驗所 104 年度科技計畫研究報告。
2. 劉秀洲、杜宗哲、C. Marie-Etancelin、李淵百、黃振芳、陳志峰。2012。褐色菜鴨殘差飼料採食量之遺傳參數估算。畜產研究 45(2)：131-140。
3. Basso, B., A. Bordas, F. Dubos, P. Morganx, and C. Marie-Etancelin. 2012. Feed efficiency in the laying duck: Appropriate measurements and genetic parameters. Poult. Sci. 91:1065-1073.
4. Bordas, A. and F. Minvielle. 1999. Patterns of growth and feed intake in divergent lines of laying domestic fowl selected for residual feed consumption. Poult. Sci. 78:317-323.
5. Bordas, A., M. Tixier-Boichard and P. Mérat. 1992. Direct and correlated responses to divergent selection for residual food intake in Rhode Island Red laying hens. Br. Poult. Sci. 33:741-754.
6. Bordas, A., and P. Mérat. 1981. Genetic variation and phenotypic correlation of food consumption of laying hens corrected for body weight and production. Br. Poult. Sci. 22:25-33.
7. Bordas, A., P. Mérat and F. Minvielle. 1996. Heterosis in egg laying lines under divergent selection for residual feed consumption. Poult. Sci. 75 :20-24.
8. Hsiao, M. C., H. C. Liu, Y. C. Hsu, Y. H. Hu, S. H. Li and S. R. Lee. 2008. Isolation and characterization of microsatellite markers in Tsaiya duck. Asian-Aust. J. Anim. Sci. 21: 624-627.
9. Koch, R. M., L. A. Swiger, D. Chambers, and K. E. Gregory. 1963. Efficiency of feed use in beef cattle. J. Anim. Sci. 22:486-494.
10. Lutting, P. and E. M. Urff. 1991. Residual feed consumption in laying hens. 2. Genetic variation and correlations. Poult. Sci. 70:1663-1672.
11. Pritchard, J. K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.
12. Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395-420.

八、育種人員姓名及其資歷

研究項目	研究年別	研究人員	職稱
1.品系性能觀察	2009	劉秀洲	副研究員兼系主任
		鄭裕信	研究員兼副所長
		蕭孟衿	約聘人員
		李舜榮	研究員兼分所長
		黃英豪	所長
2.選拔品系建立	2009	劉秀洲	副研究員兼系主任
		鄭裕信	研究員兼副所長
		蕭孟衿	約聘人員
		李舜榮	研究員兼分所長
		黃英豪	所長
3.殘差飼料採食量選育	2010~2015	劉秀洲	副研究員兼系主任
		陳志峰	中興大學教授
		李舜榮	研究員兼分所長
		黃振芳	研究員兼分所長
		鄭裕信	研究員兼副所長
		黃英豪	所長
4.命名資料整理	2016~	張怡穎	助理研究員
		劉秀洲	研究員兼分所長
		魏良原	副研究員兼系主任
		黃振芳	研究員兼副所長
		鄭裕信	所長
5.命名申請	2016~	張怡穎	助理研究員
		劉秀洲	研究員兼分所長
		魏良原	副研究員兼系主任
		林秀蓮	助理研究員
		施柏齡	副研究員
		洪哲明	副研究員
		吳明哲	研究員兼組長
		黃振芳	研究員兼副所長
		鄭裕信	所長

附錄一：褐色菜鴨高飼效品系申請新品系登記族群系譜一覽表

附錄二：褐色菜鴨高飼效品系第二代及第四代微衛星遺傳標記分析資料

附錄三、褐色菜鴨高飼效品系營運計畫書

附錄一：高飼效褐色菜鴨申請新品種登記族群系譜一覽表

序號	鴨號	性別	品系	出生日期	公鴨號	母鴨號	祖父鴨號	祖母鴨號	外祖父鴨號	外祖母鴨號
1	220003	1	S	2015/6/22	211114	210684	200460	200846	200805	200606
2	220004	2	S	2015/6/22	211114	210684	200460	200846	200805	200606
3	220005	2	S	2015/6/22	211114	210684	200460	200846	200805	200606
4	220008	2	S	2015/6/22	211114	210684	200460	200846	200805	200606
5	220010	2	S	2015/6/22	210841	210831	200555	200983	200460	200948
6	220011	1	S	2015/6/22	211030	210554	200350	200084	200611	200406
7	220013	1	S	2015/6/22	211030	210554	200350	200084	200611	200406
8	220014	2	S	2015/6/22	211030	210554	200350	200084	200611	200406
9	220015	1	S	2015/6/22	211114	210085	200460	200846	200555	200165
10	220016	2	S	2015/6/22	211114	210085	200460	200846	200555	200165
11	220018	2	S	2015/6/22	211114	210085	200460	200846	200555	200165
12	220030	2	S	2015/6/22	211114	210085	200460	200846	200555	200165
13	220031	1	S	2015/6/22	210199	210606	200161	200556	200460	200399
14	220035	2	S	2015/6/22	210199	210606	200161	200556	200460	200399
15	220036	2	S	2015/6/22	210199	210606	200161	200556	200460	200399
16	220040	1	S	2015/6/22	210841	210698	200555	200983	200441	200405
17	220043	2	S	2015/6/22	210199	210663	200161	200556	200350	200358
18	220044	1	S	2015/6/22	211095	210163	200196	200404	200460	200846
19	220045	2	S	2015/6/22	210869	210689	200161	200915	200555	200165
20	220046	1	S	2015/6/22	210064	210614	200161	200613	200196	200869
21	220048	2	S	2015/6/22	210064	210614	200161	200613	200196	200869
22	220050	2	S	2015/6/22	210199	210306	200161	200556	200585	200344
23	220051	2	S	2015/6/22	210199	210306	200161	200556	200585	200344
24	220054	2	S	2015/6/22	210819	210690	200460	200948	200555	200165
25	220055	1	S	2015/6/22	210841	210019	200555	200983	200350	200494
26	220056	1	S	2015/6/22	210841	210019	200555	200983	200350	200494
27	220058	2	S	2015/6/22	210193	210630	200161	200114	200805	200303
28	220060	2	S	2015/6/22	210193	210630	200161	200114	200805	200303
29	220061	2	S	2015/6/22	210193	210630	200161	200114	200805	200303
30	220063	2	S	2015/6/22	210193	210630	200161	200114	200805	200303
31	220065	1	S	2015/6/22	210193	210630	200161	200114	200805	200303
32	220066	1	S	2015/6/22	210841	210831	200555	200983	200460	200948
33	220068	1	S	2015/6/22	210841	210831	200555	200983	200460	200948
34	220081	2	S	2015/6/22	210841	210831	200555	200983	200460	200948
35	220083	2	S	2015/6/22	210841	210831	200555	200983	200460	200948
36	220084	1	S	2015/6/22	211095	210844	200196	200404	200555	200983
37	220085	1	S	2015/6/22	211095	210844	200196	200404	200555	200983

38	220086	2	S	2015/6/22	211095	210844	200196	200404	200555	200983
39	220088	2	S	2015/6/22	211095	210844	200196	200404	200555	200983
40	220100	2	S	2015/6/22	211095	210844	200196	200404	200555	200983
41	220101	2	S	2015/6/22	211095	210844	200196	200404	200555	200983
42	220103	1	S	2015/6/22	210006	210904	200369	200890	200611	200586
43	220104	2	S	2015/6/22	210006	210904	200369	200890	200611	200586
44	220105	2	S	2015/6/22	210006	210904	200369	200890	200611	200586
45	220106	2	S	2015/6/22	210006	210904	200369	200890	200611	200586
46	220108	1	S	2015/6/22	211030	210563	200350	200084	200196	200404
47	220110	1	S	2015/6/22	211030	210563	200350	200084	200196	200404
48	220113	1	S	2015/6/22	210193	210311	200161	200114	200460	200440
49	220115	2	S	2015/6/22	210193	210311	200161	200114	200460	200440
50	220116	2	S	2015/6/22	210193	210311	200161	200114	200460	200440
51	220118	2	S	2015/6/22	210193	210311	200161	200114	200460	200440
52	220131	1	S	2015/6/22	210869	210161	200161	200915	200460	200846
53	220133	1	S	2015/6/22	210869	210161	200161	200915	200460	200846
54	220134	2	S	2015/6/22	210869	210161	200161	200915	200460	200846
55	220135	2	S	2015/6/22	210869	210161	200161	200915	200460	200846
56	220136	1	S	2015/6/22	210006	210661	200369	200890	200350	200358
57	220138	2	S	2015/6/22	210006	210661	200369	200890	200350	200358
58	220140	2	S	2015/6/22	210006	210661	200369	200890	200350	200358
59	220141	2	S	2015/6/22	210006	210661	200369	200890	200350	200358
60	220143	1	S	2015/6/22	210199	210908	200161	200556	200611	200586
61	220144	1	S	2015/6/22	210199	210908	200161	200556	200611	200586
62	220145	2	S	2015/6/22	210199	210908	200161	200556	200611	200586
63	220146	2	S	2015/6/22	210199	210908	200161	200556	200611	200586
64	220148	2	S	2015/6/22	210199	210908	200161	200556	200611	200586
65	220150	2	S	2015/6/22	210199	210908	200161	200556	200611	200586
66	220151	2	S	2015/6/22	211030	210554	200350	200084	200611	200406
67	220153	2	S	2015/6/22	211030	210554	200350	200084	200611	200406
68	220154	1	S	2015/6/22	211114	210085	200460	200846	200555	200165
69	220155	2	S	2015/6/22	211114	210085	200460	200846	200555	200165
70	220158	1	S	2015/6/22	210841	210616	200555	200983	200196	200869
71	220160	2	S	2015/6/22	210841	210616	200555	200983	200196	200869
72	220161	2	S	2015/6/22	210841	210616	200555	200983	200196	200869
73	220163	2	S	2015/6/22	210841	210616	200555	200983	200196	200869
74	220164	2	S	2015/6/22	210841	210616	200555	200983	200196	200869
75	220165	2	S	2015/6/22	210199	210606	200161	200556	200460	200399
76	220166	1	S	2015/6/22	210006	210003	200369	200890	200350	200084

77	220168	2	S	2015/6/22	210006	210003	200369	200890	200350	200084
78	220180	2	S	2015/6/22	210006	210003	200369	200890	200350	200084
79	220181	2	S	2015/6/22	210006	210003	200369	200890	200350	200084
80	220183	1	S	2015/6/22	210819	210083	200460	200948	200369	200905
81	220184	2	S	2015/6/22	210819	210083	200460	200948	200369	200905
82	220185	2	S	2015/6/22	210819	210083	200460	200948	200369	200905
83	220188	2	S	2015/6/22	210819	210083	200460	200948	200369	200905
84	220300	1	S	2015/6/22	210658	210339	200350	200358	200441	200499
85	220301	1	S	2015/6/22	210658	210339	200350	200358	200441	200499
86	220303	2	S	2015/6/22	210658	210339	200350	200358	200441	200499
87	220304	2	S	2015/6/22	210658	210339	200350	200358	200441	200499
88	220305	2	S	2015/6/22	210658	210339	200350	200358	200441	200499
89	220306	2	S	2015/6/22	210658	210339	200350	200358	200441	200499
90	220308	1	S	2015/6/22	210064	210846	200161	200613	200555	200983
91	220310	1	S	2015/6/22	210064	210846	200161	200613	200555	200983
92	220311	2	S	2015/6/22	210064	210846	200161	200613	200555	200983
93	220313	2	S	2015/6/22	210064	210846	200161	200613	200555	200983
94	220314	2	S	2015/6/22	210064	210846	200161	200613	200555	200983
95	220315	2	S	2015/6/22	210064	210846	200161	200613	200555	200983
96	220316	1	S	2015/6/22	210064	210664	200161	200613	200350	200358
97	220318	1	S	2015/6/22	210064	210664	200161	200613	200350	200358
98	220330	2	S	2015/6/22	210064	210664	200161	200613	200350	200358
99	220331	2	S	2015/6/22	210064	210664	200161	200613	200350	200358
100	220340	1	S	2015/6/22	210869	210689	200161	200915	200555	200165
101	220343	2	S	2015/6/22	210869	210689	200161	200915	200555	200165
102	220344	2	S	2015/6/22	210869	210689	200161	200915	200555	200165
103	220345	1	S	2015/6/22	211095	210050	200196	200404	200501	200614
104	220346	1	S	2015/6/22	211095	210050	200196	200404	200501	200614
105	220348	2	S	2015/6/22	211095	210050	200196	200404	200501	200614
106	220350	2	S	2015/6/22	211095	210050	200196	200404	200501	200614
107	220351	2	S	2015/6/22	211095	210050	200196	200404	200501	200614
108	220353	2	S	2015/6/22	211095	210050	200196	200404	200501	200614
109	220354	1	S	2015/6/22	210599	210608	200501	200614	200460	200399
110	220355	1	S	2015/6/22	210599	210608	200501	200614	200460	200399
111	220356	2	S	2015/6/22	210599	210608	200501	200614	200460	200399
112	220358	1	S	2015/6/22	210819	210801	200460	200948	200501	200639
113	220361	2	S	2015/6/22	210819	210801	200460	200948	200501	200639
114	220363	2	S	2015/6/22	210819	210801	200460	200948	200501	200639
115	220364	1	S	2015/6/22	211095	210660	200196	200404	200350	200358

116	220365	1	S	2015/6/22	211095	210660	200196	200404	200350	200358
117	220366	2	S	2015/6/22	211095	210660	200196	200404	200350	200358
118	220380	2	S	2015/6/22	210841	210698	200555	200983	200441	200405
119	220381	1	S	2015/6/22	210658	210519	200350	200358	200196	200906
120	220383	2	S	2015/6/22	210658	210519	200350	200358	200196	200906
121	220384	1	S	2015/6/22	210658	210015	200350	200358	200805	200180
122	220385	1	S	2015/6/22	211095	210163	200196	200404	200460	200846
123	220386	2	S	2015/6/22	211095	210163	200196	200404	200460	200846
124	220388	2	S	2015/6/22	211095	210163	200196	200404	200460	200846
125	220389	2	S	2015/6/22	210199	210663	200161	200556	200350	200358
126	220400	2	S	2015/6/22	210199	210663	200161	200556	200350	200358
127	220401	1	S	2015/6/22	210006	210661	200369	200890	200350	200358
128	220403	2	S	2015/6/22	210006	210661	200369	200890	200350	200358
129	220405	1	S	2015/6/22	210006	210063	200369	200890	200161	200613
130	220408	2	S	2015/6/22	210006	210063	200369	200890	200161	200613
131	220410	2	S	2015/6/22	210006	210063	200369	200890	200161	200613
132	220411	2	S	2015/6/22	210006	210063	200369	200890	200161	200613
133	220415	2	S	2015/6/22	210193	210338	200161	200114	200441	200499
134	220416	2	S	2015/6/22	210193	210338	200161	200114	200441	200499
135	220418	2	S	2015/6/22	210193	210338	200161	200114	200441	200499
136	220430	2	S	2015/6/22	210193	210338	200161	200114	200441	200499
137	220431	1	S	2015/6/22	210199	210306	200161	200556	200585	200344
138	220433	1	S	2015/6/22	210199	210306	200161	200556	200585	200344
139	220434	2	S	2015/6/22	210199	210306	200161	200556	200585	200344
140	220436	2	S	2015/6/22	210193	210843	200161	200114	200555	200983
141	220438	2	S	2015/6/22	210193	210843	200161	200114	200555	200983
142	220440	2	S	2015/6/22	210193	210843	200161	200114	200555	200983
143	220441	2	S	2015/6/22	210193	210843	200161	200114	200555	200983
144	220443	1	S	2015/6/22	210193	210843	200161	200114	200555	200983
145	220444	1	S	2015/6/22	210193	210843	200161	200114	200555	200983
146	220445	1	S	2015/6/22	210658	210868	200350	200358	200161	200915
147	220446	1	S	2015/6/22	210658	210868	200350	200358	200161	200915
148	220450	2	S	2015/6/22	210658	210868	200350	200358	200161	200915
149	220451	2	S	2015/6/22	210658	210868	200350	200358	200161	200915
150	220453	2	S	2015/6/22	210658	210868	200350	200358	200161	200915
151	220454	1	S	2015/6/22	210599	210309	200501	200614	200585	200344
152	220455	2	S	2015/6/22	210599	210309	200501	200614	200585	200344
153	220456	1	S	2015/6/22	210064	210614	200161	200613	200196	200869
154	220458	2	S	2015/6/22	210064	210614	200161	200613	200196	200869

155	220460	1	S	2015/6/22	211030	210189	200350	200084	200369	200848
156	220463	2	S	2015/6/22	211030	210189	200350	200084	200369	200848
157	220464	2	S	2015/6/22	211030	210189	200350	200084	200369	200848
158	220465	2	S	2015/6/22	211030	210189	200350	200084	200369	200848
159	220466	2	S	2015/6/22	211030	210189	200350	200084	200369	200848
160	220468	1	S	2015/6/22	210064	210865	200161	200613	200369	200848
161	220480	1	S	2015/6/22	210064	210865	200161	200613	200369	200848
162	220481	2	S	2015/6/22	210064	210865	200161	200613	200369	200848
163	220483	2	S	2015/6/22	210869	210151	200161	200915	200369	200890
164	220484	1	S	2015/6/22	211030	210061	200350	200084	200161	200613
165	220485	1	S	2015/6/22	211030	210061	200350	200084	200161	200613
166	220486	2	S	2015/6/22	211030	210061	200350	200084	200161	200613
167	220488	2	S	2015/6/22	211030	210061	200350	200084	200161	200613
168	220500	2	S	2015/6/22	211030	210061	200350	200084	200161	200613
169	220501	1	S	2015/6/22	210819	210690	200460	200948	200555	200165
170	220503	1	S	2015/6/22	210819	210690	200460	200948	200555	200165
171	220504	2	S	2015/6/22	210819	210690	200460	200948	200555	200165
172	220505	2	S	2015/6/22	210819	210690	200460	200948	200555	200165
173	220506	2	S	2015/6/22	210819	210690	200460	200948	200555	200165
174	220508	2	S	2015/6/22	210819	210690	200460	200948	200555	200165
175	220510	1	S	2015/6/22	211030	211059	200350	200084	200460	200948
176	220511	1	S	2015/6/22	211030	211059	200350	200084	200460	200948
177	220513	2	S	2015/6/22	211030	211059	200350	200084	200460	200948
178	220514	2	S	2015/6/22	211030	211059	200350	200084	200460	200948
179	220515	2	S	2015/6/22	211030	211059	200350	200084	200460	200948
180	220516	2	S	2015/6/22	211030	211059	200350	200084	200460	200948
181	220518	2	S	2015/6/22	210841	210019	200555	200983	200350	200494
182	220533	2	S	2015/6/22	210841	210019	200555	200983	200350	200494
183	220534	1	S	2015/6/22	210599	211003	200501	200614	200555	200530
184	220535	1	S	2015/6/22	210599	211003	200501	200614	200555	200530
185	220536	2	S	2015/6/22	210599	211003	200501	200614	200555	200530
186	220540	2	S	2015/6/22	210599	211003	200501	200614	200555	200530
187	220541	2	S	2015/6/22	210599	211003	200501	200614	200555	200530
188	220543	2	S	2015/6/22	210064	210865	200161	200613	200369	200848
189	220544	2	S	2015/6/22	210064	210865	200161	200613	200369	200848
190	220545	1	S	2015/6/22	210599	210309	200501	200614	200585	200344
191	220546	2	S	2015/6/22	210599	210309	200501	200614	200585	200344
192	220548	2	S	2015/6/22	210599	210309	200501	200614	200585	200344
193	220550	1	S	2015/6/22	211095	211066	200196	200404	200161	200915

194	220551	1	S	2015/6/22	211095	211066	200196	200404	200161	200915
195	220554	2	S	2015/6/22	210064	210664	200161	200613	200350	200358
196	220561	2	S	2015/6/22	211095	210163	200196	200404	200460	200846
197	220563	1	S	2015/6/22	210658	210519	200350	200358	200196	200906
198	220564	2	S	2015/6/22	210658	210519	200350	200358	200196	200906
199	220565	2	S	2015/6/22	210658	210519	200350	200358	200196	200906
200	220566	2	S	2015/6/22	210658	210519	200350	200358	200196	200906
201	220580	2	S	2015/6/22	210599	210608	200501	200614	200460	200399
202	220581	1	S	2015/6/22	210199	210663	200161	200556	200350	200358
203	220583	1	S	2015/6/22	210658	210015	200350	200358	200805	200180
204	220586	1	S	2015/6/22	210819	210801	200460	200948	200501	200639
205	220600	1	S	2015/6/22	210006	210904	200369	200890	200611	200586
206	220601	2	S	2015/6/22	210006	210904	200369	200890	200611	200586
207	220603	2	S	2015/6/22	210869	210161	200161	200915	200460	200846
208	220611	2	S	2015/6/22	210064	210865	200161	200613	200369	200848
209	220615	2	S	2015/6/22	211095	211066	200196	200404	200161	200915
210	220618	2	S	2015/6/22	210599	210608	200501	200614	200460	200399
211	220630	1	S	2015/6/22	210199	210663	200161	200556	200350	200358
212	220633	2	S	2015/6/22	210064	210614	200161	200613	200196	200869
213	220634	2	S	2015/6/22	211095	210660	200196	200404	200350	200358
214	220635	2	S	2015/6/22	211095	210660	200196	200404	200350	200358
215	220636	1	S	2015/6/22	211114	210684	200460	200846	200805	200606
216	220638	2	S	2015/6/22	210869	210161	200161	200915	200460	200846
217	220640	2	S	2015/6/22	211030	210563	200350	200084	200196	200404
218	220643	2	S	2015/6/22	211030	210563	200350	200084	200196	200404
219	220644	2	S	2015/6/22	210599	210309	200501	200614	200585	200344
220	220645	1	S	2015/6/22	210819	210083	200460	200948	200369	200905
221	220646	1	S	2015/6/22	210006	210003	200369	200890	200350	200084
222	220648	2	S	2015/6/22	211030	210554	200350	200084	200611	200406
223	220650	2	S	2015/6/22	210658	210015	200350	200358	200805	200180
224	220651	2	S	2015/6/22	210658	210015	200350	200358	200805	200180
225	220653	2	S	2015/6/22	210658	210015	200350	200358	200805	200180
226	220654	2	S	2015/6/22	210658	210015	200350	200358	200805	200180
227	221001	1	C	2015/6/22	210133	211051	200398	200813	200043	200434
228	221003	1	C	2015/6/22	210133	211051	200398	200813	200043	200434
229	221011	2	C	2015/6/22	210531	210505	200981	200454	200436	200316
230	221013	2	C	2015/6/22	210368	210353	200654	200183	200436	200001
231	221014	2	C	2015/6/22	210368	210353	200654	200183	200436	200001
232	221015	2	C	2015/6/22	210368	210353	200654	200183	200436	200001

233	221018	1	C	2015/6/22	210466	210096	200310	200510	200156	200045
234	221031	2	C	2015/6/22	210466	210096	200310	200510	200156	200045
235	221041	1	C	2015/6/22	210504	211118	200436	200316	200398	200533
236	221043	1	C	2015/6/22	210504	211118	200436	200316	200398	200533
237	221044	2	C	2015/6/22	210504	211118	200436	200316	200398	200533
238	221045	2	C	2015/6/22	210504	211118	200436	200316	200398	200533
239	221046	2	C	2015/6/22	210504	211118	200436	200316	200398	200533
240	221056	1	C	2015/6/22	210368	211099	200654	200183	200544	200059
241	221058	1	C	2015/6/22	210368	211099	200654	200183	200544	200059
242	221061	2	C	2015/6/22	210368	211099	200654	200183	200544	200059
243	221063	2	C	2015/6/22	210368	211099	200654	200183	200544	200059
244	221064	1	C	2015/6/22	210088	210644	200544	200458	200466	200103
245	221065	1	C	2015/6/22	210088	210644	200544	200458	200466	200103
246	221066	2	C	2015/6/22	210088	210644	200544	200458	200466	200103
247	221068	2	C	2015/6/22	210088	210644	200544	200458	200466	200103
248	221081	2	C	2015/6/22	210133	211051	200398	200813	200043	200434
249	221083	2	C	2015/6/22	210133	211051	200398	200813	200043	200434
250	221084	1	C	2015/6/22	210504	211108	200436	200316	200435	200056
251	221085	1	C	2015/6/22	210504	211108	200436	200316	200435	200056
252	221088	2	C	2015/6/22	210504	211108	200436	200316	200435	200056
253	221101	2	C	2015/6/22	210018	210138	200466	200108	200398	200813
254	221103	2	C	2015/6/22	210018	210138	200466	200108	200398	200813
255	221104	2	C	2015/6/22	210018	210138	200466	200108	200398	200813
256	221105	1	C	2015/6/22	210665	211043	200418	200333	200156	200139
257	221108	2	C	2015/6/22	210665	211043	200418	200333	200156	200139
258	221110	2	C	2015/6/22	210665	211043	200418	200333	200156	200139
259	221111	2	C	2015/6/22	210665	211043	200418	200333	200156	200139
260	221113	1	C	2015/6/22	210094	210811	200156	200045	200981	200593
261	221114	2	C	2015/6/22	210094	210811	200156	200045	200981	200593
262	221116	2	C	2015/6/22	210094	210811	200156	200045	200981	200593
263	221118	1	C	2015/6/22	210094	210541	200156	200045	200654	200413
264	221130	1	C	2015/6/22	210094	210541	200156	200045	200654	200413
265	221133	2	C	2015/6/22	210094	210541	200156	200045	200654	200413
266	221134	2	C	2015/6/22	210094	210541	200156	200045	200654	200413
267	221135	1	C	2015/6/22	210669	210043	200435	200056	200310	200510
268	221136	1	C	2015/6/22	210669	210043	200435	200056	200310	200510
269	221138	2	C	2015/6/22	210669	210043	200435	200056	200310	200510
270	221140	2	C	2015/6/22	210669	210043	200435	200056	200310	200510
271	221141	2	C	2015/6/22	210669	210043	200435	200056	200310	200510

272	221143	1	C	2015/6/22	210466	210096	200310	200510	200156	200045
273	221145	2	C	2015/6/22	210466	210415	200310	200510	200043	200138
274	221148	2	C	2015/6/22	210466	210415	200310	200510	200043	200138
275	221151	1	C	2015/6/22	210094	210513	200156	200045	200544	200860
276	221153	1	C	2015/6/22	210094	210513	200156	200045	200544	200860
277	221154	2	C	2015/6/22	210094	210513	200156	200045	200544	200860
278	221155	2	C	2015/6/22	210094	210513	200156	200045	200544	200860
279	221158	1	C	2015/6/22	210504	211106	200436	200316	200310	200891
280	221160	1	C	2015/6/22	210504	211106	200436	200316	200310	200891
281	221161	2	C	2015/6/22	210504	211106	200436	200316	200310	200891
282	221164	2	C	2015/6/22	210504	211106	200436	200316	200310	200891
283	221183	1	C	2015/6/22	210133	211054	200398	200813	200310	200145
284	221185	2	C	2015/6/22	210133	211054	200398	200813	200310	200145
285	221186	2	C	2015/6/22	210133	211054	200398	200813	200310	200145
286	221301	1	C	2015/6/22	210018	211130	200466	200108	200418	200338
287	221303	1	C	2015/6/22	210018	211130	200466	200108	200418	200338
288	221304	2	C	2015/6/22	210018	211130	200466	200108	200418	200338
289	221305	2	C	2015/6/22	210018	211130	200466	200108	200418	200338
290	221306	2	C	2015/6/22	210018	211130	200466	200108	200418	200338
291	221308	1	C	2015/6/22	210669	210105	200435	200056	200156	200933
292	221310	1	C	2015/6/22	210669	210105	200435	200056	200156	200933
293	221311	2	C	2015/6/22	210669	210105	200435	200056	200156	200933
294	221313	2	C	2015/6/22	210669	210105	200435	200056	200156	200933
295	221314	2	C	2015/6/22	210669	210105	200435	200056	200156	200933
296	221315	2	C	2015/6/22	210504	210666	200436	200316	200418	200333
297	221316	1	C	2015/6/22	210466	210813	200310	200510	200981	200593
298	221330	1	C	2015/6/22	210133	210548	200398	200813	200814	200331
299	221331	1	C	2015/6/22	210133	210548	200398	200813	200814	200331
300	221333	2	C	2015/6/22	210133	210548	200398	200813	200814	200331
301	221334	2	C	2015/6/22	210133	210548	200398	200813	200814	200331
302	221335	2	C	2015/6/22	210133	210548	200398	200813	200814	200331
303	221336	1	C	2015/6/22	210669	210130	200435	200056	200043	200995
304	221338	2	C	2015/6/22	210669	210130	200435	200056	200043	200995
305	221340	2	C	2015/6/22	210669	210130	200435	200056	200043	200995
306	221341	2	C	2015/6/22	210669	210130	200435	200056	200043	200995
307	221343	1	C	2015/6/22	210018	210351	200466	200108	200436	200001
308	221345	2	C	2015/6/22	210018	210351	200466	200108	200436	200001
309	221346	2	C	2015/6/22	210018	210351	200466	200108	200436	200001
310	221348	1	C	2015/6/22	210531	210505	200981	200454	200436	200316

311	221350	1	C	2015/6/22	210531	210505	200981	200454	200436	200316
312	221354	2	C	2015/6/22	210088	210136	200544	200458	200398	200813
313	221355	1	C	2015/6/22	210088	210136	200544	200458	200398	200813
314	221364	1	C	2015/6/22	210665	210634	200418	200333	200398	200188
315	221365	1	C	2015/6/22	210665	210634	200418	200333	200398	200188
316	221383	1	C	2015/6/22	210485	210383	200043	200109	200654	200183
317	221384	2	C	2015/6/22	210485	210383	200043	200109	200654	200183
318	221385	2	C	2015/6/22	210485	210383	200043	200109	200654	200183
319	221386	2	C	2015/6/22	210485	210383	200043	200109	200654	200183
320	221388	1	C	2015/6/22	210466	210813	200310	200510	200981	200593
321	221400	2	C	2015/6/22	210466	210813	200310	200510	200981	200593
322	221401	1	C	2015/6/22	210531	210090	200981	200454	200544	200458
323	221403	1	C	2015/6/22	210531	210090	200981	200454	200544	200458
324	221404	2	C	2015/6/22	210531	210090	200981	200454	200544	200458
325	221406	2	C	2015/6/22	210531	210090	200981	200454	200544	200458
326	221408	1	C	2015/6/22	210485	210348	200043	200109	200544	200059
327	221410	1	C	2015/6/22	210485	210348	200043	200109	200544	200059
328	221411	2	C	2015/6/22	210485	210348	200043	200109	200544	200059
329	221413	2	C	2015/6/22	210485	210348	200043	200109	200544	200059
330	221414	2	C	2015/6/22	210485	210348	200043	200109	200544	200059
331	221431	1	C	2015/6/22	210018	210390	200466	200108	200435	200088
332	221433	1	C	2015/6/22	210018	210390	200466	200108	200435	200088
333	221434	2	C	2015/6/22	210018	210390	200466	200108	200435	200088
334	221435	2	C	2015/6/22	210018	210390	200466	200108	200435	200088
335	221436	2	C	2015/6/22	210485	210358	200043	200109	200466	200108
336	221438	1	C	2015/6/22	210368	210353	200654	200183	200436	200001
337	221440	1	C	2015/6/22	210368	210353	200654	200183	200436	200001
338	221448	2	C	2015/6/22	210368	210366	200654	200183	200466	200108
339	221450	2	C	2015/6/22	210368	210366	200654	200183	200466	200108
340	221451	1	C	2015/6/22	210368	210185	200654	200183	200435	200839
341	221453	2	C	2015/6/22	210665	210444	200418	200333	200310	200453
342	221454	2	C	2015/6/22	210665	210444	200418	200333	200310	200453
343	221455	2	C	2015/6/22	210665	210444	200418	200333	200310	200453
344	221456	1	C	2015/6/22	210665	210444	200418	200333	200310	200453
345	221458	1	C	2015/6/22	210665	210444	200418	200333	200310	200453
346	221460	1	C	2015/6/22	210088	210113	200544	200458	200418	200914
347	221463	2	C	2015/6/22	210088	210113	200544	200458	200418	200914
348	221465	2	C	2015/6/22	210088	210113	200544	200458	200418	200914
349	221466	2	C	2015/6/22	210665	210536	200418	200333	200981	200454

350	221468	2	C	2015/6/22	210665	210536	200418	200333	200981	200454
351	221480	2	C	2015/6/22	210665	210536	200418	200333	200981	200454
352	221481	1	C	2015/6/22	210669	210158	200435	200056	200418	200106
353	221483	1	C	2015/6/22	210669	210158	200435	200056	200418	200106
354	221485	2	C	2015/6/22	210669	210158	200435	200056	200418	200106
355	221486	2	C	2015/6/22	210669	210158	200435	200056	200418	200106
356	221488	1	C	2015/6/22	210094	210811	200156	200045	200981	200593
357	221510	2	C	2015/6/22	210504	210666	200436	200316	200418	200333
358	221511	2	C	2015/6/22	210088	210583	200544	200458	200435	200193
359	221514	2	C	2015/6/22	210088	210136	200544	200458	200398	200813
360	221515	2	C	2015/6/22	210088	210136	200544	200458	200398	200813
361	221516	2	C	2015/6/22	210466	210813	200310	200510	200981	200593
362	221518	2	C	2015/6/22	210018	210351	200466	200108	200436	200001
363	221530	1	C	2015/6/22	210669	210130	200435	200056	200043	200995
364	221531	1	C	2015/6/22	210485	210358	200043	200109	200466	200108
365	221533	1	C	2015/6/22	210485	210358	200043	200109	200466	200108
366	221534	2	C	2015/6/22	210485	210358	200043	200109	200466	200108
367	221535	2	C	2015/6/22	210485	210358	200043	200109	200466	200108
368	221536	2	C	2015/6/22	210665	210536	200418	200333	200981	200454
369	221538	2	C	2015/6/22	210368	210185	200654	200183	200435	200839
370	221540	2	C	2015/6/22	210368	210366	200654	200183	200466	200108
371	221543	2	C	2015/6/22	210018	210390	200466	200108	200435	200088
372	221610	1	C	2015/7/6	210368	210366	200654	200183	200466	200108
373	221611	2	C	2015/7/7	210504	211108	200436	200316	200435	200056
374	221613	1	C	2015/7/8	210665	210536	200418	200333	200981	200454
375	221614	1	C	2015/7/9	210665	210536	200418	200333	200981	200454
376	221615	2	C	2015/7/10	210665	210536	200418	200333	200981	200454
377	221616	2	C	2015/7/11	210665	210536	200418	200333	200981	200454
378	221619	1	C	2015/7/12	210466	210096	200310	200510	200156	200045
379	221630	2	C	2015/7/13	210466	210096	200310	200510	200156	200045
380	221631	1	C	2015/7/14	210018	210138	200466	200108	200398	200813
381	221633	1	C	2015/7/15	210504	210666	200436	200316	200418	200333
382	221634	1	C	2015/7/16	210368	210185	200654	200183	200435	200839
383	221635	2	C	2015/7/17	210088	210583	200544	200458	200435	200193
384	221636	2	C	2015/7/18	210088	210583	200544	200458	200435	200193

附錄二：褐色菜鴨高飼效品系第二代及第四代微衛星遺傳標記分析資料

序號	ID	Line	sex	APT001		APT 004		APT 008		APT 010		APT 012		APT 017		APT 020		APT 025		APT 026		APT 032		APT 033	
1	180104	C	1	178	202	290	294	188	192	204	212	189	189	185	185	177	177	117	117	130	146	259	259	262	266
2	180116	C	1	178	206	298	302	188	192	192	204	185	185	173	181	189	193	105	117	142	146	207	207	266	266
3	180194	C	1	178	202	290	302	192	196	204	212	185	189	185	185	193	193	117	117	142	146	259	259	262	266
4	180351	C	1	178	206	298	314	188	192	212	212	189	189	181	181	189	197	109	117	138	142	207	259	262	266
5	180361	C	1	178	178	298	298	192	196	202	212	185	193	181	185	189	197	117	117	138	142	207	259	266	266
6	180399	C	1	202	202	298	302	188	192	204	212	205	205	181	181	189	197	117	117	130	142	207	255	266	266
7	180405	C	1	178	178	292	298	188	196	204	204	189	189	185	185	193	193	109	117	142	142	207	259	262	262
8	180438	C	1	202	202	290	302	184	192	192	204	185	193	181	181	189	193	117	117	142	146	207	259	266	266
9	180450	C	1	178	202	294	298	192	192	212	212	185	193	177	185	197	201	117	117	142	142	207	259	266	266
10	180591	C	1	202	202	298	302	184	192	192	202	189	193	177	177	193	201	117	117	138	142	207	259	266	266
11	180106	C	2	178	178	290	302	184	192	192	204	185	205	185	185	189	193	105	117	142	146	207	207	262	262
12	180119	C	2	178	178	298	302	184	192	212	212	189	193	177	177	193	201	117	117	142	146	255	259	266	266
13	180180	C	2	202	202	292	296	192	196	192	212	185	189	173	181	189	189	113	117	138	142	207	259	262	262
14	180346	C	2	178	178	298	302	192	196	192	212	189	205	185	185	189	197	117	117	138	142	207	259	262	262
15	180356	C	2	178	178	298	298	184	192	192	212	185	193	181	185	189	201	117	117	142	146	207	207	262	262
16	180410	C	2	178	178	292	298	188	196	204	204	189	189	185	185	193	193	109	117	130	142	207	259	262	262
17	180555	C	2	202	202	298	302	188	192	204	204	189	193	177	181	189	193	117	117	130	138	207	255	262	262
18	180594	C	2	202	202	290	298	184	196	192	212	185	201	185	185	189	193	113	117	130	142	207	259	262	262
19	180599	C	2	178	178	290	302	192	196	204	212	185	193	181	185	193	197	117	117	138	142	207	259	266	266
20	180601	C	2	202	202	290	314	188	196	192	212	185	189	185	185	193	197	109	117	142	146	207	259	266	266
21	180610	C	2	178	178	290	298	192	192	192	204	185	189	181	185	177	189	105	117	138	146	207	259	262	262
22	180651	C	2	178	178	294	302	196	196	212	212	189	193	173	173	189	197	117	117	130	142	207	259	266	266
23	180664	C	2	178	178	290	298	184	188	204	212	185	189	177	181	193	201	117	117	138	138	255	259	266	266
24	180044	S	1	178	206	290	306	192	192	204	204	189	193	181	181	189	189	117	117	130	142	207	259	262	262
25	180093	S	1	178	178	290	298	184	192	192	204	185	189	185	185	189	193	117	117	130	146	259	259	266	266
26	180143	S	1	178	178	294	302	184	192	204	204	189	189	177	185	201	201	117	117	130	142	207	259	266	266

27	180166	S	1	178	202	290	290	192	192	204	204	193	193	177	181	189	189	117	117	142	146	207	259	266	266
28	180193	S	1	202	202	290	290	184	184	204	204	189	193	181	185	189	193	105	117	130	142	207	207	262	266
29	180338	S	1	178	206	290	302	184	184	204	204	193	193	181	181	189	193	105	117	138	142	207	259	262	266
30	180445	S	1	178	202	302	314	188	188	204	212	185	189	177	181	189	193	117	117	142	142	207	259	266	266
31	180449	S	1	178	178	298	298	192	196	202	212	185	193	181	185	189	197	117	117	138	142	207	259	266	266
32	180534	S	1	178	178	290	298	192	192	192	204	185	189	181	185	189	189	117	117	138	142	207	207	266	266
33	180095	S	2	178	178	290	294	184	192	204	204	189	189	181	185	189	193	105	117	130	146	207	259	266	266
34	180098	S	2	178	178	298	314	192	192	204	204	189	193	181	181	189	189	117	117	130	142	207	207	262	262
35	180140	S	2	202	202	298	306	184	192	192	204	193	193	177	185	193	201	117	117	138	138	207	259	266	266
36	180191	S	2	202	202	290	290	184	184	204	204	193	193	181	185	189	201	117	117	130	142	207	259	266	266
37	180343	S	2	202	202	302	302	188	192	192	212	189	193	173	181	185	189	117	117	142	142	207	259	266	266
38	180393	S	2	206	206	290	294	184	192	192	204	189	205	177	185	193	201	113	117	138	142	207	259	262	262
39	180515	S	2	206	206	298	302	184	184	204	212	193	193	177	185	193	201	105	117	138	142	207	207	266	266
40	180688	S	2	178	178	290	290	184	192	204	212	185	193	181	185	189	189	109	117	138	142	207	207	266	266
41	200041	C	2	178	178	298	302	196	196	192	212	185	193	185	185	189	197	109	117	138	146	259	259	266	266
42	200469	C	2	202	202	292	298			192	212	185	185	185	185	193	201	109	117	130	142	207	207	266	266
43	200143	C	2	178	178	302	302	184	184	212	212	189	193	177	181	193	201	117	117	130	142	207	259	262	262
44	200411	C	2	202	202	290	302	192	192	204	212	185	193	181	185	189	193	105	117	138	142	207	207	262	262
45	200531	C	2	202	202	290	290	188	188	212	212	185	189	177	185	189	201	105	117	142	146	259	259	266	266
46	200085	C	1	178	202	302	302	188	192	212	212	185	189	181	181	189	193	105	117	142	146	255	259	262	266
47	200591	C	1	178	202	294	314	192	192	192	212	185	205	181	181	189	189	117	117	138	138	207	259	266	266
48	200800	C	2	178	178	302	302	184	184	204	212	189	189	185	185	189	201	117	117	142	146	259	259	262	262
49	200054	C	1	178	202	302	302	196	196	192	212	193	193	181	185	189	197	105	117	130	142	207	259	262	266
50	200048	C	1	202	202	298	298	184	184	204	212	185	193	185	185	197	201	105	117	138	146	207	207	262	266
51	200951	C	1	178	202	290	302	184	184	192	204	185	185	181	185	189	193	113	117	130	142	207	207	266	266
52	200998	C	2	178	178	302	302	188	188	192	212	185	189	185	185	177	189	117	117	146	146	207	259	262	262
53	200963	C	2	202	202	298	302	188	188	192	212	189	205	185	185	193	197	109	117	138	142	207	259	262	262
54	200885	C	2	202	202	298	298	184	184	204	204	189	193	173	181	189	201	117	117	130	130	207	259	266	266

55	201000	C	2	202	202	302	314			204	212	185	205	181	181	189	197	117	117	142	146	207	259	266	266
56	200933	C	2	202	202	290	298	196	196	204	204	189	189	185	185	193	197	117	117	138	146	207	259	266	266
57	200860	C	2	178	178	298	298	184	184	192	192	185	185	181	185	189	201	117	117	142	146	207	207	262	262
58	200308	C	1	178	202	290	302	192	192	192	204	189	189	185	185	177	177	105	117	130	142	259	259	262	262
59	200534	C	1	202	202	290	290	184	184	204	212	193	205	181	185	193	201	117	117	130	138	207	259	262	266
60	200416	C	1	178	202	302	302	184	188	212	212	205	205	173	173	193	197	109	117	130	146	207	207	262	266
61	200184	C	1	178	178	298	298	184	184	192	212	189	205	177	185	177	201	113	117	130	138	207	259	262	266
62	200560	C	1	202	202	292	302	196	196	204	212	189	193	185	185	193	197	117	117	130	142	207	259	262	262
63	200190	C	1	178	202	302	306	184	184	192	212	189	205	177	185	193	201	105	117	138	138	207	255	266	266
64	200133	C	1	178	202	292	302	196	196	192	192	189	205	185	185	189	193	117	117	138	146	207	259	262	262
65	200110	S	2	202	202	290	314	184	188	192	204	189	193	177	181	189	193	109	117	138	142	207	207	262	262
66	200154	S	2	178	178	290	298	184	184	192	192	185	193	185	185	193	193	117	117	130	146	207	259	262	262
67	200008	S	2	202	202	290	298	184	184	202	204	193	193	185	185	193	193	117	117	138	146	207	259	262	262
68	200035	S	2	202	202	290	298	184	184	204	204	189	193	181	185	189	201	117	117	130	138	259	259	266	266
69	200199	S	2	206	206	290	290	184	184	204	212	193	193	185	185	189	201	105	117	130	142	207	207	266	266
70	200394	S	2	202	202	290	298	184	184	202	204	193	193	185	185	189	201	105	117	142	142	207	259	266	266
71	200344	S	2	178	178	298	314	188	188	204	212	185	193	185	185	193	193	105	117	138	142	207	207	266	266
72	200541	S	2	202	202	290	290	184	184	192	202	193	193	185	185	185	189	113	117	142	146	259	259	262	262
73	200440	S	2	178	178	290	298	184	188	192	202	189	193	177	185	189	193	117	117	130	146	259	259	266	266
74	200556	S	2	202	202	290	294	184	184	204	204	189	193	177	181	189	189	117	117	142	146	207	207	266	266
75	200484	S	2	206	206	290	290	184	184	192	204	189	189	181	185	189	193	117	117	138	142	207	259	262	262
76	200586	S	2	202	202	302	314	184	188	192	204	189	205	177	185	189	193	109	117	130	142	207	259	262	262
77	200031	S	2	202	202	290	306	184	184	204	204	189	193	185	185	193	201	105	117	138	142	207	207	266	266
78	200361	S	1	178	178	290	290	184	184	204	204	189	189	181	185	189	189	109	117	130	146	207	207	262	266
79	200580	S	1	202	202	290	298	184	188	204	204	193	205	177	185	193	193	117	117	142	146	207	207	262	266
80	200600	S	1	178	202	290	290	184	184	204	204	189	193	185	185	185	189	105	113	138	142	207	207	266	266
81	200865	S	1	178	202	290	314	184	184	204	204	193	193	181	181	189	189	113	117	142	142	207	207	262	262
82	200886	S	1	178	202	298	302	196	196	212	212	189	193	173	185	189	197	117	117	138	142	207	259	266	266

83	200805	S	1	202	206	290	290			204	204	189	193	181	181	189	189	117	117	138	142	207	259	262	266
84	200461	S	1	178	178	290	290	184	184	204	204	189	193	177	185	189	189	105	117	138	142	207	259	262	266
85	200369	S	1	178	206	290	290	184	184	192	204	185	189	181	185	189	201	105	109	142	142	207	259	266	266
86	200319	S	1	174	178	286	298	184	184	202	204	189	193	177	185	193	201	109	117	138	142	207	259	262	266
87	200115	S	1	178	202	302	314	188	188	204	212	189	205	181	181	189	189	117	117	142	146	207	259	262	266
88	200011	S	1	202	206	290	298	192	192	202	204	193	193	181	185	189	193	117	117	138	142	259	259	262	266
89	200384	S	1	178	202	298	302	188	188	192	204	189	189	177	185	193	201	117	117	130	146	259	259	266	266
90	200354	S	1	178	202	290	314	184	184	192	204	189	189	177	185	189	193	117	117	130	130	207	207	262	266

附錄三、褐色菜鴨高飼效品系營運計畫書

褐色菜鴨為台灣優良之蛋鴨品種，體型小，產蛋多，蛋重大，且蛋殼堅固，為我國加工蛋 (皮蛋、鹹蛋) 之主要來源鴨種。為因應飼料成本日益高漲，本品系以畜產試驗所宜蘭分所育成之褐色菜鴨畜試一號高產蛋品系為種原,選拔個體殘差飼料採食量性狀，以改進鴨隻飼料利用效率。本品系經命名登記後，將推廣給蛋種鴨場，作為純系育種或與民間褐色菜鴨雜交生產蛋用褐色菜鴨，節省蛋鴨業者飼料成本，增加產業之競爭力。

一、行銷策略：

(一) 產品定位：新育成之褐色菜鴨高飼效品系飼料轉換效率優於現有之褐色菜鴨畜試一號,後續將以褐色菜鴨高飼效品系為主要推廣之品系,以助蛋鴨業者降低飼料成本。

(二) 行銷通路：將透過技術移轉產業團體或種鴨場方式，應用育種分工架構，擴大褐色菜鴨高飼效品系族群，或生產民間所需之蛋用褐色菜鴨。

(三) 售後服務：除提供褐色菜鴨高飼效品系雛鴨外，將輔導技術移轉標的廠商種鴨繁殖及生產技術，以利產業應用。

二、商品化

褐色菜鴨高飼效品系雛鴨搭配相關繁殖及生產技術輔導，為分所商品化服務標的；技轉標的廠商將商品化生產褐色菜鴨高飼效品系、直接作為蛋用褐色菜鴨，或與民間褐色菜鴨雜交生產高飼效蛋用褐色菜鴨。

三、年度營運計畫

(一)第一年：供應褐色菜鴨高飼效品系種母雛3,000隻、種公雛300隻。

(二)第二年：供應褐色菜鴨高飼效品系種母雛3,000隻、種公雛300隻。

(三)第三年：供應褐色菜鴨高飼效品系種母雛3,000隻、種公雛300隻。

已發表之相關文獻

高飼效褐色菜鴨之選育與應用
104農科-2.5.2-畜-L1(3)

一、 中文摘要

本試驗檢定褐色菜鴨高飼效選育鴨群及對照鴨群自34週至37週齡之飼料採食量、產蛋量、體重及體重變化，以分析個體殘差飼料採食量。雛鴨於0-4週網狀高床進行育雛、5-12週網狀高床育成，於12週齡上籠，並於34至37週齡進行為期4週相關性能檢定。檢定結果顯示選拔品系之飼料採食量、總蛋重及體重變化為3,759 ± 495 g、1,609 ± 257 g 及14 ± 59 g，與對照品系之3,892 ± 466 g、1,595 ± 334 g及1 ± 66 g相比較，有略佳之性能表現。選拔鴨隻及對照品系平均殘差飼料採食量分別為-94 ± 445及79 ± 396 g；選拔品系與對照品系在平均飼料轉換效率表現無顯著差異(2.4 ± 0.5與2.6 ± 0.7)。依遺傳率為0.4進行檢定母鴨個體之育種價估算，選拔品系與對照品系之平均殘差飼料採食量育種價估測值在G1-G6分別為-13.4、-25.6、-67.5、-122.6、-128.7 及-197.9 g與7.9、0.1、-14.9、-1.7、-46.9及-58.9g，品系間育種價差異(S-C)則為-21.3、-25.7、-52.6、-120.9、-81.8及-139 g。選拔品系之殘差飼料採食量不論自表型值或是育種價預測值，皆較對照品系為佳。

關鍵詞：褐色菜鴨(Brown Tsaiya)、殘差飼料採食量(Residual feed consumption) 、選拔(Selection)

二、前言

褐色菜鴨為台灣優良之蛋鴨品種，體型小，產蛋多，蛋重大，且蛋殼堅固，不但為我國食蛋之重要來源之一，且為加工蛋(皮蛋、鹹蛋)之主要來源。為因應飼料成本日益高漲，進行提升褐色菜鴨飼料轉換效率之選拔，增加產業之競爭力。

飼料約佔家禽生產的60-70%以上的成本支出，因此增加家禽飼料轉換效率或是降低飼料浪費皆可立即降低飼養成本，增加農民經營效益。有關飼料轉換效率的測定及選拔，係一耗費時間、人力及財力的工作；傳統係藉由針對總蛋重及體重選拔，以獲得與飼料轉換效率相關的反應，而改進飼料轉換效率(Byerly *et al*., 1980; Lutting, 1990)。1990年後，部份學者在選拔工作中導入飼料消耗量的數據或包含飼料消耗量計算所得之參數，據稱能進一步增加飼料轉換效率。而殘差飼料消耗量(residual feed consumption; RFC)的測定工作早在1941年由馬里蘭大學的Dr. Byerly所提出(Byerly, 1941 as cited in Bordas and Minvielle, 1999)。而在預測工具中，體重、產蛋重量、體重變化是最常運用於線性迴歸預測採食量的因

子，而殘差飼料消耗量在諸多的研究顯示係屬於高遺傳變異率(0.4-0.5, Lutting and Urff, 1991)。洛島紅雞隻在經歷長期之飼料殘差雙向選拔後，高飼料採食品系(R+)與低飼料採食品系(R-)，在32-36週齡之殘差飼料消耗量及飼料總消耗量等性狀呈現顯著性差異(Bordas *et al*., 1992, 1996)。殘差飼料消耗量與總蛋重(egg mass)、產蛋數、蛋重及體重均無顯著相關，而與飼料採食量之遺傳相關則高達0.5。為進行飼料轉換效率選拔，首先需測定期間鴨隻的體重增加程度，其次需測定個體(或具親屬關係之群體)之飼料消耗量；多數試驗結果顯示個體檢定方式對於利用育種改善飼料轉換效率是一個較佳的方式(Klemm *et al*., 1994)。確定一個適當的測試期間，是了解性狀表現測試的必要手段之一。法國針對北京鴨進行之飼料轉換效率選拔試驗結果顯示，如果針對飼料轉換效率(feed conversion ratio; FCR)性狀進行選拔，測試期間應涵括肥育期間的最後一週(通常為第7週)；若有足夠的個別測試籠，足以容納較多數目的測試鴨隻，則其測試期間可以適量縮短(Klemm *et al*., 1994)。一般而言，針對飼料轉換效率進行選拔，並未建立所謂對照族群，而多以雙向選拔方式(divergent selection)為之，藉由多代選拔資料，了解經選拔後對於其他性狀之影響。劉 (2012)針對褐色菜鴨進行殘差飼料採食量檢定結果顯示，以四週為檢定期之殘差飼料採食量與檢定全期者間表型相關達0.90，高於為期一、二週檢定期之0.79及0.85。而四週為檢定期之殘差飼料採食量與檢定全期者間之遺傳相關介於0.93至1之間，其遺傳率則介於0.3至0.43之間，顯示產蛋鴨之殘差飼料採食量性狀係可選拔者，故擬於本試驗中進行34-37週齡個體之飼料消耗量、產蛋數等性狀檢定，並依據檢定結果進行種鴨選留繁殖下一代檢定鴨群。

三、試驗材料與方法

(一)試驗動物

以本分所選育之褐色菜鴨品系第5代作為繁殖種鴨，母鴨選留殘差飼料採食量表現較佳者48隻，及全同胞姊妹鴨殘差飼料採食量表現較佳者之公鴨12隻，繁殖第6代雛鴨；另由對照族群第5代採避免近親繁殖方式繁殖第6代對照族群。鴨隻於0-4週間在育雛舍內以紅外線燈泡保溫飼養，4週後移至平飼高床鴨舍育成，並於12週齡逢機選取正常鴨隻上籠檢定。飼料及飲水皆採任飼，0-8週齡餵飼鴨群含粗蛋白質19 %，代謝能2900 kcal/kg之育雛料，8週至初產前餵飼鴨群含粗蛋白質14 %，代謝能2800 kcal/kg之育成料，初產後則餵飼含粗蛋白質18.7 %，代謝能2900 kcal/kg之產蛋料。

(二)檢定項目

檢定期間每 3-4 天供給 700 g 產蛋料，並以面寬 15 公分之特製壓克力製飼料槽餵給，以防隔壁鴨隻盜食。飲水以乳頭式飲水器供應，兩隻共用一個引水乳頭。每 3 或 4 天定時測定鴨隻飼料消耗量，每週秤取鴨重 1 次，每天收集產蛋並秤取蛋重。測定個體在產蛋期 (34-37 週齡)之飼

料採食量、總蛋重、體重變化及平均體重，並據以計算個體之飼料殘差採食量；選拔品系與對照品系之飼料殘差採食量皆依下列方程式計算：

$$RFC = FC - pFC = FC - [a(\overline{BW})^{0.5} + b(\Delta BW) + c(EM) + d]$$

其中 FC 為實測採食量、pFC 為預估採食量、$\overline{BW}$為平均體重、ΔBW為體重變化、EM 為總蛋重。

(三)育種價估算

以動物模式採最佳線性無偏差預測值(BLUP)之計算方法，利用 PEST 程式估算個體飼料殘差採食量性狀之育種價(Groeneveld, 1990)，個體之育種價估算係參照下列數學模式估算：

$$y = Xb + Za + e$$

其中 y：觀測值；

b：為年度固定效應；

a：為逢機遺傳效應；

e：為逢機殘差效應；

X、Z：與觀測值相關因子 b 及 a 之設計矩陣。

四、結果與討論

自 34 週至 37 週齡檢定試驗鴨群之飼料採食量、產蛋量、鴨重及鴨重變化等性狀。選拔品系(S)鴨隻 4 週平均採食量為 3,759 g，每隻每天約採食 134 g，較李等(1991)籠飼組平均消耗量 189 g 減少許多，亦較賴等(2000)單籠單隻的平均消耗量 147 g 為低。鴨隻 4 週平均產蛋量為 1,609 g，每隻每天約生產 67.5 g 的蛋，如果以每顆鴨蛋 68 g 計算，則每週每隻鴨生產 5.9 枚蛋，平均產蛋率約 85%，與賴與康(2000)單籠單隻 40 週齡之平均產蛋率 77.5%相差甚多，亦較本族群 101 年度結果之 77.8%為佳；本年度對照族群同時期產蛋率亦可達 84%。選拔鴨隻 34 至 37 週齡檢定前後體重變化為微幅增重 14 g，而對照族群則在檢定前後體重幾無變化。選拔族群體表型性狀表現較選育基礎族群(G0)為佳(劉，2010)，與同時段對照品系(C) 4 週平均採食量、4 週平均產蛋量及 4 週平均體重變化分別為 3,892 g、1,595 g 及 1 g 相比較，亦有較佳之性能表現(圖 1)。

自 2009 年選育至今共計 6 個世代，各項檢定性狀表型值整理如表 1。34 至 37 週齡為期 4 週檢定期間，選拔品系之平均產蛋量略高於對照品系者(S-C=66、115、-8 、177、91 及 14 g)。同期間平均飼料採食量，則以選拔品系略少於對照品系者(S-C=-50、32、-121、-229、65 及-133 g)；然以$RFC = FC - pFC = FC - [a(\overline{BW})^{0.5} + b(\Delta BW) + c(EM) + d]$估算檢定鴨隻飼料採食量，6 代選拔品系平均估算採食量亦低於對照品系(S-C=95、200、104、-83、178 及 40 g)，然各世代選拔鴨隻平均殘差飼料採食量則極顯著少於對照品系鴨隻者(S-C=-145、-168、-225、-122、-112 及-173 g)。以平均飼料轉換效率(飼料換蛋率)而言，各世代選拔品系鴨隻表現亦略優於對照品系者(S-C=-0.6、-0.33、 -0.09、0、-0.1 及-0.2)。

諸多研究顯示家禽之飼料轉換效率皆屬中等遺傳率者(Bordas and Mérat, 1981; Fairfull and Chambers, 1984; Hartmann and Mérat,1986; Bordas *et al*., 1992)，Tixier-boichard et al.(1995)舉出公雞之遺傳率為 0.33，而母雞則為 0.27，而北京鴨之殘差飼料採食量遺傳率為 0.24 (Basso *et al*., 2012)；而褐色菜鴨隻殘差飼料採食量性狀之遺傳率為 0.3 至 0.43，亦屬中等遺傳率者(劉，2010)。依遺傳率為 0.4 進行檢定母鴨個體之育種價估算，選拔品系與對照品系之殘差飼料採食量育種價預測值如表 2 所示，2 品系在 G1-G6 平均估值分別為-15.9、-29.7、-73.1、-125.2 、-128.7 及-197.9 g 與 17.3、15.4、4.1、2.0、-46.9 及-58.9 g，品系間育種價差異(S-C)則為 -37.2、-45.1、-77.2、-127.2、-81.8 及-139 g。殘差飼料採食量與實際採食量呈現高度表型相關(0.61)，而與總蛋重、體重變化及預估採食量等性狀皆為負表型相關(表 3)，相關檢定結果與 Basso *et al*. (2010)相類似，然與飼料轉換效率則為 0.01 之表型相關。估算各檢定性狀間之淨相關(partial correlation)結果如表 4 所示，殘差飼料採食量與實際採食量呈現顯著高度淨相關(0.78)；與總蛋重、檢定期間體重變化亦呈顯著負相關；與飼料轉換效率則呈現-0.01 之淨相關。

選拔與對照 2 個品系第 1 代至第 6 代各世代母鴨平均殘差飼料採食量表型值(P)與育種價預測值(G)趨勢(圖 2)，在選拔品系部分，不論自殘差飼料採食量表型值表現及其育種價預測值曲線，皆一致呈現向下趨勢，亦即選拔品系平均殘差飼料採食量隨著選拔代數增加而逐代減少；然在對照品系部分，表現型值與育種價預測值則呈現不同走勢，平均殘差飼料採食量表型值隨代數微幅增加，唯其育種價預測值亦隨選拔代數而緩步向下。對照品系雖未針對殘差飼料採食量性狀選拔，然亦依其產蛋數多寡作為選留標準，而殘差飼料採食量與 52 週齡產蛋數為正相關(劉，2012)。

2 品系之 34-37 週齡殘差飼料採食量表型值差異在 G6 達 173 g，育種價差異亦達 139 g；換算成每隻鴨每天殘差飼料採食量則可分別減少 6 g 及 5 g 飼料採食量，以產蛋期 10 個月估算，每隻產蛋母鴨約可節省 1.8 kg 或 1.5 kg 飼料成本支出。

五、結論與建議

本試驗以褐色菜鴨 34 至 37 週齡之殘差飼料採食量為選拔指標進行選拔 6 代結果顯示，選拔品系之殘差飼料採食量不論自表型值或是育種價預測值，皆較對照品系為佳，且呈現逐代減少之趨勢。將持續進行選拔，進行新品系登記作業後，推廣民間以降低產業飼料成本支出。

六、參考文獻

李舜榮、潘生才、徐庶財、陳保基。1991。產蛋菜鴨籠飼之探討。畜產研究 24(2):177-184。

劉秀洲。2010。高飼效褐色菜鴨之選育與應用。行政院農業委員會畜產試驗所 99 年度科技計畫研究報告。
劉秀洲。2012。高飼效褐色菜鴨之選育與應用。行政院農業委員會畜產試驗所 101 年度科技計畫研究報告。
賴銘癸、康清亮。2000。褐色菜鴨籠飼面積與族群大小對生產性能之影響。畜產研究 33(3)：281-289。
Basso, B., F. Dubos, P. Morganx, A. Bordas and C. Marie-Etancelin. 2010. Feed efficiency in laying duck: Measurement, genetic variability and correlations with other traits. XIIIth European Poultry Conference. Tours. 23-27 August, I:50-55.
Basso, B., A. Bordas, F. Dubos, P. Morganx and C. Marie-Etancelin. 2012. Feed efficiency in the laying duck: Appropriate measurements and genetic parameters. Poult. Sci. 91:1065-1073.
Bordas, A. and P. Mérat. 1981. Genetic variation and phenotypic correlation of food consumption of laying hens corrected for body weight and production. Br. Poult. Sci. 22:25-33.
Bordas, A., M. Tixier-Boichard and P. Mérat. 1992. Direct and correlated responses to divergent selection for residual food intake in Rhode Island Red laying hens. Br. Poult. Sci. 33:741-754.
Bordas, A. and F. Minvielle. 1999. Patterns of growth and feed intake in divergent lines of laying domestic fowl selected for residual feed consumption. Poult. Sci. 78:317-323.
Bordas, A., P. Mérat and F. Minvielle. 1996. Heterosis in egg laying lines under divergent selection for residual feed consumption. Poult. Sci. 75 :20-24.
Byerly, T. C., J. W. Kessler, R. M. Gous and O. P. Thomas. 1980. Feed requirements for egg production. Poult. Sci. 50:2500-2507.
Byerly, T. C. 1941. Feed and other costs of producing market eggs. Page 29 in: Bulletin A1.
Fairfull, R. W. and J. R. Chambers. 1984. Breeding for feed efficiency: poultry. Can. J. Anim. Sci. 64:513-527.
Groeneveld, E. 1990. PEST User's manual, Department of Animal Sciences, University of Illinois, Urbana, IL. U.S.A.
Hartmann, W. and P. Mérat. 1986. Recent research on the breeding of egg production stock. 7th European Poultry Conference, Paris. 24-28 August, I:49-59.
Klemm, R., K. Reiter, and H. Pingel. 1994. Results and effects of direct selection for feed efficiency in the domestic ducks: 2. Report: Methodical and aspects. Arch. Geflglek. 56(5):216-221.
Lutting, P. 1990. Genetic variation of energy partitioning in laying hens: causes of

variation in residual feed consumption. World's Poult. Sci. J. 46:133-152.
Lutting, P. and E. M. Urff. 1991. Residual feed consumption in laying hens. 2. Genetic variation and correlations. Poult. Sci. 70:1663-1672.
Tixier-Boichard, M., D. Boichard, E. Groeneveld, and A. Bordas. 1995. Restrict maximum likelihood estimates of genetic parameters of adult male and female Rhode Island Red chickens divergently selected for residual feed consumption. Poult. Sci. 74:1245-1252.

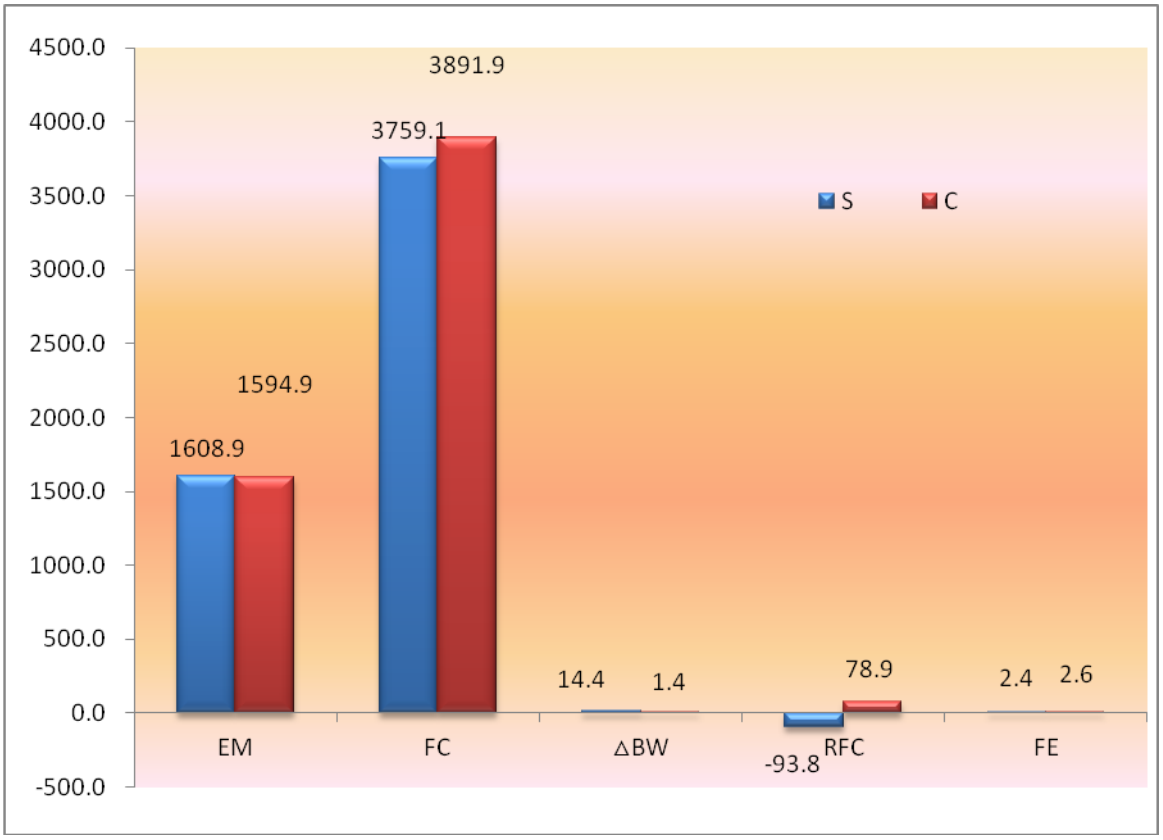


圖 1. 殘差飼料採食量選拔品系與對照品系在 34-37 週齡檢定性狀表現。EM:總蛋重、FC:實測採食量、 ΔBW:體重變化、RFC:殘差飼料採食量、FE:飼料轉換效率

表 1. 第 3 代至第 6 代選拔品系與對照品系於 34-37 週齡檢定性狀之表型值 ± 標準偏差

G.	*G3*		*G4*		*G5*		*G6*	
Traits	S line	C line	S line	C line	S line	C line	S line	C line
ΔBW(g)	-29.4 ± 76.3	-37.7 ± 67.2	-37.3 ± 64.0	-2.2 ± 69.5	-17.1 ± 86.1	0.8 ± 109.6	14.4 ± 58.6	1.4 ± 66.1
EM (g)	1422 ± 299	1430 ± 306	1043 ±328	866 ± 448	1720 ±245	1630 ±257	1609 ± 257	1595 ± 334
FC (g)	3635 ± 471	3756 ± 449	3270 ± 563	3499 ± 747	3811 ± 399	3752 ± 390	3759 ± 495	3892 ± 466
pFC (g)	3749 ± 334	3645 ± 313	3306 ± 437	3389 ± 757	3877 ± 232	3699 ± 212	3853 ± 221	3813 ± 232
RFC (g)	-114.7 ± 302	111 ± 315	-36.2 ± 329	85.9 ± 331	-59.6 ± 314	52.6 ± 309	-93.8 ± 445	78.9 ± 396
FE	2.7 ± 0.78	2.79 ± 0.80	3.2 ± 0.8	3.2 ± 0.9	2.3 ± 0.5	2.4 ± 0.6	2.4 ± 0.5	2.6 ± 0.7

EM:總蛋重、FC:實測採食量、ΔBW:體重變化、pFC:預估採食量、 RFC:殘

差飼料採食量、FE:飼料轉換效率

表 2. 各世代選拔品系與對照品系 34-37 週齡殘差飼料採食量育種價預測值 ± 標準偏差

	S line (g)	*C line (g)*	*S-C (g)*
G0	-0.3 ± 118.4	4.2 ± 189.6	-4.5
G1	-15.9 ± 162.3	17.3 ± 229.7	-37.2
G2	-29.7 ± 114.4	15.4 ± 108.2	-45.1
G3	-73.1 ± 103.0	4.1 ± 103.3	-77.2
G4	-125.2 ± 101.0	2.0 ± 152.9	-127.2
G5	-128.7 ± 121.0	-46.9 ± 118.7	-81.8
G6	-197.9 ± 128.7	-58.9 ± 131.9	-139.0

表 3. 選拔品系各性狀間表型相關

	RFC	△BW	EM	FC	pFC
RFC					
△BW	-0.06				
EM	-0.01	-0.10			
FC	0.61**	0.14	0.56**		
pFC	-0.01	0.23**	0.72**	0.78**	
FE	0.01	0.09	-0.60**	-0.34**	-0.44**

EM:總蛋重、FC:實測採食量、ΔBW:體重變化、pFC:預估採食量、RFC:殘差飼料採食量、FE:飼料轉換效率

**表極顯著相關(P<0.01)

表 4. 選拔品系各性狀間表型淨相關(partial correlation)

	RFC	△BW	EM	FC
△BW	-0.36**			
EM	-0.58**	-0.24**		
FC	0.78**	0.42**	0.70**	
FE	-0.01	0.13*	-0.41**	-0.03

EM:總蛋重、FC:實測採食量、ΔBW:體重變化、RFC:殘差飼料採食量、FE:飼料轉換效率

**表極顯著相關(P<0.01)、*顯著相關(P<0.05)

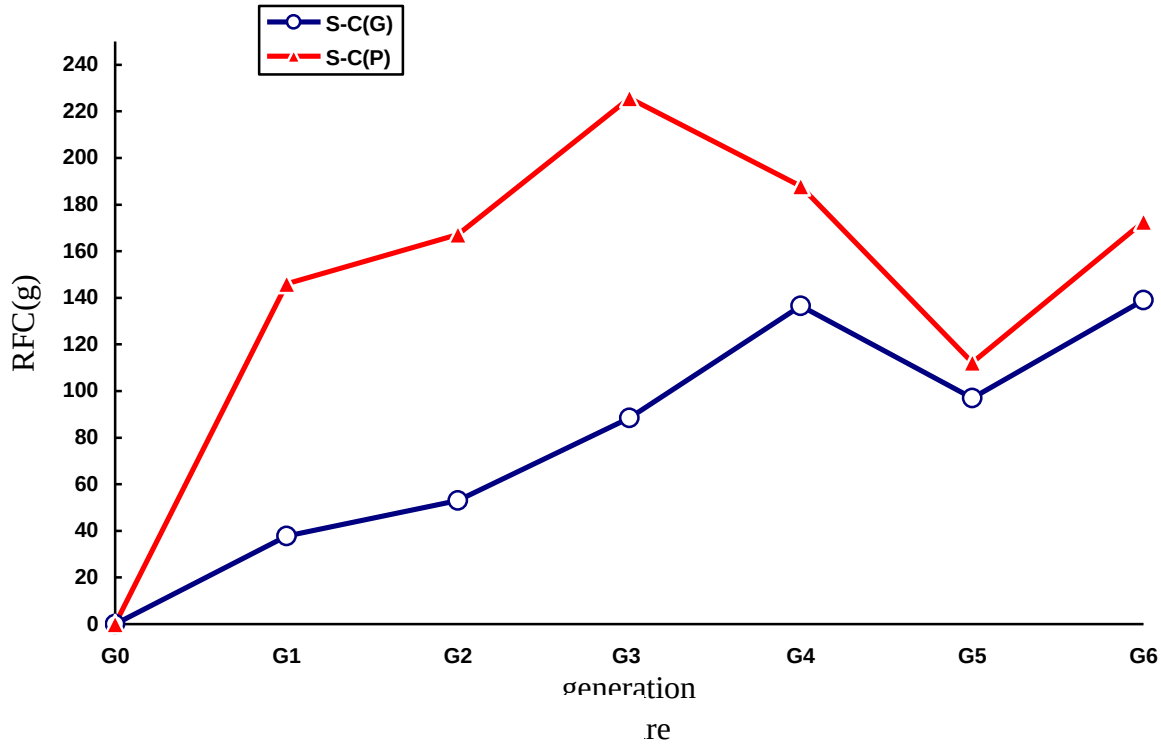


圖 2. 各世代選拔品系之殘差飼料採食量表型值(P)或育種價預測值(G)差異(S-C)。

褐色菜鴨殘差飼料採食量之遺傳參數估算(1)

劉秀洲(2) 杜宗哲(3) Christel Marie-Etancelin(4) 李淵百(3)
黃振芳(2) 陳志峰(3)(5)

收件日期：100 年 6 月 1 日；接受日期：100 年 12 月 22 日

摘要

本試驗旨在估算褐色菜鴨產蛋期間各週齡之殘差飼料採食量，並分析為期不同週數的檢定期與全檢定期間殘差飼料採食量之相關程度，以作為改進褐色菜鴨飼料轉換率之選拔依據。褐色菜鴨檢定族群雛鴨於12週齡上籠，並於22至52週齡進行各週飼料採食量、蛋產量、體重及體重變化檢定。週平均飼料採食量、週平均產蛋量、週平均飼料轉換率、週平均體重、週平均體重變化分別為894 ± 192 g、398 ± 100 g、2.56 ± 1.49 g、1276 ± 133 g及-2 ± 46 g。結果顯示，以四週為檢定期之殘差飼料採食量與檢定全期者間表型相關達0.90，高於為期一、二週檢定期之0.79及0.85。而四週為檢定期之殘差飼料採食量與檢定全期者間之遺傳相關介於0.93至1之間，其遺傳率則介於0.3至0.43之間，顯示產蛋鴨之殘差飼料採食量性狀係屬中等遺傳率者，預期針對殘差飼料採食量性狀之遺傳選拔應屬有效者。後續族群將於34至37週齡進行殘差飼糧採食量相關性狀檢定，數代後再進行相關選拔效率評估。

關鍵詞：褐色菜鴨、飼料殘差採食量、飼料轉換率。

緒言

褐色菜鴨為台灣優良之蛋鴨品種，體型小、產蛋多、蛋重大，且蛋殼堅固，不但為我國食蛋之重要來源之一，亦為加工蛋（皮蛋、鹹蛋）之主要來源。飼料約佔家禽生產的 60%以上的成本支出，所以只要增加飼料轉換率（feed conversion ratio; FCR）、降低飼料浪費皆可立即降低飼養成本，增加農民之經營效益。有關飼料轉換率的測定及選拔，係一耗費時間、人力及財力的工作；傳統模式係藉由針

(1) 行政院農業委員會畜產試驗所研究報告第 1723 號。
(2) 行政院農業委員會畜產試驗所宜蘭分所。
(3) 國立中興大學動物科學系。
(4) National Institute of Agricultural Research, SAGA, Toulouse, France
(5) 通訊作者，E-mail: cfchen@dragon.nchu.edu.tw。

對蛋產量及體重選拔，以獲得與飼料轉換率相關的反應數據，進而改進飼料轉換率（Lutting, 1990）。1990年後，部分學者在選拔工作中導入飼料消耗量的數據或包含飼料消耗量計算所得之參數，據稱能進一步增加選拔效率。而殘差飼料採食量（residual feed consumption; RFC）的測定工作早在1941年由馬里蘭大學的Byerly所提出（Byerly, 1941 as cited in Bordas and Minvielle, 1999)，在預測工具中，體重、產蛋重量、體重變化是最常運用於線性迴歸預測採食量的因子。由多次迴歸方程式估算所得預期採食量，與實際採食量間的差值即為殘差飼料採食量（Koch *et al*., 1963; Bordas and Mérat, 1981），動物的飼料採食量與生產、維持的能量需要及體組成分的改變等參數有密切相關。就產蛋鴨隻而言，這些參數與產蛋數、體重有密切相關，也可能與體重的改變量集體組成分有關（Byerly *et al*., 1980; Chen *et al*., 2008; Marie-Etancelin *et al*., 2008），亦即高殘差飼料採食量（R+）的動物，其飼料轉換率較低殘差飼料採食量者（R-）為差。而雞隻殘差飼料採食量性狀在諸多的研究顯示係屬於高遺傳變異率者（0.4－0.5, Lutting and Urff, 1991）。洛島紅雞隻在經歷長期之飼料殘差雙向選拔後，高飼料量採食品系（R+）與低飼料量採食品系（R-）間，在32－36週齡之殘差飼料採食量及飼料總消耗量等性狀呈現顯著性差異（Bordas *et al*., 1992, 1996）。殘差飼料消耗量與蛋產量（egg mass）、產蛋數、蛋重及體重均無顯著相關，而與飼料採食量之遺傳相關則高達0.5。為進行飼料轉換率選拔，首先需測定期間鴨隻的體重增加程度，其次需測定個體（或具親屬關係之群體）之飼料消耗量；多數試驗結果顯示個體檢定方式對於利用育種改善飼料轉換率是一個較佳的方式（Klemm *et al*., 1994）。確定一個適當的測試期間，是了解性狀表現測試的必要手段之一；在北京鴨的試驗結果顯示，如果針對飼料轉換率性狀進行選拔，測試期間應涵括肥育期間的最後一週（通常為第7週）；若有足夠的個別測試籠，足以容納較多數目的測試鴨隻，則其測試期間可以適量縮短（Klemm *et al*., 1994）一般而言，針對飼料轉換率進行選拔，並未建立所謂對照族群，而多以雙向選拔方式（divergent selection）為之，藉由多代選拔資料，了解經選拔後對於其他生產性狀之影響。至於褐色菜鴨之飼料轉換率選拔部分文獻闕如，故本試驗針對各週齡個體之飼料消耗量、產蛋數等性狀資料進行收集，估算褐色菜鴨產蛋期間各週齡之殘差飼料採食量，並分析不同週數檢定期與全檢定期間之相關程度，以作為改進褐色菜鴨飼料轉換率之選拔依據。

材料與方法

I. 試驗動物

選留本分所繁殖孵化之褐色菜鴨雛鴨，0－3週間在育雛舍內以紅外線燈泡保溫飼養，3週後移至平飼高床鴨舍育成，並於12週齡逢機選取正常鴨隻上籠檢定。飼料及飲水皆採任飼，0－8週齡餵飼鴨群含粗蛋白質19%，代謝能2900 kcal/kg之育雛料，8週至初產前餵飼鴨群含粗蛋白質14%，代謝能2800 kcal/kg之育成料，初產後則餵飼含粗蛋白質18.7%，代謝能2900 kcal/kg之產蛋料，檢定期間每3－4天供給700 g 產蛋料，並以面寬15公分之特製壓克力製飼料槽餵給，以防隔壁鴨隻盜食。飲水以乳頭式飲水器供應，兩隻共用一個引水乳頭。每3或4天定時測定鴨隻飼料消耗量，每週秤取鴨重1次，每天收集產蛋並秤取蛋重。

II. 檢定項目

(i) 檢定期間測定個體在產蛋期（22－52 週齡）之飼料採食量、蛋產量、體重變化及平均體重，並據以計算個體之飼料殘差採食量；飼料殘差採食量依下列方程式計算：$R = FI - \overline{FI} = FI - [a(\overline{BW})^{0.5} + b(\Delta BW) + c(EM) + d]$，其中 FI 為實測採食量、$\overline{FI}$ 為預估採食量、$\overline{\mathrm{BW}}$ 為平均體重、ΔBW 為體重變化、EM 為蛋產量。

(ii) 統計分析

利用 SAS 統計軟體相關模式（CORR procedure）進行檢定期間（22 – 52 週齡）每一週（24, 25, 26, …, 52 週齡, n＝29）、每二週（24 – 25, 25 – 26, 26 – 27, … ,51-52 週齡, n＝28）及每四週（24 – 27, 25 – 28, 26 – 29, …, 49-52 週齡, n= 26）飼料殘差採食量與整個檢定期間（22 – 52 週齡）飼料殘差採食量之表型相關分析。

利用 VCE4 軟體（Groeneveld, 1996）進行變方及遺傳參數分析，方程式如下：

$$y_{ij} = \mu + g_i + e_{ij}$$

y_{ij} ＝ 觀測值

μ ＝ 族群平均值

g_i ＝ 逢機累加基因效果

e_{ij} ＝ 逢機標準機差

結果與討論

試驗鴨群自 22 週至 52 週齡，每週檢定每隻鴨之飼料採食量、產蛋量、鴨重及鴨重變化。試驗結果顯示，整個檢定期間鴨隻平均週採食量為 894 ± 192 g (圖 1) ，檢定後期鴨隻較前期者採食較多之飼料量，鴨隻飼料採食量隨週齡增加及蛋重增加而增加。每隻鴨每天約採食 127 g，較李等（1991）籠飼組平均消耗量 189 g 減少許多，亦較賴等（2000）單籠單隻的平均消耗量 147 g 為低，推測應為檢定方式及飼料槽設計差異影響。

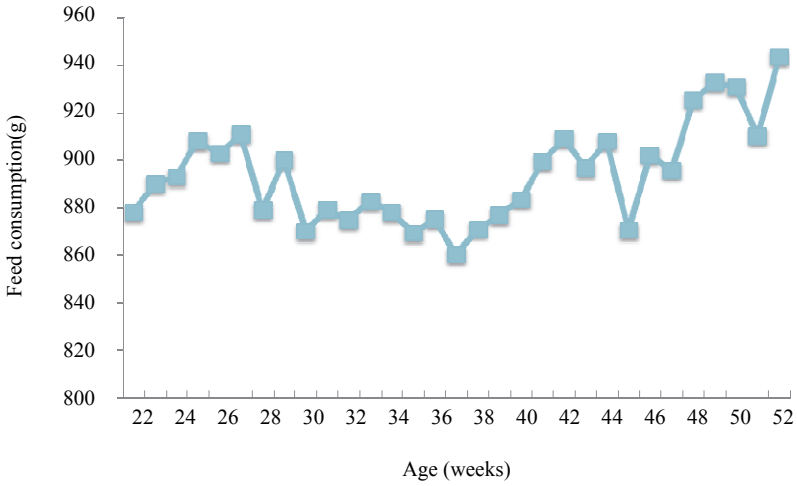


圖 1. 褐色菜鴨單週平均飼料採食量。

Fig. 1. The average weekly feed consumption of Brown Tsaiya ducks.

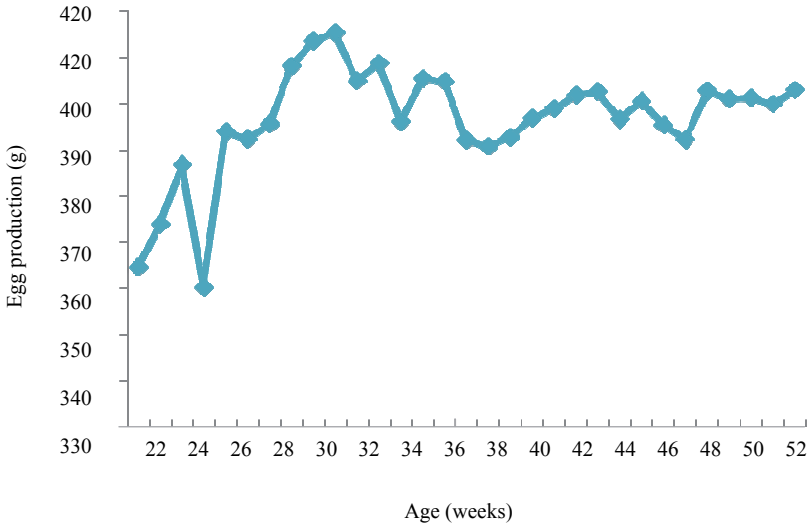


圖 2. 褐色菜鴨單週平均蛋產量。

Fig. 2. The average weekly egg production of Brown Tsaiya ducks.

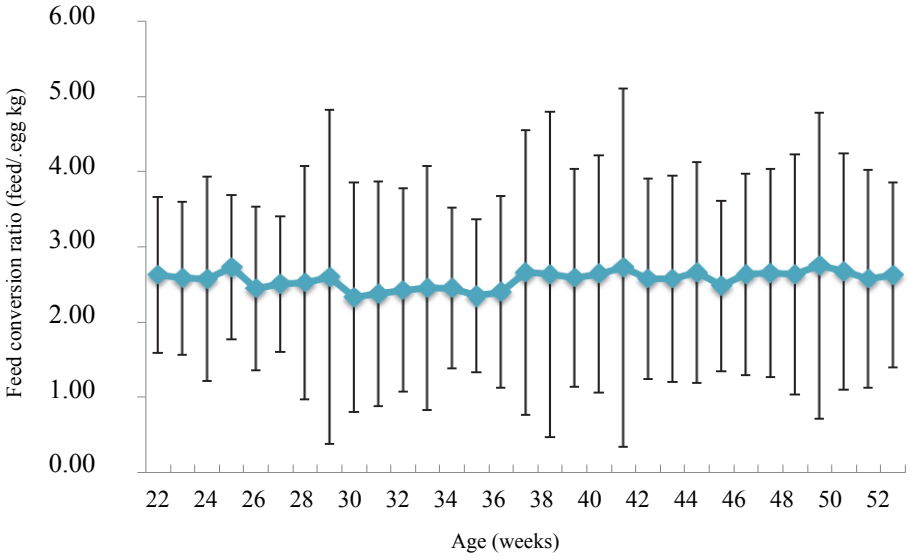


圖 3. 褐色菜鴨單週平均飼料轉換率。

Fig. 3. The average weekly feed conversion ratio of Brown Tsaiya ducks.

檢定期間，鴨隻平均週產蛋量為 398 g (圖 2)，每隻每天約生產 57 g 的蛋，如果以每顆鴨蛋 65 g 計算，則每週每隻鴨生產 6.1 枚蛋，換算成平均產蛋率為 87%，則較賴等（2000）單籠單隻 40 週齡及 72 週齡之平均產蛋率 77.5 及 70.9%為高。產蛋高峰於 31 週齡出現 415 g，隨後下降，直至檢定結束仍維持每週每隻鴨平均 400 g 蛋產量，顯示檢定鴨群並未因頻繁之飼料秤重、鴨隻體重測定等人為干擾，對鴨隻產蛋性能產生嚴重不良影響。

鴨隻週平均體重為 1276 ± 133 g，除產蛋初期數週（23 – 26 週齡）鴨隻平均體重超過 1.3 kg 外，其餘檢定週齡鴨隻平均體重皆不足 1.3 kg，且平均每週體重變化為-2 g。檢定鴨隻 22 – 52 週週平均飼料轉換率為 2.56 ± 1.49 (圖 3)， 其中 29、38、41 及 49 週齡之標準偏差大於其他檢定週齡，分別為 2.60 ± 2.22、2.63 ± 2.16、2.73 ± 2.38 及 2.75 ± 2.03。檢定期間以 30 – 36 週齡之飼料轉換率較檢定前期及後期為低，而最小飼料轉換率標準偏差落於第 27 週，其餘各檢定週齡飼料轉換率標準偏差頗大，暗示褐色菜鴨選拔飼料轉換率的可能性。

根據檢定數據進行遺傳估算，若分別以 1、2、4 週為檢定期，其與全期（22 – 52 週齡）之殘差飼料採食量表型相關如圖 4 所示。為期 1 週之檢定期，其表型相關介於 0.65 至 0.86 間，又以 32、34 及 40 週齡之 0.86 最高；為期 2 週之檢定期，其平均表型相關介於 0.77 至 0.92 間，以 32 – 33 週齡之 0.91 最高；為期 4 週之檢定期，其平均表型相關介於 0.85 至 0.95 間，以 32 – 35 週齡皆高於 0.93。再以 4 週為檢定期，進行與全期間之殘差飼料採食量遺傳相關分析，則發現在 26 個樣本中，僅有 50%的樣本收斂，並估算得遺傳相關介於 0.93 至 1，且以 30 – 34 週齡為檢定期與全期之遺傳相關高達 0.99 為最高 (圖 5)。

研究顯示家禽之飼料轉換效率皆屬中等遺傳率者 (Bordas and Mérat, 1981; Fairfull and Chambers, 1984; Hartmann and Mérat,1986; Pauw, 1987; Lutting and Urff, 1991; Bordas *et al*., 1992)，Tixier-Boichard *et al*.（1995）試驗估算公雞之飼料轉換效率遺傳率為 0.33，而母雞者則為 0.27。而本試驗褐色菜鴨各週齡殘差飼料採食量之遺傳率估算結果如圖 6 所示，26 個觀測值中，僅有 16 觀測值收斂並估算得遺傳率，其餘 10 個觀測值未估算得收斂後之遺傳值。4 週檢定期飼料殘差之遺傳率介於 0.30 至 0.43，並以 34 – 37 週齡之 0.43 為最高者，其遺傳率亦有隨鴨隻年齡增加而降低之趨勢。經估算結果，褐色菜鴨飼料殘差性狀屬中等遺傳率者，此數據與 Basso *et al*.（2010）估算北京鴨之殘差飼料採食量遺傳率為 0.31，且與飼料採食量之遺傳相關高達 0.94 結果相類似，顯示產蛋鴨之殘差飼料採食量，如同其它物種，係可選拔者。

針對北京鴨進行之飼料轉換率選拔試驗結果顯示，測試期間應涵括肥育期間的最後一週（通常為第 7 週），若有足夠的個別測試籠，足以容納較多數目的測試鴨隻，則其測試期間可以適當縮短（Klemm *et al*., 1994）。為獲得可信賴的數據，慎選殘差飼料採食量相關性狀測定週齡及測定期間是必要考量因素，測定週齡太早將面臨飼料浪費嚴重問題，同時如果測定期間短於 2 週，對於整個產蛋期間的資料代表性則略嫌不足（Basso *et al*., 2010）。所以，根據資料分析結果，選擇 34 – 37 週齡為期 4 週期間作為檢定期，在褐色菜鴨產蛋期間的殘差飼料採食量選拔上，應為可行之測定時間，除可避免產蛋早期飼料浪費之問題產生外，亦應可獲致較佳的選拔效果。

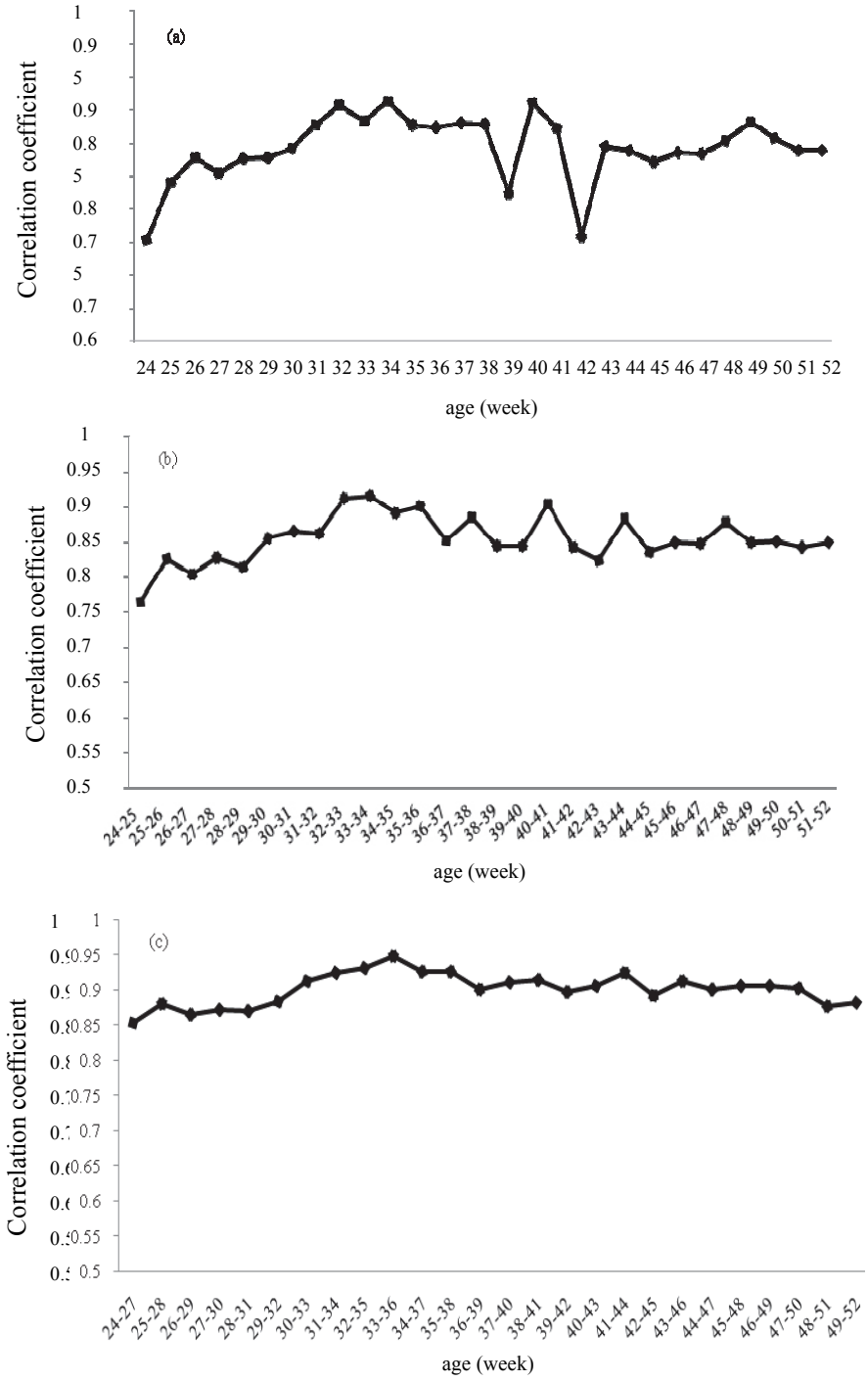


圖 4. 為期 1、2、4 週檢定期與全檢定期間殘差飼料採食量之表型相關。

Fig. 4. Phenotypic correlation of residual feed consumption between (a) one-week, (b) two-week, (c) four-week recording duration and whole recording duration.

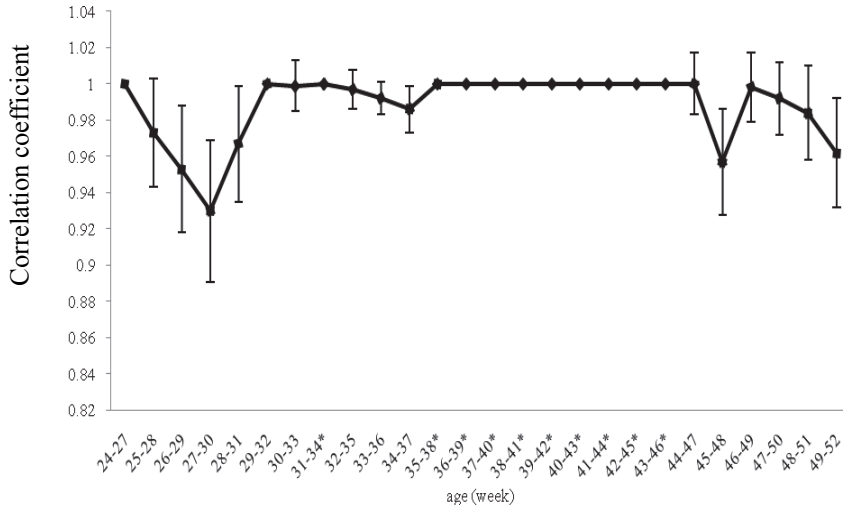


圖 5. 為期 4 週檢定期與全期檢定期間殘差飼料採食量之遺傳相關。

Fig. 5. Genetic correlation of residual feed consumption between four-week recording duration and whole recording duration.

* Optimization did not finish with status 1. Standard errors are therefore not meaningful.

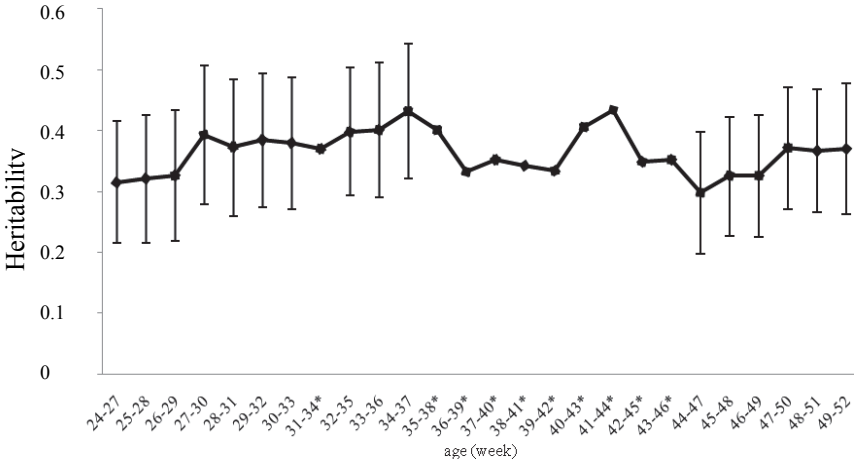


圖 6. 估算每 4 週檢定期殘差飼料採食量之遺傳率。

Fig. 6. Heritability of residual feed consumption for each four-week recording duration.

* Optimization did not finish with status 1. Standard errors are therefore not meaningful.

致謝

本試驗承行政院農業委員會經費支持（98 農科-2.1.3-畜-L1(1)）。試驗期間蒙蕭孟衿小姐、林美葉小姐、林秀齡小姐、馮國銘先生、黃順和先生、柯惠君小姐等同仁協助現場工作，謹申謝忱。

參考文獻

李舜榮、潘生才、徐庶財、陳保基。1991。產蛋菜鴨籠飼之探討。畜產研究 24(2)：177－184。

賴銘癸、康清亮。2000。褐色菜鴨籠飼面積與族群大小對生產性能之影響。畜產研究 33(3)：281－289。

Basso, B., F. Dubos, P. Morganx and C. Marie-Etancelin. 2010. Feed efficiency in laying duck: measurement, genetic variability and correlations with other traits. 13th European Poultry Conference, Tours. I: 621.

Bordas, A. and F. Minvielle. 1999. Patterns of growth and feed intake in divergent lines of laying domestic fowl selected for residual feed consumption. Poult. Sci. 78: 317-323.

Bordas, A., M. Tixier-boichard and P. Mérat. 1992. Direct and correlated responses to divergent selection for residual food intake in Rhode Island Red laying hens. Br. Poult. Sci. 33: 741-754.

Bordas, A. and P. Mérat. 1981. Genetic variation and phenotypic correlation of food consumption of laying hens corrected for body weight and production. Br. Poult. Sci. 22: 25-33.

Bordas, A., P. Mérat and F. Minvielle. 1996. Heterosis in egg laying lines under divergent selection for residual feed consumption. Poult. Sci. 75: 20-24.

Byerly, T. C. 1941. Feed and other costs of producing market eggs. Page 29 in: Bulletin A1. Maryland Agriculture Experimental Station, University of Maryland, College Park, MD.

Byerly, T. C., J. W. Kessler, R. M. Gous and O. P. Thomas. 1980. Feed requirements for egg production. Poult. Sci. 50: 2500-2507.

Chen, C. F., N. Z. Huang, D. Gourichon, Y. P. Lee, M. Tixier-Boichard and A. Bordas. 2008. Effect of introducing the naked neck gene in a line selected for low residual feed consumption on performance in temperature or subtropical environments. Poult. Sci. 87: 1320-1327.

Fairfull, R. W. and J. R. Chambers. 1984. Breeding for feed efficiency: poultry. Can. J. Anim. Sci. 64: 513-527.

Groeneveld E. 1996. VCE4, User's guide and reference, Manual version 1.0, Institute of Animal Husbandry and Animal Behavior, Neustadt, Germany.

Hartmann, W. and P. Mérat. 1986. Recent research on the breeding of egg production stock. 7th European Poultry Conference, Paris. I: 49-59.

Klemm, R., K. Reiter and H. Pingel. 1994. Results and effects of direct selection for feed efficiency in the domestic ducks: 2. Report: Methodical and aspects. Arch. Geflglek. 56(5)：216-221.

Koch, R. M., L. A. Swiger, D. Chambers and K. E. Gregory, 1963. Efficiency of feed use in beef cattle. J. Anim. Sci. 22:：486-494.

Lutting, P. 1990. Genetic variation of energy partitioning in laying hens: causes of variation in residual feed consumption. World's Poult. Sci. 46: 133-152.

Lutting, P. and E. M. Urff. 1991. Residual feed consumption in laying hens. 2. Genetic variation and correlations. Poult. Sci. 70: 1663-1672.

Marie-Etancelin, C., H. Chapuis, J. M. Brun, C. Larzul, M. M. Mialon-Richard and R. Rouvier. 2008.

Genetics and selection of mule ducks in France: a review. World's Poult. Sci. J. 64: 187-208.

Pauw, R. 1987. Analysis of several selection criteria evaluating efficiency in laying hens. Doctoral Thesis. University of Bonn, Fed. Rep. Germany.

SAS. 1999. SAS User's Guide: Statistics, 6.12 ed., SAS Inst., Cary, NC. USA.

Tixier-Boichard, M., D. Boichard, E. Groeneveld and A. Bordas. 1995. Restrict maximum likelihood estimates of genetic parameters of adult male and female Rhode Island Red chickens divergently selected for residual feed consumption. Poult. Sci. 74: 1245-1252.

Genetic parameters of residual feed intake in the Brown Tsaiya duck (1)

Hsiu-Chou Liu(2) Tsung-Che Tu(3) Christel Marie-Etancelin(4)
Yen-Pai Lee (3) Jeng-Fang Huang(2) and Chih-Feng Chen(3) (5)

Received: Jun. 1, 2011; Accepted: Dec. 22, 2011

Abstract

The objective of this study was to estimate weekly residual feed consumption (RFC) of Brown Tsaiya duck during egg laying period, for understanding the correlation between different recording periods and whole recording period, and as reference for further selection. Ducks were individually caged after 12 weeks of age, and data of feed consumption, egg mass, body weight and body weight change were collected every week from 22 to 52 weeks of age. The results indicated that average feed consumption per week, average egg mass per week, average weekly feed conversion ratio, average body weight per week and average body weight change were 894 ± 192 g, 398 ± 100 g, 2.56 ± 1.49 g, 1276 ± 133 g and -2 ± 46 g, respectively. Phenotypic correlations of RFC between four-week recording duration and whole recording duration was 0.95, higher than that of one- and two-week recording duration. The genetic correlations of RFC between four-week recording duration and whole recording duration ranged from 0.93 to 1.00. Heritability of RFC ranged from 0.30 to 0.43. This indicated that RFC was moderately to highly heritable and would respond to genetic selection. Data regarding of residual feed consumption related traits will be collected at age from 34 to 37 weeks for continual generations, and then the selection efficiency evaluation will be conducted.

Key Words: Brown Tsaiya duck, Residual feed consumption, Feed conversion ratio.

---

(1) Contribution No. 1723 from Livestock Research Institute, Council of Agriculture, Executive Yuan.
(2) Ilan Branch Institute, COA-LRI, Ilan 26846, Taiwan, R. O. C.
(3) Department of Animal Science, National Chung-Hsing University, Taichung 40227, Taiwan, R. O. C.
(4) National Institute of Agricultural Research, SAGA, Toulouse, France
(5) Corresponding author, E-mail:cfchen@dragon.nchu.edu.tw

Feed efficiency in the laying duck: Appropriate measurements and genetic parameters[1]

B. Basso,*[2] A. Bordas,† F. Dubos,‡ P. Morganx,‡ and C. Marie-Etancelin*

*INRA, UR631 Station d'Amélioration Génétique des Animaux, B.P. 52627, 31326 Castanet-Tolosan, France; †INRA—AgroParisTech, UMR1313 Génétique Animale et Biologie Intégrative, 78352 Jouy-en-Josas, France; and ‡INRA, UE89 UEPFG d'Artiguères, 40280 Benquet, France

ABSTRACT The objective of this study was to characterize residual feed intake (RFI) in common laying ducks by a) adjusting position and duration of the measurement period and b) estimating genetic parameters of RFI. The feed intake (FI), BW, and egg mass laid (EML) were recorded for 64 I444 common ducks at the beginning (–35 wk of age) and the middle (41–48 wk of age) of the laying curve. Much feed wastage was observed at the beginning of the laying curve and led to biased FI data. However, when laying was well-established, weekly and fortnightly FI measurements were well correlated phenotypically (Rp from 0.84 to 0.92 and from 0.91 to 0.94, respectively for weekly and fortnightly FI) with the measurements over the whole 2-mo period. Regarding egg mass laid, phenotypic correlations between the one-week measurements and the measurements over the whole 2-mo period were more variable than those for FI, ranging from 0.74 to 0.94, and similar to whatever was the period of measurement. The RFI was investigated in a second experiment based on 384 common female ducks, for which FI, EML, BW, and BW gain were recorded at 39 wk of age. The RFI was determined by multiple regression of FI on metabolic BW and EML. Heritability values of FI and RFI were 0.34 and 0.24, respectively. In addition, if the heritability values obtained for BW (0.65) and BW gain (0.09) were consistent with studies in chickens, the very low EML estimates (0.06) were unexpected. The RFI was strongly genetically linked to FI (Rg = +0.89) but appeared to be independent from BW. Selection based on RFI should therefore reduce the FI of animals without clearly modifying the other components. Moreover, the correlated responses on reproductive traits seem favorable because lower RFI values increase the number of eggs produced per year as well as the hatchability and fertility rates.

Key words: feed efficiency, laying duck, genetic parameter, residual feed intake

2012 Poultry Science 91:1065–1073
http://dx.doi.org/10.3382/ps.2011-02008

INTRODUCTION

In France, the common duck is mainly used in cross-breeding programs with Muscovy drakes to produce mule ducks for fatty liver production. These ducks, the dams of mule ducks, are almost exclusively reared in hatcheries. In contrast, in Asia, the biggest market for the Pekin duck, with about 400 million animals sold each year (Klein-Hessling, 2007), the common duck is bred mainly to produce eggs for human consumption and as a source of meat at an older age. The selection of these waterfowl, therefore, focuses essentially on reproductive traits, such as egg production and female fertility in pure and cross breeds (Marie-Etancelin et al., 2008b). The French breeder market is mainly dominated by 2 companies: Orvia and Grimaud Frères Sélection; the latter is also the leader for Pekin duck selection in Asia by selecting grand-parental stocks in France and multiplying parental stocks in China.

Nevertheless, as for all livestock, feeding the ducks constitutes the primary farming cost, especially because the animals have a long productive career. Feed efficiency is therefore a crucial economic trait for farmers and breeders. In 2009, Orvia marketed a new strain of Muscovy duck improved for its feed efficiency, as determined using the feed conversion ratio (**FCR**). However, due to the possible nonlinearity of the component traits of the FCR, the genetic improvement is complex and could lead to undesirable changes, such as the modification of BW and egg production in laying hens (Bordas et al., 1992). We therefore decided to investigate the concept of residual feed intake (**RFI**) in ducks. By definition, RFI is a fraction of total food intake phenotypically unlinked to maintenance and production requirements. But the genetic correlations between

Received November 7, 2011.
Accepted January 10, 2012.
[1]This study was supported in part by the ANR-Genanimal, AGENAVI/CIFOG, and the Inter-regional Midi-Pyrénées Aquitaine Council.
[2]Corresponding author: Benjamin.Basso@toulouse.inra.fr

RFI and its components differ by species and animals experiments. The use of this selection criterion has fully demonstrated its effectiveness in mammals and in avian breeds, such as laying hens (Bordas et al., 1992), and many studies have been carried out to evaluate its genetic or physiological consequences (Tixier-Boichard et al., 2002). However, as far as we know, the genetic aspects of feed intake (**FI**) in laying ducks have not been studied until now. First, this paper aims to determine the best and cheapest way to evaluate feed efficiency and its component traits during laying, and second, to estimate the genetic determinism of RFI in common laying ducks and whether it is genetically linked to other economic traits.

MATERIALS AND METHODS

Populations, Rearing, and Feeding System

Results were obtained from 2 animal designs, for which hatching, breeding, and measurements were carried out at the INRA waterfowl experimental farm (UEPFG, Benquet, France). Experimental procedures were performed in accordance with French National Guidelines for the care and use of animals for research purposes (Certificate of Authorisation to Experiment on Living Animals no. 7740, Ministry of Agriculture and Fish Products).

In both experiments, the FI of common laying ducks was measured. As FI depends on production, maintenance requirements, and changes in body composition, we recorded feed consumption, egg mass laid, BW, and BW gain (difference between BW at the end of the test period and at the beginning of the test period). The aim of the first design was to define the best period of the laying curve to record feed efficiency information, whereas the second design aimed to estimate the genetic determinism of feed efficiency in laying ducks.

First Design. A flock of 64 common female ducks from the INRA I444 strain (a light Pekin strain selected for fertile period duration) was procreated in one hatching batch. The birds received an ad libitum starting diet containing 16.0% CP and 2,900 kcal of ME/kg from birth to 5 wk old. Then they received a commercial pelleted duck feed containing 16.5% CP and 2,700 kcal of ME/kg. Animals were bred in a collective pen with a natural lighting program until 12 wk of age, and then in individual cages with individual feeders with a lighting program of 16L:8D per day. During 2 periods of 8 wk (**P1** and **P2**, respectively; Figure 1), the weekly FI was recorded for each duck. The FI was defined as the amount of distributed feed at the beginning of the week minus the remaining uneaten feed at the end of the week. Similarly, daily laid eggs were weighed to estimate the daily egg mass laid. Animal ages varied from 28 to 35 wk during P1 (**W28**–**W35**) and from 41 to 48 wk during P2 (**W41**–**W48**).

Second Design. In total, 384 common female ducks were produced over a period of 2 yr and 3 annual batches, from a cross between 2 INRA strains, I444 and I37 (a synthetic heavy common duck strain) belonging to a QTL detection program (Marie-Etancelin et al., 2008a). The experimental set up was similar to the first design for feeding and breeding conditions, except that animals were transferred to individual cages at 10 to 12 wk old, according to hatching batches. Feed intake measurements were performed at an age of 39 wk on average. Testing was performed during one week in the first year and during 2 wk in the second year. During each test period, daily laid eggs were weighed, including abnormal eggs (the weights of broken eggs were assumed to be equal to the average weight of normal

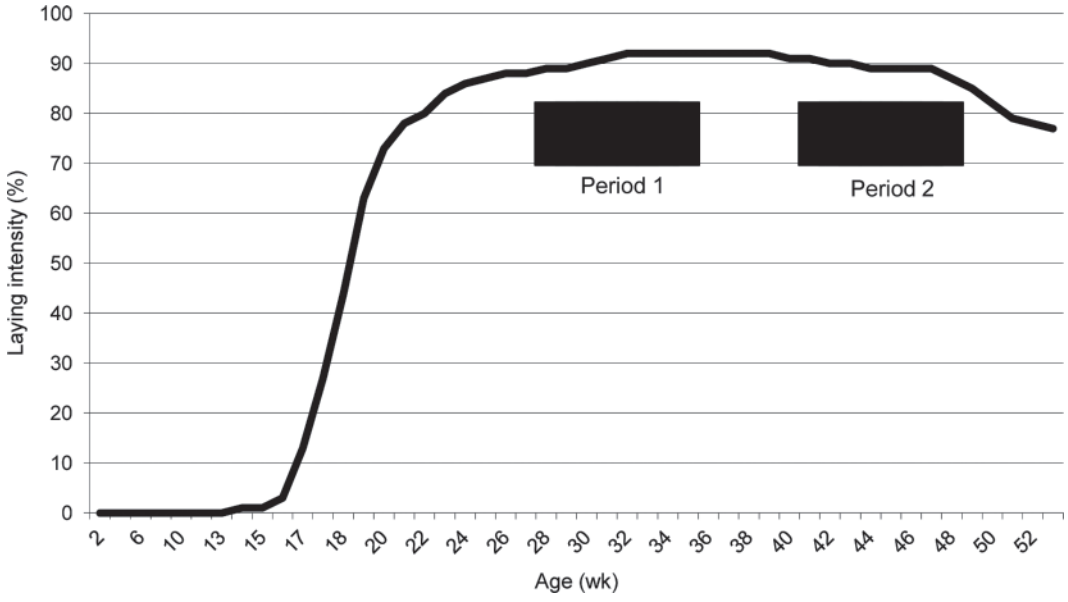


Figure 1. Periods (P1 and P2) of feed intake recording on the theoretical duck laying curve.

eggs laid by the duck during the week). Animals were weighed at the beginning (for both years) and at the end (only for the second year) of the test period.

The reproductive traits of the ducks were recorded up to 52 wk of age. First, the total number of eggs laid per female and the rate of abnormal eggs (broken eggs, soft eggs, porous shell, double yolk) were recorded, and second, the fertility rate (fertilized eggs/incubated eggs) and the hatchability rate (hatched eggs/fertilized eggs) were recorded over a 4-wk period with 2 artificial inseminations per week (crossbred mating with Muscovy drakes, INRA strain I66).

Statistical Analysis

For the first design, correlation coefficients among intake performances were estimated using the CORR procedure of SAS (SAS Institute, 1999).

For the second design, all performances were adjusted to remove the fixed effects of year and hatching batches with the following model:

$$y = \mu + \text{year} + \text{hatch} + e, \quad [1]$$

where y is the measurement of FI, egg mass laid (**EML**), BW, BW gain (**ΔW**), FCR (the ratio between FI and the EML), number of eggs at 1 yr, abnormal eggs rate, fertility rate, and hatchability rate; μ is the intercept; hatch and year are fixed effects; and e is the random residual assumed to be normally distributed with mean 0 and variance σ^2. Phenotypic correlation coefficients among adjusted performances were estimated using the GLM procedure of SAS (SAS Institute, 1999).

To qualify feed efficiency, we estimated the RFI as a linear function of FI and outputs, such as EML, ΔW, and maintenance requirements ($BW^{0.75}$). The power value of BW was chosen equal to 0.75 because it is the most frequently used. Moreover, in the laying hen, the power value may vary in the range of 0.5 to 1 without affecting the efficiency of the prediction (Tixier-Boichard et al., 2002).

So, for a given population, we defined RFI by multiple regression equations, already used by Byerly (1941) to limit food costs in laying hens:

$$\text{FI} = \mu + a \times \text{BW}^{0.75} + b \times \Delta\text{W} + c \times \text{EML} + d \quad [2a]$$

$$\text{or FI} = \mu + a \times \text{BW}^{0.75} + c \times \text{EML} + d \text{ (if } \Delta\text{W not available)}, \quad [2b]$$

where a, b, and c are regression coefficients of FI on metabolic BW ($BW^{0.75}$), ΔW (if available), and EML, respectively. The variable d (or RFI) is the residual of the previous equation: an animal with strong negative RFI values should be more efficient because it consumes less than the regression predicts it should. By definition, RFI is phenotypically independent from the traits for which it has been adjusted.

The effects of year on RFI were tested using the following model:

$$\text{FI} = \mu + \text{year} + a \times \text{BW}^{0.75}(\text{year}) + c \times \text{EML}(\text{year}) + d. \quad [3]$$

Coefficients of the multiple regression and their P-value were estimated using the GLM procedure of SAS (SAS Institute, 1999),

To estimate genetic parameters, the pedigree data included 501 records obtained by tracing back 5 generations of ancestors (on both male and female sides) for 264 laying females measured in the study. The variance components (additive genetic and residual) were estimated using VCE 4.2.5 software (Groeneveld, 1997) in multi-trait analysis, according to an animal model including the year effect. The model was the same for each trait and had the following structure:

$$y = \mu + \text{year} + \alpha + e, \quad [4]$$

where y is the measurement of FI, EML, BW, ΔW, RFI, number of eggs at 1 yr, abnormal eggs rate, fertility rate, and hatchability rate; μ is the intercept; year is the only fixed effect (removed for RFI trait); α is the random additive genetic effect; and e is the random residual effect.

RESULTS

Analysis of FI Measurements at Distinct Periods During the Laying Curve (First Design)

The average FI of the flock was 313 ± 96 g/d during P1 and only 256 ± 58 g/d during P2. These FI values are broadly consistent with the estimates usually obtained with food distributed in linear collective feeders. Individual average FI over 8-wk period distributions were clearly different in mean; $P < 0.0001$ for period fixed effect. Moreover, comparison of mixed models (with or without heterogeneous variance for period random effect) showed a significant difference in dispersion between the 2 periods, with a stronger dispersion of data during P1 (Figure 2). Presence of feed wasters during P1 may explain the upper range of the data. In fact, the mean FI value during period 2 is consistent with the usual feed consumption of a flock of 45 wk old ducks (about 250 g/d), whereas the consumption level of the first period seems to be too high and variable.

To appreciate the representativeness of a short period of FI measurements in laying ducks, we estimated, for each period of 8 wk, phenotypic correlations (Rp) between a) each weekly intake measurement over the period and b) weekly FI measurements and the AFI of the corresponding period. Similar analysis was carried out for the EML trait.

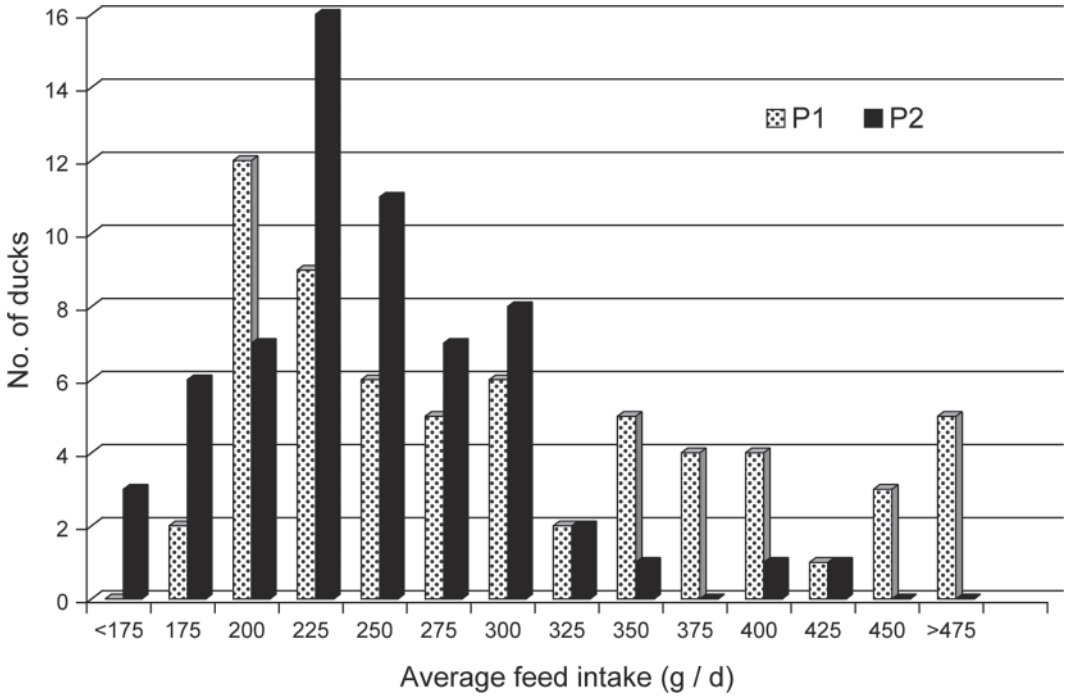


Figure 2. Distribution of individual feed intakes for each measurement period (P1 and P2).

Phenotypic correlations between the weekly intakes over a period were highly variable, ranging from 0.50 to 0.95 (results not shown); the mean values of these correlations were 0.75 ± 0.11 and 0.76 ± 0.05, respectively for P1 and P2 (average of the correlations between each pairs of weekly intakes). The higher SE in P1 confirmed the nonrobustness of FI measurements during this period. Weekly measurements displayed high correlations with the average FI of the corresponding period (Figure 3), ranging from 0.83 to 0.93, whatever the period. The correlations were particularly high for the middle weeks of the period. Moreover, if we averaged 2 consecutive weeks of FI measurements, correlations with the period FI increased even more (Figure 3). For EML, correlations between week measurements and the average of the period were more variable than for FI, ranging from 0.74 to 0.94. As for FI, correlations increased when fortnightly estimations were used. Nevertheless, the results for P2 did not provide a more robust prediction of the EML during the 8 wk than P1. The result was unexpected as at the later ages, most of the egg weights are more consistent, and then correlations between fortnight egg mass and period egg mass were expected to be higher.

Residual FI: Phenotypic and Genetic Analysis (Second Design)

Phenotypic Aspects of RFI and Its Components.

The FI of laying ducks was determined for 192 females each year. As a food-wasting phenomenon specific to some females was reported (pellets observed under the cages, but not weighed), which could lead to a bias in FI estimations, we excluded from the analysis the identified wasteful females. The FI was therefore recorded for 104 and 160 ducks in 2005 and 2006, respectively. For this data set (Table 1), the mean FI value was 265 g/d, whereas the BW and EML were respectively approximately 2.62 kg and 73.3 g/d. Using model 1, a significant ($P < 0.001$) difference appeared for FI and BW between the 2 yr: In 2005, animals were heavier at the beginning of the test period and ate more feed during the test. The FCR during laying was approximately 4.1, with no difference between the 2 yr. The ΔW (only estimated in 2006) was about 8.9 g/d, which is equivalent to only 12% of the average daily egg mass. Regarding the reproductive traits, the total number of eggs laid at the age of 1 yr was 206.4, with no difference observed between the 2 yr, whereas the fertility and hatchability rates (65.9 and 70.1%, respectively) were higher in 2006 compared with 2005. The percentage of abnormal eggs also increased in 2006, with more than 18.0% compared with 10.6% in 2005. Whatever the trait, there was no significant effect of the hatching batch.

The phenotypic correlations between FI and the other components of feed efficiency were computed separately year by year (Table 2). The FI showed correlations with BW (0.24 and 0.29 in 2005 and 2006, respectively) and egg mass (0.18 and 0.34 in 2005 and 2006, respectively). The phenotypic correlation between FI and ΔW was 0.25 in 2006. Correlations were higher during the second year, when measurements were recorded over a 2-wk period instead of only 1 wk during the first year.

To compute the RFI, we applied multiple regressions, first on the 2006 data set only, with (model 2a) or without (model 2b) the ΔW (Table 3). The pheno-

Table 1. Mean, SD, and ANOVA of studied traits (second design)

Item	Trait[1]	N	Mean ± SD	Fixed effect		Least square means	
				Year	Hatch	2005	2006
Test period							
Feed intake (g/d)	FI	264	265 ± 46	***	ns	277	253
Total eggs mass (g/d)	EML	264	73.3 ± 15.6	ns	ns	—	—
BW (g)	BW	264	2,620 ± 289	***	ns	2,733	2,514
BW gain (g/d)	ΔW	158	8.9 ± 4.7	—	ns	—	—
Feed conversion ratio	FCR	264	4.1 ± 3.1	ns	ns	—	—
Out of test period							
Eggs number at 1 yr		264	206.4 ± 39.9	ns	ns	—	—
Abnormal eggs (%)		264	15.1 ± 16.0	***	ns	10.6	18.3
Fertility rate		261	65.9 ± 23.7	***	ns	56.5	70.7
Hatchability rate		256	70.1 ± 20.4	***	ns	62.9	74.4

[1]FI = feed intake; EML = egg mass laid; ΔW = BW gain; and FCR = feed conversion ratio.
***$P < 0.001$.

typic variance accounted for by the models was globally moderate (R^2 around 0.26). The regression coefficients for BW and EML were strongly significant ($P < 0.001$), whereas the coefficient for ΔW was only weakly significant ($P = 0.02$). The variation percentage explained by the model was little improved (model 2b vs. model 2a) with the introduction of the ΔW trait. Indeed, R^2 increased only by 2 points (from 0.25 to 0.27), and the correlation between RFI was 0.98 for both equations (results not shown). When model 3 was applied to the entire data set, RFI was shown to be significantly different for the 2 yr. The part of the phenotypic variance of FI accounted for by the model was also 26%. The coefficients of the intercept and the EML trait (Table 4) were significantly different between years ($P < 0.001$). Regarding the metabolic BW coefficients, the level of significance was nearly reached ($P = 0.053$) for the year effect. Therefore, further analyses were performed using the within-year multiple-regression model (model 3).

Heritabilities and Genetic Correlations. The heritability estimates were approximately 0.34 for FI and 0.24 for RFI (Table 5). Body weight heritability was high (0.65), whereas EML and ΔW heritabilities were close to zero. Body weight was genetically linked to FI (Rg = +0.55), but EML and ΔW were not. Hence, the genetic correlations estimated between FI and the feed efficiency components demonstrate that a selection on food consumption will mainly modify the maintenance requirements (BW).

The strong genetic correlation observed between FI and RFI (Rg = +0.89) showed that the more effective the animal is, the less the animal eats. The RFI seemed to be genetically independent of BW (Rg = +0.12 ± 0.23). Genetic correlations between RFI with EML and ΔW are very strong (−0.65 and 0.57, respectively) but very uncertain, as shown by very high SE (0.60 and 0.41, respectively).

The heritabilities of reproductive traits were consistent with values found in literature (Marie-Etancelin et al., 2008b), ranging from 0.20 for the number of eggs at 1 yr of age to 0.46 for the abnormal eggs rate (Table 6). Genetic correlations between reproductive traits and RFI showed favorable relationships: The decrease of RFI should enhance reproductive traits, by increasing the amount of eggs produced per year, the hatchability rate, and, to a lesser extent, the fertility rate, by reducing the percentage of abnormal eggs.

DISCUSSION

The objective of the first part of this study was to determine the adequate measurement period for collecting accurate RFI data in common laying ducks. At the beginning of laying period, FI measurements seemed to be spurious because of substantial food wastage under the cages. Moreover, the laying status of the female ducks (well-established laying or still sporadic) also affected the energetic requirements of the females and modified the ranking of animals for FI from one week to another. Therefore, to obtain reliable measurements of feed consumption, the beginning of laying period (P1) should be avoided.

When the laying was well established (P2), correlations between weekly measurements and the FI over the whole 8-wk period were always higher than 0.84 (up to 0.92) and were consistent with results found by Bordas and Mérat (1975) in laying hens. They estimated that the correlation between food consumption recorded over a short period (16 d) and a long period (3 mo) was 0.84. When we averaged our FI measurements over 2 consecutive weeks, correlations with the period FI average increased even more and reached 0.94; we hence obtained a better and more stable estimation of consumption during the given period. Thus, the ex-

Table 2. Phenotypic correlations between feed efficiency components

Year	FI with		
	BW	EML	ΔW
2005	0.24	0.18	—
2006	0.29	0.34	0.25

FI = feed intake; EML = egg mass laid; ΔW = BW gain

Table 3. Coefficients of the multiple regression estimated on 2006 data, depending on the model[1]

Model	R square	Value	*a*	*b*	*c*	*d*
2a	0.27	Coefficient	0.65	1.72	1.23	−81.5
		P-value	***	*	***	ns
2b	0.25	Coefficient	0.66	/	1.33	−78.4
		P-value	***		***	ns

[1]Model 2a: FI = $\mu + a \times BW^{0.75} + b \times \Delta W + c \times EML + d$; Model 2b: FI = $\mu + a \times BW^{0.75} + c \times EML + d$ (if ΔW not available).

*$0.005 < P < 0.01$; ***$P < 0.001$.

trapolation of a measurement carried out over a short 2-wk period can provide an accurate estimation of the FI of the duck during a long laying period (at least 8 wk). Nevertheless, the longer the measurement period is, the more reliable the estimation will be. Therefore, the choice of the FI measurement period was crucial to obtain reliable data: The test period must not be placed too early during the laying curve and must be at least 2 wk long.

The objective of the second study was to estimate the genetic parameters of RFI. In 2006, even though measurements were carried out over a 2-wk period, the genetic correlations we observed between FI and BW, EML, or ΔW were low (0.29, 0.34, and 0.25, respectively) compared with those obtained by Bordas and Mérat (1974) in laying hens with a 4-wk measurement period. Their estimates were 0.57 for FI and BW, 0.49 for FI and egg mass produced, and 0.42 for FI and BW change. As a consequence, our predictive model explained only 26% of the phenotypic variance. We could hypothesize that the duration over which measurements were collected affects the precision of the measurements, and that over longer periods, the technical bias that could be introduced (for FI or other traits) is reduced for the estimation of RFI. In laying hens, the part of the phenotypic variance of FI explained by the model is generally higher than that observed here, between 30 and 90% (Luiting and Urff, 1987; Byerly et al., 1980). Nevertheless, the multiple regression coefficients for female Rhode Island Red reported by Bordas

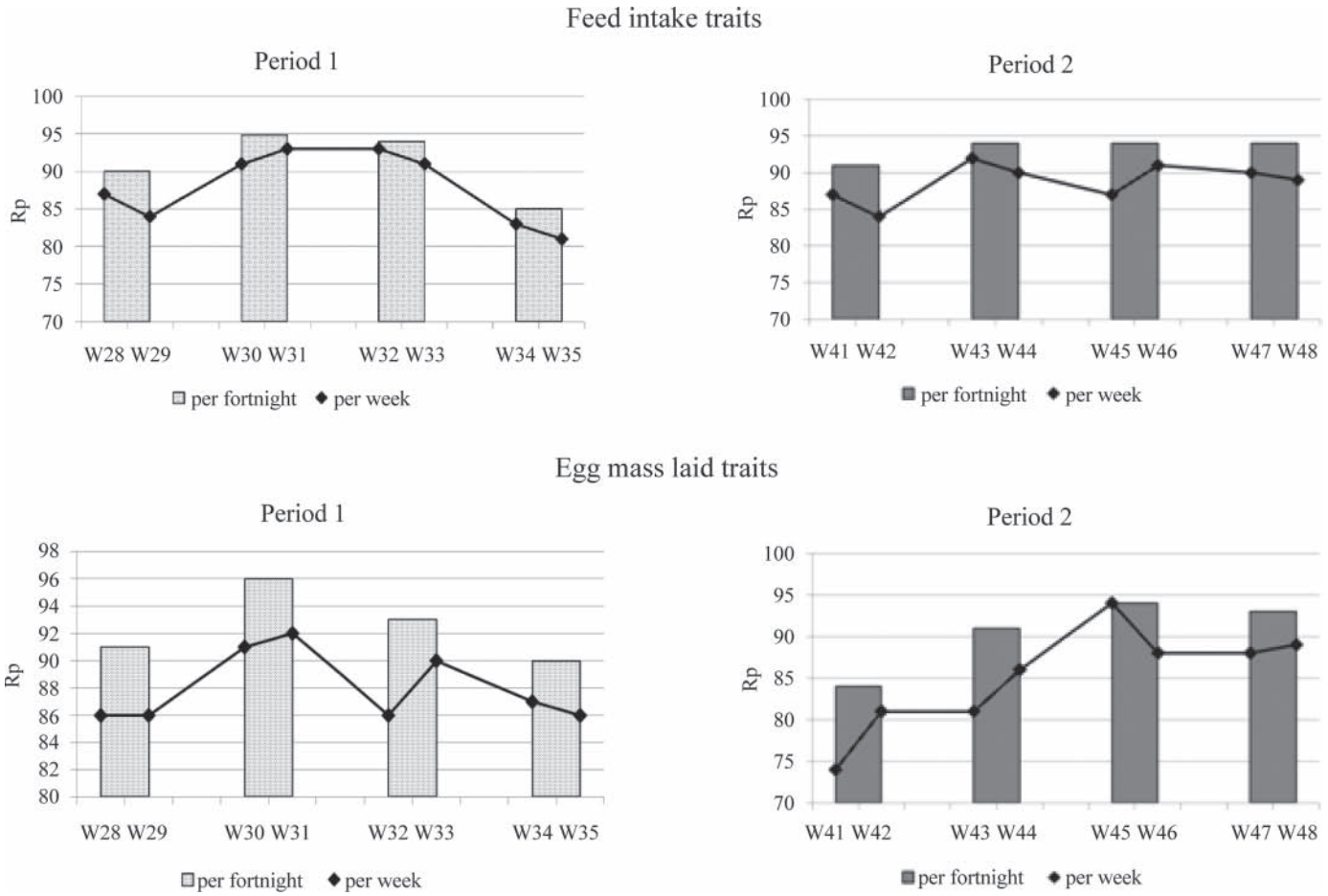


Figure 3. For both periods (P1 and P2) and both traits studied (feed intake and egg mass laid), correlations (Rp) between the average period trait and each average week trait (lines) or each average fortnight trait (histogram).

Table 4. Within-year regression coefficients of feed intake on metabolic BW, egg mass laid (EML), and intercept

Year	Intercept	$BW^{0.75}$	EML
2005	37.17	0.31	0.32
2006	23.89	0.66	1.31
P-value	**	ns (5.3%)	**

** $0.001 < P < 0.005$.

et al. (1992) was close to our coefficients (−83 vs. −81.5 for the intercept; 1.75 vs. 1.72 for ΔW, 0.69 vs. 1.26 for EML; and 3.75 vs. 0.65 for metabolic BW; the coefficients for metabolic BW and the intercept reported by Bordas et al. (1992) were divided by 28 to correct for the duration of measurements). Including the change in BW in the multiple regression improved the accuracy of the model very marginally, we could hypothesize that our r^2 were lower than those in hens because the precision of EML measurements over a 2-wk period was lower in ducks than in hens. As food composition is unchanged in our experiment, another possibility to improve our model's prediction would be to take into account the ambient temperature, which was, according to Combs (1968), the most relevant environmental factor to be considered because high temperatures are known to decrease RFI in pigs, for example (Labroue et al., 1999).

The estimation of the heritability for RFI was of approximately 0.24, which is situated in the large range of published estimations for laying hens (0.05–0.60; Arboleda et al., 1976; Luiting and Urff, 1987; Katle and Kolstad, 1991; Tixier-Boichard et al., 1995). This result confirmed that RFI is, as in other species, selectable in laying ducks. The high heritability values for BW (0.65) and FI (0.34), as well as the low heritability value for ΔW (0.09), were in agreement with the estimations of Luiting and Urff (1991a) and Tixier-Boichard et al. (1995). Indeed, the latter authors found heritability values of 0.56, 0.43, and 0.14 for BW, FI, and ΔW, respectively, for female Rhode Island Red hens. Cheng et al. (1995) reported a heritability value close to 0.50 for the BW of Tsaiya ducks. Our heritability estimation for EML was very low (0.06) and contrasted with that of Tixier-Boichard et al. (1995), who established an EML heritability value of 0.69. Nevertheless, Luiting and Urff (1991b) reported a heritability value for EML in White Leghorn hens lower than 0.30, which decreased over time and occasionally reached zero values.

The strong genetic correlation estimates we show here between FI and RFI is higher than that usually reported for laying hens (from 0.20–0.60 for Luiting and Urff, 1991b; 0.40 ± 0.04 for Tixier-Boichard et al., 1995) or growing turkeys (0.47 for Case et al., 2010). Nevertheless, all results show that the more effective the animal is, the less it eats. Moreover, as for laying hens (Fairfull and Chambers, 1984), estimates of genetic correlations between RFI and the components of feed efficiency in ducks were low (0.12 for BW) or with a high SE for other traits; RFI was therefore only genetically linked clearly to FI. This result is of great interest, as a selection based on RFI would decrease FI without modifying the genetic level of BW. Additional results are needed to control the effect of RFI selection on EML and ΔW.

As reviewed by Tixier-Boichard et al. (2002), correlated responses to selection on RFI were investigated across species and 5 mains groups of traits were identified: heat production, behavior, body composition, physiological indicators, and reproduction. Our favorable genetic correlations between reproductive traits and RFI were consistent with results in hens. In laying hens, Bordas and Mérat (1993) found that, at the 17th generation of lines selected for feed efficiency, the hatching rate was 30% higher than in a low RFI line, with delayed embryo development in the high RFI line, with about 10 h difference in the time of hatching. Morisson et al. (1997) confirmed these results and also showed that high RFI lines display poorer performances at fertilization and lesser production of motile spermatozoa.

Before considering selection on RFI for fatty liver production, we needed to evaluate the correlations between RFI and other traits, such as the fattening of the animals, behavior, and heat adaptation. Concerning heat adaptation, Bordas et al. (1992) reported that selection on RFI criterion modifies mean heat production or dissipation. In laying hens, El-Kazzi et al. (1995) demonstrated that adiposity and lipid contents at various parts of the carcass were higher in the low RFI line. For growing ducks, Guy et al. (2002) showed that the selection on FCR led to a lower growth rate but also to a higher adult BW and less fattening, which was not the goal of such a selection. As in fatty duck production, the most important economic trait is feed efficiency

Table 5. Genetic parameters (heritabilities on the diagonal and genetic correlations above) of feed intake components[1]

	FI	BW	ΔW	EML	RFI
FI	0.34 ± 0.06	0.55 ± 0.17	−0.07 ± 0.47	0.09 ± 0.33	0.89 ± 0.07
BW		0.65 ± 0.11	−0.84 ± 0.43	−0.45 ± 0.38	0.12 ± 0.23
ΔW			0.09 ± 0.13	0.88 ± 0.19	−0.65 ± 0.60
EML				0.06 ± 0.05	0.57 ± 0.41
RFI					0.24 ± 0.11

[1]FI = feed intake; EML = egg mass laid; ΔW = BW gain; RFI = residual feed intake.

Table 6. Genetic parameters of reproductive traits[1]

Item	No. of eggs at 1 yr	Abnormal eggs (%)	Fertility rate	Hatchability rate
Heritability	0.20 ± 0.11	0.46 ± 0.11	0.32 ± 0.10	0.24 ± 0.11
Correlations with RFI	−0.68 ± 0.39	0.57 ± 0.23	−0.32 ± 0.26	−0.67 ± 0.20

[1]RFI = residual feed intake.

during growth, it would also be essential to evaluate the relationship between RFI during reproduction and RFI during growth. In pigs, Gilbert et al. (2010) estimated a genetic correlation of 0.35 to 0.37 between these 2 estimates of RFI. However, in beef cattle, Renand et al. (2010) showed that these 2 traits were independent. In turn, Luiting et al. (1994) indicated that genetic selection for a low RFI lead to the production of less active animals. This aspect was confirmed by Altan et al. (2004) in Japanese quail, where an increase of RFI lead to a longer duration and lesser induction of tonic immobility.

We did not overlook the impact of all environmental factors and genotype-environment interactions. Despite showing that RFI is a good way to improve feed efficiency, Carré et al. (2008) highlighted in broilers the existence of many interactions between poultry genetics and feeding parameters. Thus, in laying hens, Bordas and Minvielle (1997) have shown that the overconsuming line was better adapted to high temperatures with a lesser decrease in egg production and ΔW.

ACKNOWLEDGMENTS

The authors thank all of the INRA-UEPFG (Benquet, France) staff for their technical work on the experiments and L. Drouilhet (INRA-SAGA) for her advice after reviewing.

REFERENCES

Altan, O., I. Oguz, Y. Akbas, and M. Aksit. 2004. Genetic variability of residual feed consumption (RFC) and its relationships with some production traits and fear response in Japanese quail hens (*Coturnic coturnix japonica*). Archiv. Fur. Geflug. 68:223–229.

Arboleda, C. R., D. L. Harris, and A. W. Nordskog. 1976. Efficiency of selection in layer-type chickens by using supplementary information on feed consumption. Theor. Appl. Genet. 48:75–83.

Bordas, A., and P. Mérat. 1974. Variabilité génétique et corrélations phénotypiques caractérisant la consommation alimentaire de poules pondeuses après correction pour le poids corporel et la ponte. Ann. Gén. Sél. Anim. 6:369–379.

Bordas, A., and P. Mérat. 1975. Enregistrement sur une courte période de la consommation d'aliment chez la poule pondeuse pour l'étude génétique de l'efficacité alimentaire. Ann. Gén. Sél. Anim. 7:331–334.

Bordas, A., and P. Mérat. 1993. Durée d'incubation et effet du stockage des œufs sur le taux d'éclosion dans des lignées de poules sélectionnées sur la consommation alimentaire résiduelle. Genet. Sel. Evol. 25:397–402.

Bordas, A., and F. Minvielle. 1997. Réponse à la chaleur de poules pondeuses issues de lignées sélectionnées pour une faible (R−) ou forte (R+) consommation alimentaire résiduelle. Genet. Sel. Evol. 29:279–290.

Bordas, A., M. Tixier-Boichard, and P. Mérat. 1992. Direct and correlated responses to divergent selection for residual food intake in Rhode Island Red laying hens. Br. Poult. Sci. 33:741–754.

Byerly, T.C. 1941. Feeds and other costs of producing market eggs. Univ. Maryland Agric. Exp. Sta., Bull. No. A1.

Byerly, T. C., J. W. Kessler, R. M. Gous, and O. P. Thomas. 1980. Feed requirements for egg production. Poult. Sci. 59:2500–2507.

Carré, B., S. Mignon-Grasteau, and H. Juin. 2008. Breeding for feed efficiency and adaptation to feed in poultry. World's Poult. Sci. J. 64:377–390.

Case, L., S. Miller, and B. Wood. 2010. Genetic parameters of feed efficiency traits in the turkey (*Meleagris gallopavo*). Proc. 9th World Congr. Genet. Appl. Lives. Prod., Leipzig, Germany.

Cheng, Y. S., R. Rouvier, J. P. Poivey, and C. Tai. 1995. Genetic parameters of body weight, egg production and shell quality traits in the Brown Tsaiya laying ducks. Genet. Sel. Evol. 27:459–472.

Combs, G. F. 1968. Amino acid requirements of broilers and laying hens. Proc. Maryland. Nutr. Conf. Univ. Maryland, College Park.

El-Kazzi, M., A. Bordas, and G. Gandemer. 1995. Divergent selection for residual food intake in Rhode Island Red egg-laying lines: Gross carcass composition, carcass adiposity and lipid contents of tissues. Br. Poult. Sci. 36:719–728.

Fairfull, R. W., and J. R. Chambers. 1984. Breeding for feed-efficiency—Poultry. Can. J. Anim. Sci. 64:513–527.

Gilbert, H., Y. Billon, H. Lagant, J. A. Calderon, P. Guillouet, J. P. Bidanel, J. Noblet, P. Sellier, and S. Hermesch. 2010. Correlated responses in sow feed intake, body composition and reproduction after divergent selection for residual feed intake in the growing pig. Proc. 9th World Congr. Genet. Appl. Lives. Prod., Leipzig, Germany.

Groeneveld, E. 1997. VCE 4.2 User's Guide and Reference Manual. Institute of Animal Husbandry and Animal Behaviour, Mariensee, Germany.

Guy, G., C. Larzul, and M. D. Bernadet. 2002. Efficacité alimentaire chez le mulard. 5th Journées de la Recherche sur les Palmipèdes à Foie Gras, Pau, France.

Katle, J., and N. Kolstad. 1991. Selection for efficiency of food utilisation in laying hens: Direct response in residual food consumption and correlated responses in weight gain, egg production and body weight. Br. Poult. Sci. 32:939–953.

Klein-Hessling, H. 2007. Pekin duck breeders require special management. World's Poult. Sci. J. 23:14–18.

Labroue, F., R. Guéblez, M. C. Meunier-Salaün, and P. Sellier. 1999. Feed intake behaviour of group-housed Piétrain and Large White growing pigs. Ann. Zootech. 48:247–261.

Luiting, P., and E. M. Urff. 1987. Residual feed intake: A new source of genetic variation for feed efficiency. Proc. 38th Ann. Meet. Europ. Assoc. Anim. Prod. Lisbon, Portugal.

Luiting, P., and E. M. Urff. 1991a. Residual feed consumption in laying hens. 1. Quantification of phenotypic variation and repeatabilities. Poult. Sci. 70:1655–1662.

Luiting, P., and E. M. Urff. 1991b. Residual feed consumption in laying hens. 2. Genetic variation and correlations. Poult. Sci. 70:1663–1672.

Luiting, P., E. M. Urff, and M. W. A. Verstegen. 1994. Between-animal variation in biological efficiency as related to residual feed consumption. Neth. J. Agric. Sci. 42:59–67.

Marie-Etancelin, C., J. M. André, E. Baéza, B. Basso, D. Bastianelli, M. D. Bernadet, J. M. Brun, S. Davail, F. Dubos, X. Fernandez, K. Gontier, D. Guémené, G. Guy, H. Manse, M. M. Mialon, and C. Larzul. 2008a. Dispositif de detection de QTL chez le canard commun. 8th Journées de la Recherche sur les Palmipèdes à Foie Gras, Arcachon, France.

Marie-Etancelin, C., H. Chapuis, J. M. Brun, C. Larzul, M. M. Mialon-Richard, and R. Rouvier. 2008b. Genetics and selection of mule ducks in France: A review. World's Poult. Sci. J. 64:187–207.

Morisson, M., A. Bordas, J. M. Petit, C. Jayat-Vignoles, R. Julien, and F. Minvielle. 1997. Associated effects of divergent selection for residual feed consumption on reproduction, sperm characteristics, and mitochondria of spermatozoa. Poult. Sci. 76:425–431.

Renand, G., A. Vinet, and D. Krauss. 2010. Genetic relationship between feed intake of growing bulls and adult cows. Proc. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany.

SAS Institute. 1999. SAS User's Guide: Statistics. SAS Inst. Inc., Cary, NC.

Tixier-Boichard, M., D. Boichard, E. Groeneveld, and A. Bordas. 1995. Restricted maximum likelihood estimates of genetic parameters of adult male and female Rhode Island Red chickens divergently selected for residual feed consumption. Poult. Sci. 74:1245–1252.

Tixier-Boichard, M., A. Bordas, G. Renand, and J. P. Bidanel. 2002. Residual food consumption as a tool to unravel genetic components of food intake. Proc. 7th World Congr. Genet. Appl. Lives. Prod. Institut National de la Recherche Agronomique, Montpellier, France.

Patterns of Growth and Feed Intake in Divergent Lines of Laying Domestic Fowl Selected for Residual Feed Consumption

A. BORDAS and F. MINVIELLE[1]

Laboratoire de Génétique Factorielle, Institut National de la Recherche Agronomique, 78352 Jouy en Josas Cedex, France

ABSTRACT At Generations 15 (males) and 18 (females) of a divergent selection experiment on residual feed consumption (RFC) in laying poultry, patterns of growth, feed consumption, and associated traits were monitored between the ages of 4 and 34 wk. This monitoring was done to determine how the well-established RFC and feed intake (FI) divergences between adults of the low intake R^- line and of the high intake R^+ line took place, in relation to the evolution of correlated traits. In males and females, BW and BW gain were higher in the R^- line in the first weeks of test, but patterns of BW were quite similar in both lines afterwards. However, R^- hens remained heavier than R^+ females to the end of the experiment. Line difference for BW was achieved by 28 wk of age, at 2,974 and 2,094 g, and the R^- line was then 135 and 133 g heavier, respectively, for males and females. After a fast initial increase in both sexes, FI diverged quickly around 16 to 18 wk of age (sexual maturity), to attain 13 g/d in males and 28 g/d in females, well before the end of the experiment. At the same time, a divergence was observed for wattle length, which was 21% higher in R^+ females. Residual feed consumption gradually diverged in the two lines, and the difference became significant at 14 wk of age. At the same age, shanks were 8% longer in R^+ hens. Finally, levels of triiodothyronine decreased faster in the R^- line, as FI divergence was increasing. These results indicate that the RFC difference between R^- and R^+, lines obtained in adult birds by divergent selection lines, starts somewhat early in life. In addition, the RFC difference appears to be associated with specific growth periods, around 14 wk and 18 to 20 wk. It is at this time that large differences appear in morphological traits involved in body heat loss. No significant correlations were found between early (6 wk) measures of FI, feed efficiency, and RFC, with adult values (32 to 34 wk) for RFC and FI.

(*Key words*: growth curve, feed consumption pattern, wattle length, shank length, triiodothyronine)

1999 Poultry Science 78:317–323

INTRODUCTION

Since 1976, a divergent selection experiment has been carried out on adult Rhode Island Red fowl for residual feed consumption (RFC), a trait associated with feed intake (FI) and feed efficiency. Residual feed consumption was first measured in poultry by Byerly (1941). After 19 generations, the high feed intake R^+ line and the low feed intake R^- line show considerable divergence for both RFC and total feed consumption measured between 32 and 36 wk of age (Bordas *et al*., 1992, 1996). However, little is known about the early growth of those lines, as only one comparison has been carried out, with cockerels from 4 to 8 wk of age. This measurement was made on Generation 10, at which time no difference of feed efficiency was observed (Tixier-Boichard *et al*., 1988). On the other hand, at the second generation of a similar selection experiment, Katle (1992) reported that feed efficiency of the two divergent lines was significantly different at both adult and early (0 to 5 wk) ages.

The present work had two objectives. The first was to study and compare the patterns of feed intake traits, BW, and morphological measurements up to the age at which RFC is evaluated for selection in the R^- and R^+ lines. The second objective was to detect the correlated changes to selection that were associated with those patterns of growth, and to evaluate the age at which the marked divergence observed in adult birds appeared.

MATERIAL AND METHODS

Experimental Animals and Measurements

The comparison between the two lines has been done separately for males and for females because of limited

Received for publication February 27, 1998.
Accepted for publication October 6, 1998.

[1]To whom correspondence should be addressed: minvielle@dga2.jouy.inra.fr

Abbreviation Key: DIT = diet-induced thermogenesis; FI = feed intake; RFC = residual feed consumption; T_3 = triiodothyronine; T_4 = thyroxine.

facilities and labor, but with the same experimental procedures. Male chicks of the two lines kept for this work were hatched on April 24, 1992. They belonged to Generation 15 of the selection experiment and were sired by the breeders of each line. After a 3-wk adaptation period with 2 animals per cage, 28 cockerels per line were housed in individual cages from 4 to 32 wk of age (with a transfer to larger cages at 10 wk). In the same way, 49 and 50 female chicks, from lines R^- and R^+, respectively, hatched on November 4, 1994 (Generation 18), were kept in individual cages from 4 to 34 wk of age.

The ambient temperature was decreased regularly from 30 C at hatch to 25 C at 4 wk and to 22 C at 6 wk. The lighting regimen was 10 h light/d until 18 wk and 14 h light/d after 18 wk. A starter diet (200 g total protein and 2,800 kcal ME/kg) from 0 to 10 wk, a grower diet (153 g total protein and 2,750 kcal ME/kg) from 10 to 18 wk, and a breeder diet (155 g total protein and 2,650 kcal ME/kg) thereafter, were provided for *ad libitum* consumption.

Individual BW and feed consumption were measured every other week from 4 to 32 wk (males) or 34 wk (females), and numbers of eggs laid by hens and egg weights were recorded daily. Between 4 and 34 wk of age, wattle length was measured eight and four times, respectively, in males and females. During this same time period, shank length was obtained six times in females and twice in males (20 and 24 wk).

Three other individual traits were derived from BW and FI. Body weight gain was the difference of BW over a 2-wk period. Relative BW gain was BW gain divided by BW measured at the start of the 2-wk period. Feed efficiency was the ratio of FI on BW gain over the same 2-wk period.

Hormone Assays

Blood sampling of fed hens was performed on 25 R^- and 27 R^+ individuals, from all sire families, at the ages of 4, 8, 12, and 17 wk. Plasma was obtained and kept frozen at –80 C. It was then transported on dry ice to the Catholic University of Leuven (Belgium), where hormone assays were performed. Plasma level of triiodothyronine (T_3) was determined by RIA, using a ICN[2] antiserum. Plasma thyroxine (T_4) determination was performed with Amersham[3] standard solutions. Intra-and interassay variability for T_3 and T_4 tests were 2.9 and 6.2%, and 3.2 and 3.3%, respectively.

Statistical Analysis

Values of residual feed consumption for each 2-wk period were obtained as the residuals of the linear regression of individual FI on BW, weight gain, and egg mass (in females) during that period, separately for males and females (Bordas *et al*., 1992). For all traits measured repeatedly over time, one-way ANOVA were carried out to test the difference between the two lines at each period. The level of significance was set at $P = 0.01$ in an effort to account for the possible effect of correlated errors on significance, as the same individuals were measured in each successive period.

To describe and compare the shape of growth curves (BW, wattle length) and the evolution of associated feed measurements (FI, RFC), nonlinear regressions were fitted to the data. Three models were used. First, BW was analyzed using the monomolecular model (France *et al*., 1996) with the three-parameter (A, B, and k) equation $BW = A - B \exp(-kt)$, where t was the age in weeks, A was the asymptotic BW in grams; B was the range of BW from initial BW (t = 0) to asymptotic BW; and k was the relative rate of growth. Then, to assess the plateaus (value p and age $t_{plateau}$) reached by BW, FI, and RFC, the following segmented quadratic model (SAS Institute, 1988) was adjusted iteratively to the data:

$$X = at^2 + bt + c \text{ if } t < t_{plateau}, \text{ and}$$
$$X = p \text{ if } t \geq t_{plateau},$$

where t was the age in weeks; X was the measure of the trait (BW, FI, or RFC), a, b, and c were the coefficients of the quadratic equation; and p was the plateau if convergence was reached. For BW, female FI, and RFC, all observations were used to evaluate the plateau but, for male FI, only those from 16 wk on could be included because the plateau was lower than the maximum FI value reached around 16 wk of age (Figure 5). Finally, as the complete (recorded in males only) pattern of wattle growth had a sigmoid shape, it was then described by using the Gompertz model (France *et al*., 1996) with the three-parameter (W_0, μ_0 and D) equation

$$W = W_0 \exp [\mu_0 (1 - \exp(-Dt))/D],$$

where t was the age in weeks; W_0 was the wattle length (millimeters) at t = 0; μ_0 was the specific growth rate at t = 0; and D was a decay parameter. Hormone levels, measured repeatedly on females during the experiment, were analyzed also by linear regression on age.

Correlations between traits, or within a trait and between measures taken at different periods were estimated. All statistical analyses were performed by using the GLM, CORR, and NLIN procedures of the SAS® library of programs (SAS Institute, 1988).

RESULTS

Body Weight

Males of both lines had a very similar growth pattern between 4 and 32 wk (Figure 1A), with no difference in

[2]IMMUCHEN™ ICN Pharmaceuticals, Diagnostics Division, Costa Mesa, CA 92626.

[3]Amersham International, Buckinghamshire, HP7 9LL, U.K.

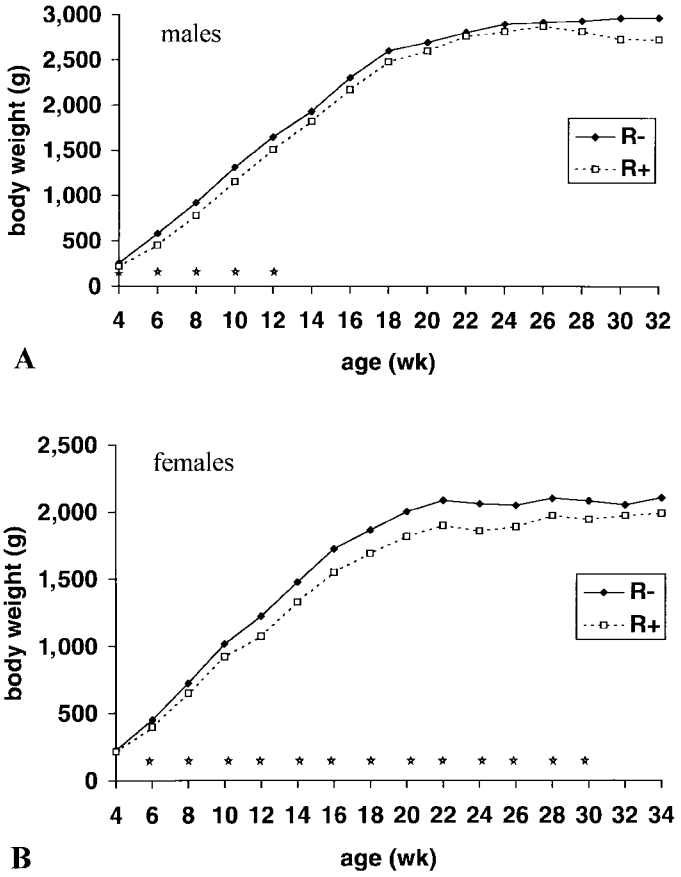


FIGURE 1. Body weight in low intake R– and high intake R+ lines. *$P < 0.01$.

growth curve parameters (Table 1) and close plateau characteristics (Table 2). However, between 4 and 12 wk, cockerels of the R– line were significantly heavier, up to 27% above the mean of the two lines (Figure 1A). On the other hand, females of the high and low intake lines had more different growth patterns (Figure 1B), with a flatter curve for the R+ line, corresponding to a smaller rate k but with similar mature weight A (Table 1) and close BW at the plateau (Table 2). Correspondingly, female R– chicks were heavier during most of the experiment, from 13% at 6 wk down to 7% at 30 wk (Figure 1B).

Relative Weight Gain

Evolution of weight gain relative to BW was similar in both lines (Figure 2), thereby confirming that the growth process was not different overall in the high and low intake lines. However, higher relative weight gain was obtained in R+ males at the beginning of the experiment (Figure 2A).

Feed Intake

In both sexes, FI of the R– line was significantly higher (23 and 10%, respectively, for males and females) in the first experimental period between 4 and 6 wk of age (Figure 3). In males of both lines, FI increased up to 16 to 18 wk of age and then decreased towards intermediate and stable values, but at a 14% higher (relative to the mean of the two lines) level for the R+ line (Figure 3A and Table 2). Feed intake of females increased in the two lines similarly up to about 15 wk, when it reached a plateau for the R– line. Thereafter, FI continued to increase in the R+ line up to a plateau with a 27% higher FI value (Figure 3B and Table 2). Although homologous line differences were observed for FI in both sexes, FI did not show any marked decreasing trend in females. This observation was certainly related to the onset of egg laying around 20 wk of age.

Feed Efficiency

Feed efficiency (feed to gain) was estimated at 6, 8, and 10 wk, and it was significantly lower in the R– lines, both for males (2.4 vs 3.2, with SEM = 0.25) and females (2.9 vs 3.3, with SEM = 0.04) at 6 wk, and at 8 wk (3.2 vs 3.4, with SEM = 0.04) for females only. There were no other significant differences between lines.

Residual Feed Intake

In males of both lines, patterns of change in RFC were linear (Figure 4A and Table 2). The difference between lines became significant at 14 wk, when it was 75 g or

TABLE 1. Growth curve parameters (least squares means) of males and females between 4 and 32 or 34 wk of age from the R– and R+ lines[1] as described by the monomolecular model A–B exp(–kt)[2]

	Male (n = 40)			Female (n = 99)		
	Line			Line		
Parameter	R–	R+	CV	R–	R+	CV
A	3,444	3,368	12.2	2,307	2,220	14.1
B	4,754	4,719	6.4	3,341	3,067	11.1**
k	0.0891	0.0845	24.1	0.1042	0.0930	19.1**

[1]R–, low-feed intake line; R+, high-feed intake line.

[2]A = asymptotic BW (grams). B = range of BW from initial BW (t = 0) to asymptotic BW (grams). k = relative rate of growth. t = age (weeks).

**Significant line difference ($P \leq 0.01$).

TABLE 2. Characterization of the plateaus[1] reached by males and females of the R– and R+ lines,[2] for BW, feed intake (FI), and residual feed consumption (RFC) between 4 and 32 or 34 wk of age

	BW		FI		RFC	
Line	R–	R+	R–	R+	R–	R+
Males (n = 40)						
Value, g	2,974	2,839	1,228	1,411	No plateau	
Age, wk	27.8	28.0	22.6	28.8		
Females (n = 99)						
Value, g	2,094	1,961	1,280	1,676	–248	243
Age, wk	26.3	28.0	14.9	25.4	>34	>34

[1]Value (p) and age ($t_{plateau}$) at plateau were obtained when convergence was reached, by fitting iteratively segmented quadratic model to the data X: $X = at^2 + bt + c$ if $t < t_{plateau}$, and $X = p$ otherwise, where t is the age (wk) and a, b, c, are the coefficients of the quadratic curve.

[2]R–, low-feed intake line; R+, high-feed intake line.

about 4% of mean feed consumption. At the end of the experiment, the difference had reached 306 g, that is 12% of mean FI. In females, evolution of RFC was parallel to that of males (Figure 4B), but a late plateau was predicted both for R– and R+ hens (Table 2). Significant divergence between lines appeared earlier, at 8 wk, and became established at 12 wk (80 g or 7% of mean FI). It reached 407 g (27% of mean FI) at 34 wk. To assess whether the regular increase of the divergence was not simply a scaling effect, for each period, the RFC difference between lines (divergence) was divided by the standard deviation of RFC in that period. Then, the regression of the standardized divergence on time was estimated (Table 3). For males and females, the model was highly significant, and it explained 95 and 92% of the variance, respectively.

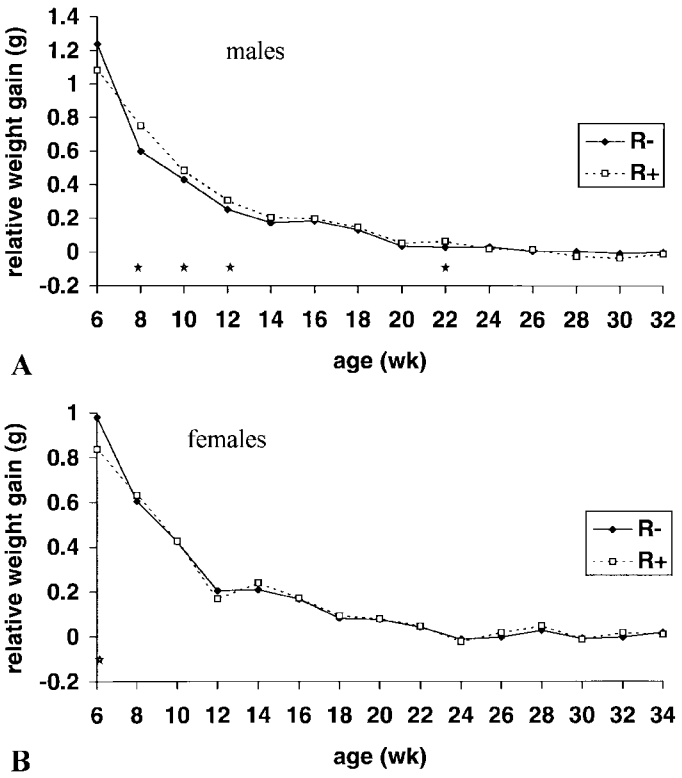


FIGURE 2. Relative body weight gain in low intake R– and high intake R+ lines. *$P < 0.01$.

Egg Production

Egg mass from R+ hens, which started laying somewhat later (data not shown), had caught up with that from R– hens by 26 wk of age, and it was similar in both lines from then on (Table 4).

Shank Length

Shank length of R+ females was significantly higher from the age of 14 wk (Table 4), and the difference remained constant from that age on. This observation was in agreement with the significant 6% difference between lines found in males around 22 wk (data not shown).

Wattle Length

Wattle growth of males was similar in high and low intake lines until 20 wk (Figure 5A); then it diverged significantly, and the difference appeared to increase until the end of the experiment, from 5.3 mm (10% of the mean wattle length) to 12.6 mm, or 20%. For R– and R+ males respectively, estimated values of the parameters of the Gompertz model were 0.254 and 0.936 for W_0, 0.812 and 0.488 for μ_0, and 0.148 and 0.109 for D. The decay parameter D was different ($P \leq 0.05$) in the two lines. Correspondingly, asymptotic values of male wattle length from the Gompertz function were, respectively, 60.3 and 81.6 mm in the R– and R+ lines. Growth of wattles was recorded in females from 14 wk of age only, and wattles are shorter in that sex. Yet, the same overall pattern did occur, as consistently larger wattles were measured in R+ hens as early as on Week 18 (Figure 5B).

Hormone Levels in Hens

Levels of T_3 decreased linearly with age in the two lines and a significant line difference was observed at 17 wk

(Table 4). The regression slopes differed significantly, and were equal to –0.090 (±0.0093) and to –0.063 (±0.0088) ng/mL per wk, respectively, in the R– and R+ lines. On the other hand, level of T_4 did not change linearly as hens grew older (Table 4).

Correlations

Nonsignificant correlations (data not shown) were obtained between early (6 wk) and late (32 or 34 wk) RFC as well as between early measures of feed efficiency and adult RFC.

DISCUSSION

Although male and female data were obtained more than 2 yr apart, they will be discussed together, as they showed similar trends. Data collected on adult birds at the end of this experiment (32 to 34 wk) corresponded well with earlier reports on these same lines (Bordas *et al.*, 1992, 1996), with little, if any, difference in BW, BW gain, and egg mass, but with a large line effect on RFC, FI, and morphological measures. It was reasonable, then, to suppose that the trends up to 32 or 34 wk obtained in this work adequately represented the growth period in the R– and R+ lines, which have been always selected on the basis of feed tests on adult birds only.

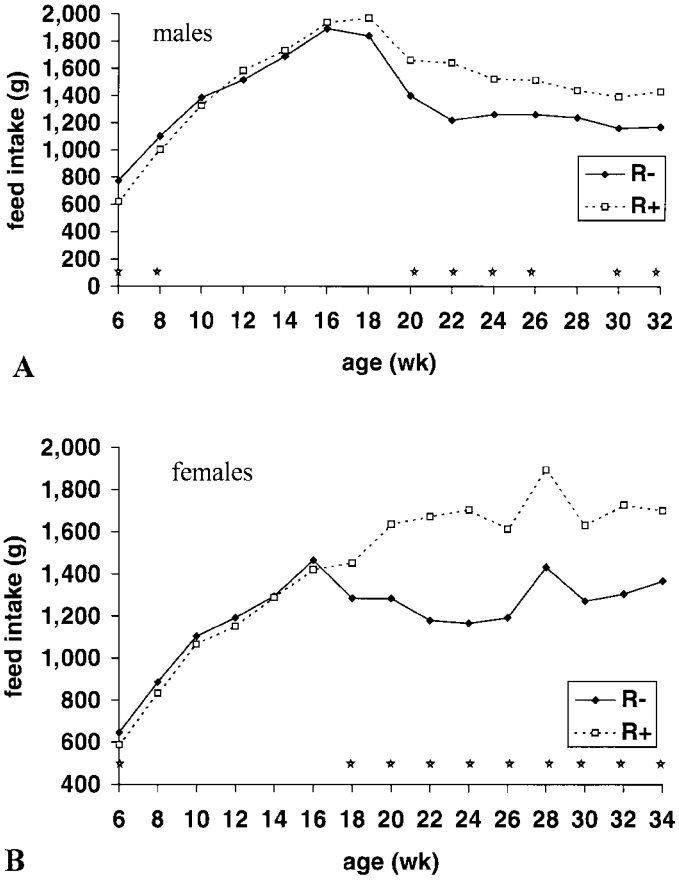


FIGURE 3. Feed consumption in low intake R– and high intake R+ lines. *P < 0.01.

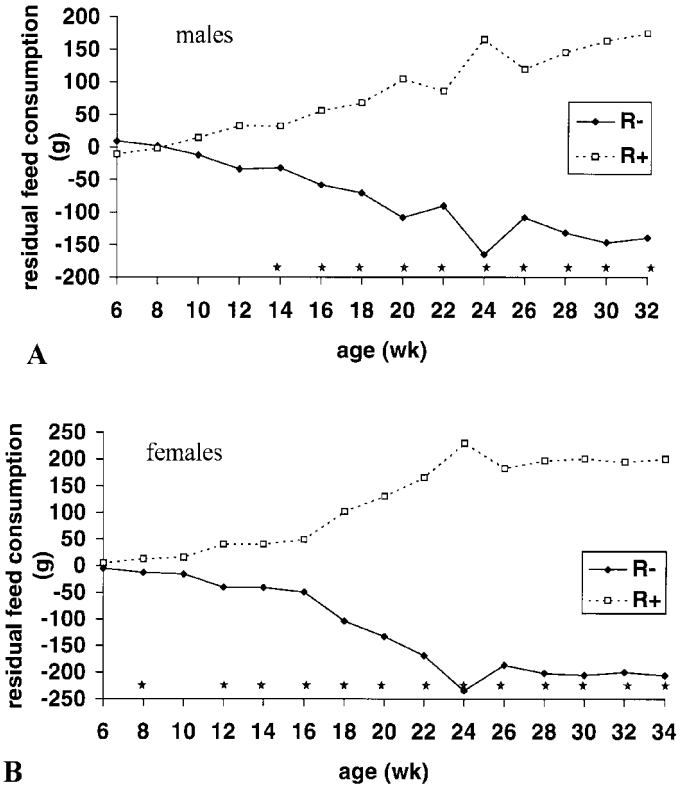


FIGURE 4. Residual feed consumption in low intake R– and high intake R+ lines. *P < 0.01.

Around 6 to 10 wk of age, higher BW and relative BW gain, and relatively lower FI in the R– line produced lower values of feed efficiency (more efficient growth) in both males and females of that line, as observed by Katle (1992) in her own lines. After that period, relative BW gain became similar over time in both lines. Early BW advantage of R– birds was marginal in males but significant in females across most of the experiment.

After a steady and similar increase in the two lines and for both sexes, FI diverged abruptly and permanently around 16 wk of age. In males, when approaching the end of growth, FI decreased, and the line difference had became stable at 13 g/d by the end of the experiment. In females, as egg laying started around 20 wk, FI continued to increase in the R+ line but did not

TABLE 3. Quadratic regression[1] of standardized line RFC difference (Div) on the age (wk)

Line	a (± SE)	b (± SE)	c (± SE)
Males	–0.00389 ± 0.00062***	0.204 ± 0.024***	–1.41 ± 0.20***
Females	–0.00151 ± 0.00048**	0.104 ± 0.019***	–0.215 ± 0.175 NS

[1]Div = a(wk)^2 + b(wk) + c.

**P ≤ 0.01.

***P ≤ 0.001.

TABLE 4. Least squares means and coefficient of variation of traits measured on females of R⁻ and R⁺ lines[1]

Trait	R⁻	R⁺	CV
2-wk egg mass, g			
26 wk	441	441	40
30 wk	520	500	34
34 wk	497	463	38
Shank length, cm			
10 wk	9.1	9.3	4.4
14 wk	9.8	10.5	3.3**
22 wk	9.9	10.7	2.9**
34 wk	9.9	10.8	2.9**
Triiodothyronine, ng/mL			
4 wk	1.97	1.86	27
8 wk	1.55	1.74	22
12 wk	1.49	1.67	27
17 wk	0.72	0.99	44**
Thyroxine, ng/mL			
4 wk	10.8	10.4	34
8 wk	12.3	12.3	28
12 wk	11.7	9.7	28
17 wk	11.9	11.2	25

[1]R⁻ = low-feed intake line; R⁺ = high-feed intake line.
**Significant line difference ($P \leq 0.01$).

decrease in the R⁻ line. However, FI reached a plateau earlier in the R⁻ line than in the R⁺ line. Nevertheless, the stable final line difference (28 g/d) was reached well before 34 wk of age. This female line difference appeared to be surprisingly large when considering that the R⁻ hens were consistently heavier and produced marginally higher egg mass than R⁺ females. In both sexes, these large line FI differences were triggered when sexual maturity was being reached, at the same time as wattle size, a secondary sexual character that helps dissipate body heat, sharply diverged. Therefore, early and late values of FI certainly did not represent measures of the same trait from a genetic standpoint. Correspondingly, correlations between early and late FI, with values equal to 0.07 and 0.09 in males and females, were not significant. On the other hand, shank length already differed in females at 14 wk of age, well before sexual maturity was established.

The overall RFC time trend was quite similar for both sexes, with a fairly regular linear increase in the R⁺ line and the mirror image decrease in the R⁻ line. These results, confirmed by the curvilinear increase of standardized RFC line differences, show that RFC divergence between lines accumulates from early in life onwards. Indeed, the first highly significant RFC line difference was obtained in males and females contemporaneously to shank length difference, which is over 1 mo before FI divergence. Although the line divergence for RFC was established gradually, the fact that it was only well-identified at 14 wk of age may account for the absence of a significant correlation between early and late RFC. Finally, the decreasing trend for T_3 observed in both lines led to an increasing line difference for the level of this hormone that might be associated with divergence of FI.

Although previous genetic analyses of the selection experiment on RFC, through calculations of genetic correlations (Tixier-Boichard *et al.*, 1995), showed that wattle length, shank length, and FI line differences were the main correlated responses to direct selection on RFC, the biological relationships between those traits were still unclear. In the present work, those line differences were observed again, but we found that they became established in different ways and at different times during the growth of the birds. Divergence for RFC was early and synchronous to that of shank length. This observation seemed to be in fair agreement with the recent detection of segregation for a major gene associated with that morphological trait in F_2 crosses of R⁻ and R⁺ lines (Tixier-Boichard *et al.*, 1997). But, another stage of divergence between the two lines takes place around maturity, for both FI and wattle length, certainly contributing to the continued divergence observed for RFC. By then, however, shank length difference has become constant. As each one of those stages involved divergence for both a feed consumption trait and a morphological trait involved in body heat dissipation, it seems possible that selection on RFC may have had also a direct impact on those morphological traits and, through their heat loss function, a secondary action on FI.

Also, RFC diverged to the same extent in both sexes. Divergence occurred before the onset of laying, and it

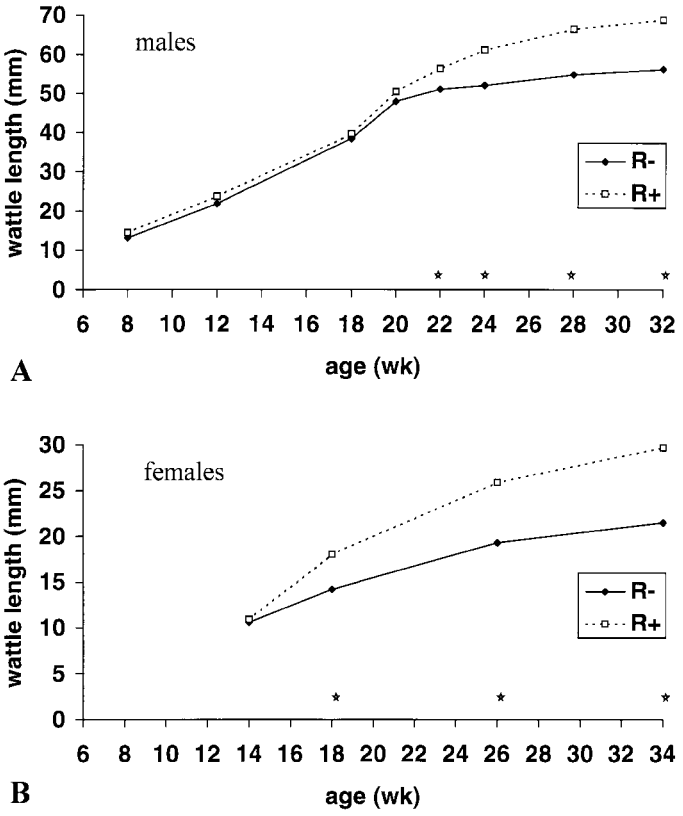


FIGURE 5. Wattle length in low intake R⁻ and high intake R⁺ lines. *$P < 0.01$.

was independent from line differences in feed intake. At that time, then, it appeared that RFC was the same trait in males and females, in contradiction with the absence of genetic and phenotypic correlations between sexes reported earlier in adult birds (Bordas *et al*., 1992).

From a practical poultry breeding standpoint, detailed observation of BW and FI patterns over time, along with plateaus observed by 28 wk of age, definitely validates the early report on RFC (Bordas and Mérat, 1975), which recommended that feed trial evaluation of RFC in adult birds be carried out over a short period only.

The present results suggest that energy metabolism of birds from the two lines should be monitored early (around 12 to 14 wk of age) and near sexual maturity (at 18 to 20 wk) to help understand discrepancies between some previous quantitative genetics results and the actual patterns described in this work. It would help also to better identify the physiological processes involved in the build up of RFC and in its divergence in adults between the high and low lines. It is known to be centered around diet-induced thermogenesis, the heat production due to the increase of metabolic rate after feed ingestion (Gabarrou *et al*., 1997). For example, one might test whether the relative contribution of basal metabolism and of diet-induced thermogenesis to RFC line differences would change with time, with appetite-related diet-induced thermogenesis only becoming determinant for RFC later in life. It remains also to be verified whether these observed patterns of co-evolution of BW, FI, RFC, and heat dissipating traits were only the consequence of artificial selection on RFC, or whether they represented a more general phenomenon that occurs during the growth of birds.

ACKNOWLEDGMENTS

The authors are very grateful to J. Buyse and E. Decuypere for hormone assays performed in the Laboratory of Physiology and Immunology of Domestic Animals, Catholic University of Leuven, Belgium, and they wish to thank B. Rivet for technical assistance in Jouy-en-Josas.

REFERENCES

Bordas, A., and P. Mérat, 1975. Enregistrement sur une courte période de la consommation d'aliment chez la poule pondeuse pour l'étude génétique de l'efficacité alimentaire. Ann. Génét. Sél. Anim. 7:331–334.

Bordas, A., M. Tixier-Boichard, and P. Mérat, 1992. Direct and correlated response to divergent selection for residual food intake in Rhode Island Red laying hens. Br. Poult. Sci. 33:741–754.

Bordas, A., P. Mérat, and F. Minvielle, 1996. Heterosis in egg laying lines under divergent selection for residual feed consumption. Poultry Sci. 75:20–24.

Byerly, T. C., 1941. Feed and other costs of producing market eggs. Page 29 *in*: Bulletin A1. Maryland Agriculture Experimental Station, University of Maryland, College Park, MD.

France, J., J. Dijkstra, and M. S. Dhanoa, 1996. Growth functions and their application in animal science. Ann. Zootech. 45s:165–174.

Gabarrou, J. F., P. A. Geraert, M. Picard, and A. Bordas, 1997. Diet-induced thermogenesis in cockerels is modulated by genetic selection for high or low residual feed intake. J. Nutr. 127:2371–2376.

Katle, J., 1992. Relationship between food consumption in growing chicks and their residual food consumption as adult laying hens. Acta Agric. Scand. 42:14–19.

SAS Institute, 1988. SAS® User's Guide: Statistics. SAS Institute Inc., Cary, NC.

Tixier-Boichard, M., A. Bordas, and P. Mérat, 1988. Divergent selection for residual feed intake in laying hens: effects on growth and fatness. Pages 129–132 *in*: Leanness in Domestic Birds. B. Leclercq and C. Whitehead, ed. Butterworths, London, U.K.

Tixier-Boichard, M., D. Boichard, E. Groenveld, and A. Bordas, 1995. Restricted maximum likelihood estimates of genetic parameters of adult male and female Rhode Island Red chickens divergently selected for residual feed consumption. Poultry Sci. 74:1245–1252.

Tixier-Boichard, M., A. Bordas, J. L. Coville, R. Bergé, F. Pitel, and A. Vignal, 1997. QTL detection in a cross between lines divergently selected for residual food consumption. Pages 123–130 *in*: Proceedings of the 3rd European Poultry Breeders Roundtable, Pruhonice, Czech Republic.

British Poultry Science (1992) 33: 741–754

DIRECT AND CORRELATED RESPONSES TO DIVERGENT SELECTION FOR RESIDUAL FOOD INTAKE IN RHODE ISLAND RED LAYING HENS

A. BORDAS, M. TIXIER-BOICHARD AND P. MÉRAT

Laboratoire de Génétique Factorielle, INRA, 78 352 Jouy-en-Josas Cedex, France

Received for publication 15th January 1991

Abstract 1. Divergent selection was undertaken in a Rhode Island Red population for residual food intake, measured in males and females, using mass selection.

2. In the absence of a control line, selection response during 14 generations was estimated by the within-year divergence between lines.

3. The direct response in residual food intake was found to be significant in both sexes, the divergence reaching almost three phenotypic standard deviations in each sex.

4. Significant correlated responses were obtained for food efficiency; it was improved in the low residual food intake line. Shank length, wattle length and rectal and comb temperature showed higher values in the high line, suggesting an increased heat production or dissipation. Inconsistent changes were observed for other egg production traits.

INTRODUCTION

Genetic improvement of efficiency of food utilisation in laying hens may be obtained by direct or indirect selection. Most of the traits which are currently under selection in laying stocks are positively correlated with the efficiency of conversion of food into eggs. Egg mass, body weight and change of body weight have been shown to account for variability in food consumption Fairfull and Chambers, 1984). However, some variability which remains unaccounted for offers a different approach to improving efficiency (Bordas and Mérat, 1981). By fitting a multiple linear regression equation a predicted food intake can be obtained as a function of egg mass, adult body weight and its change (Byerly *et al.*, 1980). The deviation of observed intake from its predicted value represents the residual food intake (R). This has been shown to be moderately heritable (Bordas and Mérat, 1981; Fairfull and Chambers, 1984; Hartmann and Mérat, 1986; Pauw, 1987).

Divergent selection offers a good experimental way of investigating the possible genetic improvement of a trait and its relationships with other eco-

nomically important traits. Such an experiment was undertaken for the R trait in a Rhode Island Red population, which showed interesting preliminary results after 8 generations of selection (Bordas and Mérat, 1984). Fourteen generations of selection have now been completed. Direct and some correlated responses are presented here.

ANIMALS AND SELECTION PROCEDURE

The Rhode Island Red base population used in this study was first measured for food intake during three generations of random mating. In 1975 (generation 0), males and females were chosen by selection within sire families on individual R record. Males and females of the upper half for R values became the parents of the first generation of the R^+ line while an equal number of individuals of the lower half were the parents of the R^- line. The two lines have been bred separately from 1976 on, with one generation each year, avoiding sib and half-sib matings. No unselected control line was maintained. Males and females have always been selected on their own individual performance for the R trait. A second stage of selection was added for the females from generation 5 on, consisting in culling the hens with the lowest egg number, (less than 10% of the birds).

Population size, environmental conditions and recording procedure have been described previously (Bordas and Mérat, 1984). Briefly, the number of sires for each line was originally 8 and was increased to 9 in generation 4. Each sire had a selected son, but dam family was not taken into account in the selection procedure. The number of dams varied from 4 to 5 per sire. The number of adult males with complete records ranged between 32 and 40 per line and year. The number of recorded adult females per generation ranged from 60 to 120 per line until generation 6 and from 100 to 160 thereafter. The R^+ line often had fewer available chicks because of a lower hatching rate. The data set finally used included only survivors' records between 1972 and 1989 for females ($n=3780$), and between 1975 and 1989 for males ($n=1064$). Complete pedigree data were used from 1972 to 1989, with a total number of 257 sires with progeny and 959 dams with progeny.

The food contained approximately 160 g total protein and 10·6 MJ of ME/kg, except in 1986 when a more concentrated food was used. The lighting regimen was 14 h light per day. The average ambient temperature was 20 ± 2°C (until 1985) and 22 ± 2°C thereafter.

Data registered included age at first egg (AFE), number of eggs to 39 weeks of age (EN) and egg weight (EW) at a mean age of 37 weeks. Information on food efficiency was obtained between 33 and 37 weeks of age for both sexes: individual food intake during 28 days (FI), mean adult body weight (BW) and change in body weight (ΔW); total egg mass (EM) was recorded over 28 d in females. The R criterion was the deviation of FI from a multiple regression equation which was fitted for each sex within each year. The same equation was used in both lines, in 1989 it was:

$$R=FI-(105{\cdot}1\ BW^{0.5}+1{\cdot}75\ \Delta W+0{\cdot}69\ EM-2326) \text{ for females,}$$

undertaken for the R trait
ed interesting preliminary
d Mérat, 1984). Fourteen
Direct and some correlated

d in this study was first
f random mating. In 1975
lection within sire families
upper half for R values
R^+ line while an equal
rents of the R^- line. The
with one generation each
ed control line was main-
l on their own individual
ction was added for the
hens with the lowest egg

cording procedure have
. Briefly, the number of
9 in generation 4. Each
ken into account in the
m 4 to 5 per sire. The
between 32 and 40 per
per generation ranged
100 to 160 thereafter.
f a lower hatching rate.
ords between 1972 and
89 for males ($n=1064$),
with a total number of

rotein and 10·6 MJ of
l was used. The lighting
nperature was 20±2°C

number of eggs to 39
37 weeks. Information
weeks of age for both
adult body weight (BW)
as recorded over 28 d
a multiple regression
The same equation was

6) for females,

$$R=FI-(91{\cdot}0\ BW^{0.5}+3{\cdot}67\ \Delta W-1907) \text{ for males.}$$

The value of 0·5 taken as the power of BW was found to be the most suited to our experimental conditions (Prod'homme, 1965). Food efficiency (FE) was calculated as the ratio of EM/FI. Wattle length (WL), shank length (SL), rectal temperature (TRE) and comb temperature (CT) were measured on both sexes at 9 months of age at the end of the period of feed recording. Records were available for all traits in all years, except for rectal temperature (missing before 1976 and from 1978 until 1981) and for comb temperature (missing before 1985).

METHODS

Divergence between lines

Environmental and genetic effects were analysed as fixed effects with the following linear model:

$$Y_{ijkl}=\mu+(\text{year}-\text{hatch})_{ij}+(\text{year}-\text{line})_{ik}+e_{ijkl} \quad (1)$$

where Y_{ijkl} is the record of the lth individual belonging to the kth line (k=base or R^- or R^+) born in the jth hatch ($j=1$, 2 or 3) of the ith year ($i=1972,\ldots,1989$) and e_{ijkl} is a random residual error normally distributed.

Analysis of variance and estimated effects were obtained with the GLM procedure of the SAS library. Each sex was analysed separately. The within-year divergence between R^+ and R^- lines provided an estimate of the response to selection, assuming the absence of interaction between line and environment. The estimated divergence was compared to the null hypothesis in order to conclude whether lines differed significantly or not. This was referred to as test 1, which only accounted for sampling variance. Drift variance also was important to consider because of the large number of generations. Estimates of sampling variance of selection response including drift variance were derived from Hill (1971; 1972).

The formula used was the following:

$$V(G_t)=\frac{3}{8}\sigma^2{}_m\left(\frac{th^2(1-h_0{}^2r^2)}{N_m}+\frac{(t-1)h^2h_0{}^2r^2}{M_m}\right)$$

$$+\frac{3}{8}\sigma^2{}_f\left(\frac{th^2(1-h_0{}^2r^2)}{N_f}+\frac{(t-1)h^2h_0{}^2r^2}{M_f}\right)+\sigma^2_p\left(\frac{MH+ML}{MH.ML}\right)$$

where G_t is the total response measured in generation t from the difference between the mean performance in the high line (sample size=MH) and the mean performance in the low line (sample size=ML), $\sigma^2{}_p$ is the phenotypic variance of the measured trait, different values for each sex were also used, $\sigma^2{}_m$

TABLE 1

Phenotypic means and standard deviations of production traits in females of the R^- and R^+ lines

Year	1975	1980		1985		1989	
Line	Base population	R^-	R^+	R^-	R^+	R^-	R^+
Sample size	73	128	127	161	102	147	151
Variables concerning food efficiency							
R (g)	2 ± 184[1]	−100	102	−132	209	−294 ± 200	286 ± 255
FI (g)	3087 ± 445	3005	3419	2654	3334	2934 ± 324	3625 ± 366
BW (g)	1967 ± 195	1988	2205	1998	2156	2046 ± 192	2091 ± 185
ΔW (g)	48·7 ± 73·7	25	31	7	14	38 ± 66	35 ± 72
EM (g)	947 ± 256	1073	1045	881	1008	1089 ± 205	1159 ± 219
FE	0·305 ± 0·066	0·356	0·305	0·331	0·303	0·373 ± 0·069	0·321 ± 0·062
Egg production variables							
AFE (ds)	177 ± 19	176	183	168	171	154 ± 12	158 ± 11
EN	55 ± 17	73	65	67	74	96 ± 25	97 ± 23
EW (g)	57·0 ± 4·4	56·7	57·4	52·4	54·7	51·6 ± 3·2	52·7 ± 3·2
Variables concerning heat production or dissipation							
TRE (°C)	...	...	...	40·37	40·36	40·46 ± 0·35	40·57 ± 0·36
CT (°C)	...	...	...	31·88	33·03	30·44 ± 2·01	32·38 ± 2·31
SL (mm)	109·4 ± 3·7	105·8	107·9	105·1	110·2	104·7 ± 3·4	109·5 ± 3·0
WL (mm)	28·14 ± 3·87	26·6	34·2	24·2	32·2	22·4 ± 3·4	31·0 ± 3·8

[1] Phenotypic standard deviation.

SL (mm)	109·4 ± 3·7					22·4 ± 3·4	
WL (mm)	28·14 ± 3·87	26·6	34·2	24·2	32·2		

[1] Phenotypic standard deviation.

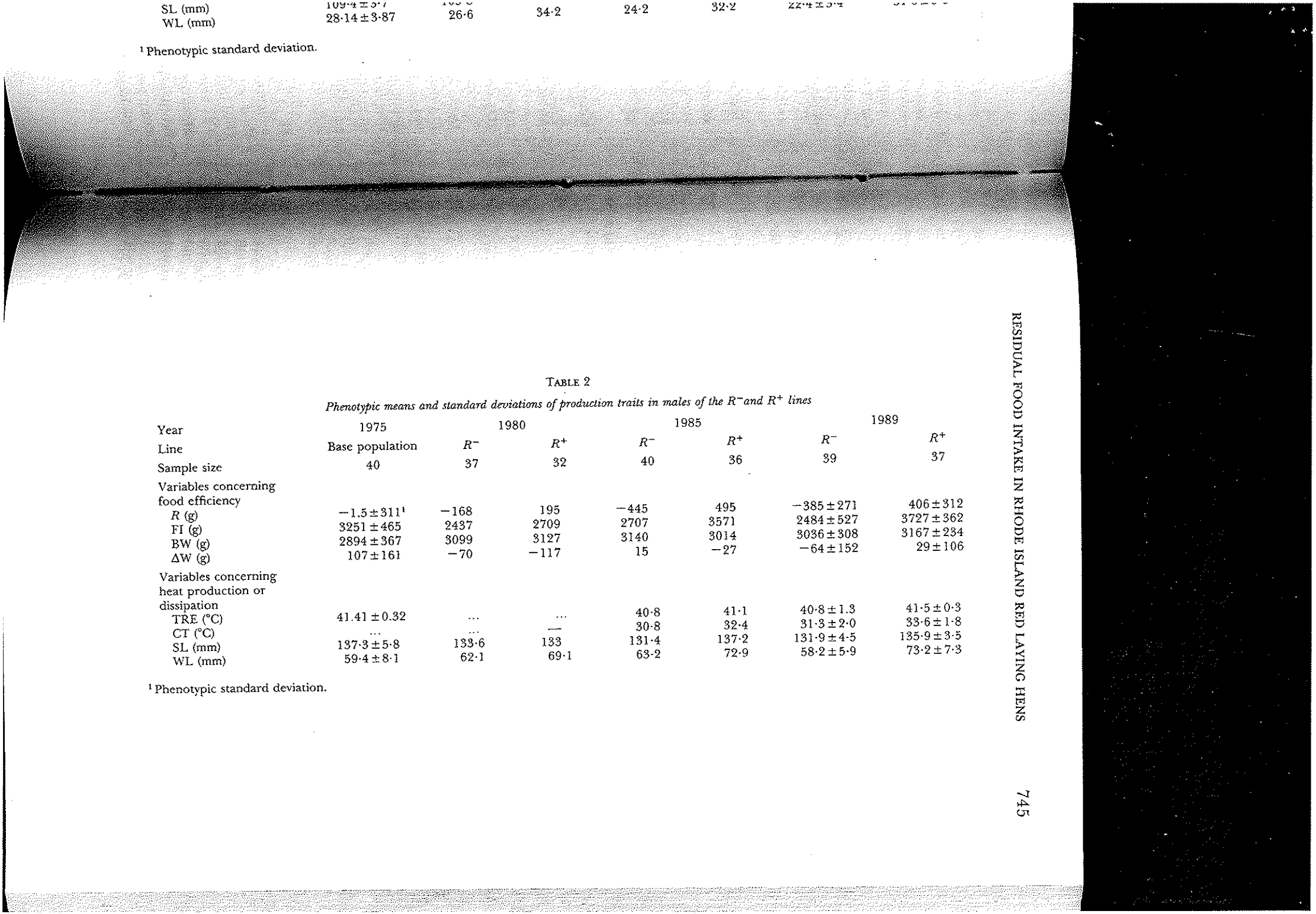


TABLE 2

Phenotypic means and standard deviations of production traits in males of the R^- and R^+ lines

Year	1975	1980		1985		1989	
Line	Base population	R^-	R^+	R^-	R^+	R^-	R^+
Sample size	40	37	32	40	36	39	37
Variables concerning food efficiency							
R (g)	−1.5 ± 311[1]	−168	195	−445	495	−385 ± 271	406 ± 312
FI (g)	3251 ± 465	2437	2709	2707	3571	2484 ± 527	3727 ± 362
BW (g)	2894 ± 367	3099	3127	3140	3014	3036 ± 308	3167 ± 234
ΔW (g)	107 ± 161	−70	−117	15	−27	−64 ± 152	29 ± 106
Variables concerning heat production or dissipation							
TRE (°C)	41.41 ± 0.32	...	...	40·8	41·1	40·8 ± 1.3	41·5 ± 0·3
CT (°C)	...	...	—	30·8	32·4	31·3 ± 2·0	33·6 ± 1·8
SL (mm)	137·3 ± 5·8	133·6	133	131·4	137·2	131·9 ± 4·5	135·9 ± 3·5
WL (mm)	59·4 ± 8·1	62·1	69·1	63·2	72·9	58·2 ± 5·9	73·2 ± 7·3

[1] Phenotypic standard deviation.

TABLE 3

Cumulative selection differentials in males and females of both lines, weighted by number of daughters

Line	Males		Females	
	R^-	R^+	R^-	R^+
Traits				
R (g)	−3255	3841	−1833	2249
FI (g)	−2894	5543	−639	3220
BW (g)	1033	1373	108	−248
EN	—	—	82·4	113·6
EW (g)	—	—	3·36	−3·36
TRE (°C)	−0·029	0·0366	0·1229	0·004
SL (mm)	−6·24	11·59	−2·80	2·54
WL (mm)	10·26	38·68	−0·59	7·64

TABLE 4

Total direct and indirect selection responses

Trait	Total response $(R^+)-(R^-)$ Generation 14 (year 1989)	Test for null hypothesis[1] in generation 14 Test 1	Test 2[2]	Test 3[3]
Direct				
R Males	759	***	***	***
R Females	580	***	***	***
Indirect				
FI (g) Males	1186	***	***	***
FI Females	691	***	***	***
BW g Males	223	***	NS	NS
BW Females	45	*	NS	NS
ΔW (g) Males	66	*	NS	NS
ΔW Females	−3	NS	NS	NS
EM (g)	70	**	NS	NS
FE g/j	−0·051	***	†	NS
EN	0·4	NS	NS	NS
AFE (d)	3·9	*	NS	NS
EW (g)	1·2	**	NS	NS
TRE °C Males	0·62	***	***	***
TRE Females	0·11	**	NS	NS
SL (mm) Males	3·8	***	NS	NS
SL Females	4·8	***	†	†
WL (mm) Males	15·9	***	***	***
WL Females	8·6	***	**	*

[1] Estimated response to selection differed from zero at $P<0·10$ (†); $P<0·05$ (*); $P<0·01$ (**); $P<0·001$ (***).
[2] Genetic correlation with selected trait assumed to be one.
[3] Genetic correlation with selected trait assumed to be zero.

for males and σ^2_f for females, h^2 is the heritability of the measured trait, h_0^2 is the heritability of the selected trait, r^2 is the squared genetic correlation between the measured and the selected traits, N_m and N_f are, respectively, the number of sires and dams used per generation in each line, M_m and M_f are, respectively, the number of recorded males and females per generation and line, a geometric mean was used to account for the change in population size from generation 0 to generation t.

Constant heritability values were used across generations. Inbreeding was

not taken into account. Because genetic correlations were generally not available, variance of correlated response was calculated for the two extreme situations where r^2 is set to one (test 2) or to zero (test 3). This also applied to the variance of direct response in each sex, so one may consider the possibility that residual food intake may be a different trait according to sex.

Selection differential

The residuals of linear model (1) represented the individual deviations from the within year × hatch × line contemporaries of the same sex. Selection differentials were obtained by averaging the residuals of model (1) weighted by the number of daughters for all the sires and dams in each line. Cumulative weighted selection differentials were then computed over the 14 generations of selection.

Inbreeding

Individual inbreeding coefficients were computed for all the animals present between 1972 and 1989, according to the algorithm of Quaas (1976). They were averaged by generation for each line and sex.

RESULTS

Phenotypic trends

Mean performances of both lines are shown in Table 1 for females and in Table 2 for males. Phenotypic standard deviations are given for the base population and for the last generation of selection. Fig. 1 shows the yearly phenotypic means of males and females in each line for residual food intake.

In generation 14, R^+ and R^- females differed by more than two standard deviations for the selected criterion, R, and males by 2·5 standard deviations, the R trait being more variable in males. A similar situation was observed for food intake. Body weight was slightly higher in R^+ birds. Gain in body weight during the recording period did not differ between lines.

In generation 14, egg mass and egg weight were slightly higher in R^+ hens. Food efficiency differed by nearly one standard deviation between lines with better efficiency in R^- hens. Yet both lines were more efficient than was the base population, which can be related to the improvement of egg number. Age at first egg also decreased in both lines but was always slightly higher in R^+ females.

Rectal temperature, comb temperature, shank length and wattle length were all greater in R^+ males and females. Divergence reached two standard deviations for wattle length. For rectal temperature, divergence between lines appeared relatively more important in males than in females.

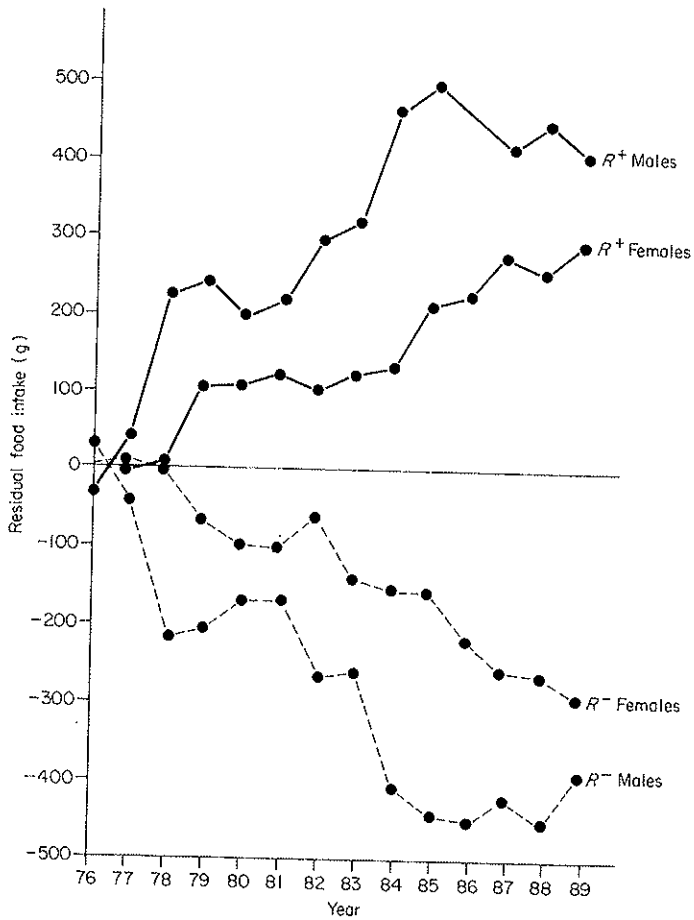


FIG. 1.—Generation means for residual food intake in both sexes.

Selection differentials

Selection pressure fluctuated in the first generations, but has remained fairly stable since 1982, being 22·5% on average in males of both lines, 30% in R^- females and 35% in R^+ females. Weighted selection differentials (WSD) are presented in Table 3 for sires and dams. Clearly, they are of opposite sign between lines for the R trait, food intake and shank length, in both sexes. Absolute levels of WSD were generally higher in males than in females and in the R^+ than in the R^- line. Selection differentials for body weight were positive in males but close to zero in females and slightly negative in R^+ females. For the traits measured in females only, WSD were positive for egg number in both lines, with slightly higher values in the R^+ line. For egg weight, WSD were of the same magnitude, but opposite in direction, with positive values in R^- and negative values in R^+.

Finally, some discrepancies appeared between lines and sexes for rectal

temperature and wattle length. WSD for rectal temperature were positive for R^+ males and R^- females, but negative for R^- males. A very small selection pressure was applied to R^- females for wattle length while positive WSD appeared for R^+ females. Males showed positive WSD for wattle length with much lower values in the R^- line.

Genetic trends

Selection responses are shown in Table 4. A significant difference between lines was said to be found when the null hypothesis was rejected according to test 2, (r^2 between traits = 1), or test 3, (r^2 between traits = 0). In contrast to test 1, which accounted only for sampling variance, some differences were no longer significant, particularly body weight and egg production traits.

A highly significant direct response to selection was found for the R criterion. The divergence between lines was in the expected direction and was significant since generation 3 for males and 4 for females. The difference between lines remained relatively constant for males after generation 11, but was still increasing in females.

In the case of the traits involved in the prediction equation for R, a significant correlated response was found for food intake in both sexes from generation 7 on (Fig. 2). The traits used to predict the residual were expected to show very weak correlations with the selection criterion. Indeed, the correlated responses did not differ from zero after accounting for estimated drift variance. Some significant divergences were found in generations 1 and 5 for body weight in females, because of higher values in the R^+ line. A difference of the same magnitude remained thereafter, but was not significant because of the increasing importance of drift variance. This situation suggests an effect of initial sampling rather than a correlated response to selection.

Body weight change during the food recording period did not show any

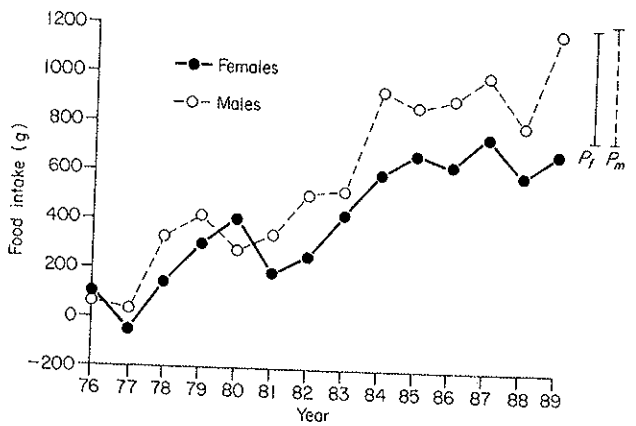


FIG. 2.—Divergence between lines (R^+ and R^-) in food intake (g). Phenotypic standard deviation of the base population is indicated by P_m for males and P_f for females.

correlated response. The divergence between lines showed some fluctuations for egg mass, generally to the advantage of the R^+ line, but no significant difference remained after accounting for drift (Table 4). Correlated response for food efficiency has been significant in several generations, with performance better in the R^- line (Fig. 3). Tests 2 and 3 gave different results regarding the significance level of the correlated response on food efficiency in the most recent generation (Table 4). Additional calculations of the sampling and drift variances for this response have been performed with different values for the genetic correlation (r) between food efficiency and residual food intake. The significance of the divergence ($P<0{\cdot}10$) obtained by test 2 for $r=1$, was true for any value of the genetic correlation (r) above 0·25.

The lines did not differ for egg number after accounting for drift variance (Table 4). Inconsistent fluctuations were found from year to year. Differences between lines for age at first egg and egg weight were rarely significant (Table 4), with a trend for higher values in the R^+ line. Again, the difference found for egg weight remained constant across generations, suggesting an effect of initial sampling rather than a correlated response to selection.

Divergence between lines was observed with higher values in R^+ birds for shank length (Fig. 4a) and wattle length (Fig. 4b) in both sexes and for rectal temperature in males (Table 4). Correlated response to selection became significant very early for wattle length (generation 2) and later for rectal temperature in males (generation 7), or for shank length (generation 9).

The inbreeding level increased regularly during the selection experiment, at a slightly higher rate in the R^+ line. Final values are moderate, being 28% in the R^+ line and 25% in the R^- line.

DISCUSSION

Reliability of the methods

The estimates of selection response could be biased in the case of genotype × environment interaction. Recent observations suggested that the differences between lines for food intake and R were not affected by the percentage of total protein in the food (Bordas and Mérat, 1991), excluding one cause of genotype × nutrition interaction. An interaction between genotype and ambient temperature is possible and further experimental data are needed.

Expected drift variance was taken into account to compare the differences between lines with the null hypothesis. A constant value of heritability was assumed to calculate the drift variance, while it is known that selection and inbreeding will decrease genetic variance. Therefore, drift variance might have been over-estimated. However, the direct response still appeared to be highly significant. The effect of drift on correlated traits depends on the importance of the genetic correlations with the selected trait. These correlations should be studied in the future. However, significant correlated responses have been found for some traits, even after maximising the drift variance (test 3). The

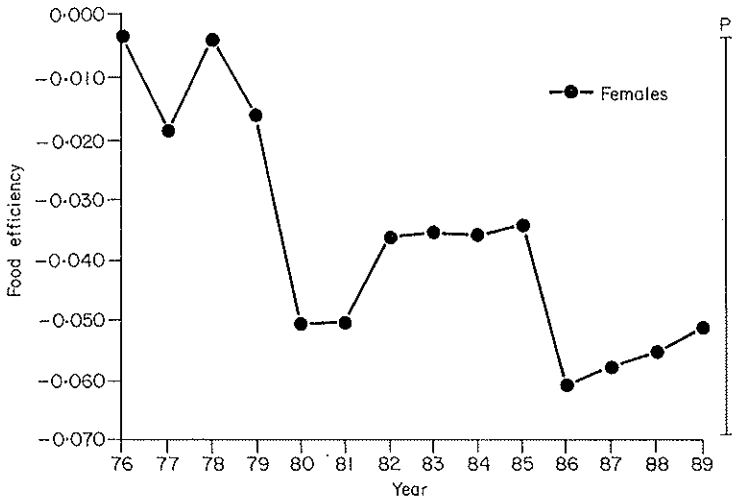


FIG. 3.—Divergence between lines (R^+ and R^-) in food efficiency (egg mass/food intake), of laying hens. Phenotypic standard deviation is indicated by P.

very small difference in inbreeding between lines should not affect the comparison of the genetic trends between lines.

Direct response and realised heritability

Selection succeeded in creating a clear divergence in the R trait, which represented a similar proportion of the phenotypic standard deviation in both sexes. Correlations between brother and sister for the R trait were close to zero within each generation (A. Bordas, unpublished data), which suggests that R is a different trait in males and females. A rough estimate of realised heritability could be obtained from the ratio of total selection response (estimated by the divergence between lines at the most recent generation for each sex), to cumulative selection differentials (Table 4). If the genetic correlation between sexes for R was one, the realised h^2 for R would be 0·12; if the correlation was zero, the realised h^2 for R would be 0·28 in females and 0·21 in males. These estimates are lower than those found by Katle and Kolstad (1991), when selecting only on the R trait of White Leghorn females. Lower estimates had been found in an unselected RIR population (Bentsen, 1983). In unselected White Leghorn strains, estimates of h^2 were found to lie within the same range of variation (Pauw, 1987). More recent studies have yielded heritability values for the R trait between 0·23 (Muller, 1989) and 0·45 (Luiting and Urff, 1991).

Correlated responses on production traits

Food efficiency responded to selection and lines differed by one phenotypic standard deviation, which represents more than 20% of the performance level of the base population. Other selection experiments on residual food

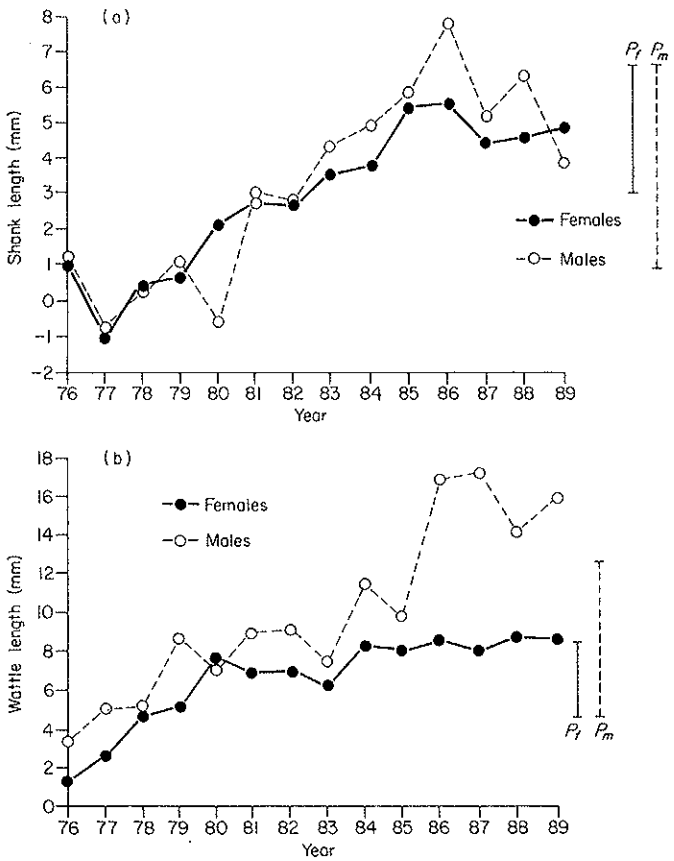


FIG. 4.—Divergence between lines (R^+ and R^-): (a) in shank length (mm). (b) in wattle length (mm). Phenotypic standard deviation of the base population is indicated by P_m for males and P_f for females.

intake appear also to be successful in improving food efficiency without affecting egg production (Katle and Kolstad, 1991; Liuttula, 1989).

No significant correlated responses were observed on egg number, age at first egg and egg weight. Positive selection differentials were found for egg number in both lines, which must be related to the addition of this trait in the selection procedure. A change in age at first egg could have been expected on the basis of a negative correlation found with the R trait by Bentsen (1983). However, no significant phenotypic correlation was found between the R trait and age at first egg in our population (Bordas and Mérat, 1981). In a RIR line a low and positive phenotypic correlation was found between residual food intake and egg weight at the age of 60 weeks (Bentsen, 1983). The present study involves egg weight at a different age (37 weeks), for which the correlation might be different.

Correlated responses on traits related to heat production and dissipation

Selection in males seemed to have particularly affected some traits related to heat production or dissipation (TRE, CT, SL, WL), as shown by selection differentials and genetic trends. The same situation can be found in females for shank length but not so clearly for rectal temperature. Indeed, selection for the R trait of female White Leghorns showed no correlated response on rectal and comb temperature nor on wattle length, but had an effect on shank length (Katle, 1991). Shank length is related not only to heat dissipation, but also to body size and maintenance requirement (Damme and Pirchner, 1984). Very low phenotypic correlations were found in White Leghorns between R and wattle or shank length (Katle *et al.*, 1984). Conversely, a positive phenotypic correlation was found in our base population between the R trait and wattle length in both sexes (Bordas and Mérat, 1981).

Heat production as measured in a respiration chamber was found to be increased in a R^+ line (Katle, 1991), and was also found to differ between R^- and R^+ type White Leghorn hens (Luiting, 1986). Both components of activity-related and activity-free heat production were increased in White Leghorn layers showing a high residual food consumption (Luiting *et al.*, 1991). Layers from the present R^- and R^+ lines have been studied in respiration chambers. A difference between lines was found, especially for heat production in the fed state and for diet-induced thermogenesis (Geraert *et al.*, 1991). Abdominal fatness tends to be slightly, but not significantly higher in adult chickens from the R^- line (Tixier *et al.*, 1988). Lines do not differ greatly in feather condition at 40 weeks of age.

Behavioural differences were found between lines divergently selected on the R trait (Braastad and Katle, 1989). The present lines also differed in food pecking activity (A. Bordas, unpublished data), which was higher in R^+ line.

ACKNOWLEDGEMENTS

The authors are very grateful to Didier Boichard from the Station de Génétique Quantitative et Appliquée at INRA, Jouy-en-Josas, for providing the program for inbreeding analysis.

REFERENCES

BENTSEN, H. (1983) Genetic variations in feed efficiency of laying hens at constant body weight and egg production. II Sources of variation in feed consumption. *Acta Agriculturа Scandinavica*, **33**: 305–320.

BORDAS, A. & MÉRAT, P. (1981) Genetic variation and phenotypic correlations of food consumption of laying hens corrected for body weight and production. *British Poultry Science*, **22**: 25–33.

BORDAS, A. & MÉRAT, P. (1984) Correlated responses in a selection experiment on residual feed intake of adult Rhode Island Red cocks and hens. *Annales Agriculturae Fenniae*, **23**: 233–237.

BORDAS, A. & MÉRAT, P. (1991) Sélection divergente pour la consommation alimentaire 'résiduelle' de la poule en période de ponte: réponse au taux protéique de l'aliment. *Genetics Selection Evolution*, **23**: 249–256.

BRAASTAD, B.O. & KATLE, J. (1989) Behavioural differences between laying hen populations selected for high and low efficiency of food utilisation. *British Poultry Science,* **30:** 533–544.

BYERLY, T.C., KESSLER, J.W., GOUS, R.M. & THOMAS, O.P. (1980) Feed requirements for egg production. *Poultry Science,* **50:** 2500–2507.

DAMME, K. & PIRCHNER, F. (1984) Predicting daily feed intake of laying hens by supplementary information on wattle, shank and bare body surface. 35th Annual Meeting of the EAAP, Den Haag, 6–9 August 1984, 10 p.

FAIRFULL, R.W. & CHAMBERS, J.R. (1984) Breeding for feed efficiency: poultry. *Canadian Journal of Animal Science,* **64:** 513–527.

GERAERT, P., GUILLAUMIN, S., BORDAS, A. & MÉRAT, P. (1991) Evidence of a genetic control of diet induced thermogenesis in poultry. *12th Symposium on Energy and Metabolism in Farm Animals,* Kartause Ittingen, 1–7 September 1991, *EAAP Publications,* **58:** 380–383.

HARTMANN, W. & MÉRAT, P. (1986) Recent research on the breeding of egg production stock. *7th European Poultry Conference, Paris,* 24–28 August 1986, **I:** 49–59.

HILL, W.G. (1971) Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics,* **27:** 293–311.

HILL, W.G. (1972) Estimation of realised heritabilities from selection experiments. I. Divergent selection. *Biometrics,* **28:** 747–765.

KATLE, J. (1991) Selection for efficiency of food utilisation in laying hens: causal factors for variation in residual food consumption. *British Poultry Science,* **32:** 955–969.

KATLE, J. & KOLSTAD, N. (1991) Selection for efficiency of food utilisation in laying hens: direct response in residual food consumption and correlated responses in weight gain, egg production and body weight. *British Poultry Science,* **32:** 939–953.

KATLE, J., BENTSEN, H.B. & BRAASTAD, B.O. (1984) Correlated traits with residual feed consumption. *Proceedings of the XVII World's Poultry Congress,* Helsinki, 8–12 August 1984, 136–138.

LIUTTULA, M. (1989) Selection experiment for feed efficiency in egg laying hens. *Nordish Jordbrugsforckning,* **71:** 428.

LUITING, P. (1986) Metabolic differences between White Leghorns selected for high and low residual feed consumption. *7th European Poultry Conference, Paris,* 24–28 August 1986, **I:** 118–121.

LUITING, P., URFF, E.M. (1991) Residual feed consumption in laying hens. 2. Genetic variation and correlations. *Poultry Science,* **70:** 1663–1672.

LUITING, P., SCHRAMA, J.W., VAN DER HEL, W. & URFF, E.M. (1991) Metabolic differences between White Leghorns selected for high and low residual feed consumption. *British Poultry Science,* **32:** 763–782.

MULLER, S. (1989) Investigations of selection for feed conversion in laying hens. *8th International Symposium on Current Problems in Avian Genetics, Smolenice,* April 3–6, 43–47.

PAUW, R. (1987) Analysis of several selection criteria evaluating efficiency in laying hens. Doctoral Thesis. University of Bonn, Fed. Rep. Germany, 165 pp.

PROD'HOMME, J. (1965) Calcul d'une équation permettant de prévoir la consommation alimentaire de pondeuses en fonction de leurs performances. *Annales de Zootechnie,* **14:** 335–339.

QUAAS, R.L. (1976) Computing the diagonal elements and inverse of a large numerator relationships matrix. *Biometrics,* **32:** 949–953.

TIXIER, M., BORDAS, A. & MÉRAT, P. (1988) Divergent selection for residual feed intake in laying hens: effects on growth and fatness, in: LECLERCQ, B. & WHITEHEAD, C.C. (Eds) *Leanness in Domestic Birds,* 129–132 (London, Butterworths).

British Poultry Science, 22 25-33. 1981 0007-1686/81/00040025$02.00

GENETIC VARIATION AND PHENOTYPIC CORRELATIONS OF FOOD CONSUMPTION OF LAYING HENS CORRECTED FOR BODY WEIGHT AND PRODUCTION

A. BORDAS AND P. MERAT

Laboratoire de Genetique Factorielle, Centre National de Recherches Zootechniques, INRA, 78350 Jouy en Josas, France

Received for publication 15th April 1980

1. On samples from two brown-egg strains between 1967 and 1979, " residual " food consumption (*i.e.* food consumption corrected for body weight, weight change and egg mass) of laying hens was investigated with respect to variation between sire families and phenotypic correlations with production traits, egg traits, morphological and physiological variables. A similar residual food intake (corrected for body weight and body-weight gain) was obtained for samples of males in some years.

2. Highly significant differences between sire families were observed for residual food consumption in both strains for females and in one strain for males.

3. In females, highly significant positive phenotypic correlations were found in both populations for residual food intake with wattle length and shank temperature, and in one population with shank length and width. On the whole, a negative correlation was observed with egg-shell thickness. For two variables recorded only in one strain, proportion of broken eggs and yolk : albumen ratio, there were highly significant positive correlations.

4. For males also, wattle length and shank temperature were positively correlated with residual food intake.

5. The physiological meaning and possible use of correlated variables as criteria for indirect selection for food efficiency of layers are discussed.

INTRODUCTION

The efficiency of food conversion by laying hens depends on their egg production and body weight and it has been considerably improved by breeding to increase the former and to decrease the latter. This genetic improvement, however, may have limits, hence it is of interest to consider that part of the variance in food intake which is independent of egg production and body size.

The significance of this " residual " variance is likely to depend on the population being analysed. Arboleda (1971), Nordskog *et al.* (1972), Lee and Nordskog (1975) and Arboleda *et al.* (1976) did not find this type of variation in White

Leghorn layers. Conversely, significant variation in this respect was found by Bordas and Mérat (1974), Watanabe *et al.* (1975) and Hagger (1977*a*, *b*) within lines, and Filmer (1974) and Farrell (1975) between lines. Results from Hurnik *et al.* (1977) within lines, and Glazener and Blow (1954) between lines also indicate such variation. On the other hand, Mérat (1968), Mérat and Bordas (1972, 1974, 1979), Mérat *et al.* (1979) and Bordas and Mérat (1976) showed the influence of genotype at particular loci on this residual component of food consumption.

Little information is available concerning the correlations between residual food consumption and production traits as well as physiological or morphological traits. Reviews on various aspects of food efficiency (*e.g.* Filmer, 1974; Balnave, 1974) mention variations in basal metabolism, plumage condition or physical activity as likely components of this efficiency. Within two experimental strains Mérat and Bordas (1974) observed several significant phenotypic correlations of food intake independent of egg production, body weight and weight change, especially with size of unfeathered appendages, comb, wattles and shanks. In addition Heil (1976) and Orlov and Tuchemski (1974) suggested a relation with internal composition of the egg. From another viewpoint the relation between feathering of laying hens and food consumption has been investigated mainly by comparisons between strains by Charles (1976), Emmans and Charles (1976), Wathes (1976) and Leeson and Morrison (1978).

The present paper describes further our own results that complete and enlarge those previously published (Bordas and Mérat, 1974).

MATERIALS AND METHODS

Genetic material and husbandry

Two populations were used: a " synthetic " strain (termed Jouy) of brown-egg-laying type, containing segregations at known loci, and a Rhode Island Red strain (M99) derived from a strain of Le Magneraud Research Station. In each population, individual food consumption was recorded on samples over several years within the period 1967 to 1979. The first strain was hatched in autumn and the second in spring.

Chicks were reared in floor pens to 16 weeks of age, then pullets kept for measurement of production were transferred to individual laying cages. Recording of food intake, body weight and egg production began at 8 months of age and continued for three successive 28-d periods (Jouy) or one such period (M99). Hens were fed *ad libitum* a layer mash, or pellets in 1970 and 1971, with 160 g crude protein and 10·6 MJ ME/kg. The diet formula was the same for all years. Food wastage could not be accurately evaluated but was considered to be small.

In certain years body weight and food intake were measured in similar conditions and for the same duration for a sample of males of the same origin as the females.

Measurements and observations

For each bird the average of the three 28-d periods (Jouy) or the value for the one period (M99) was taken as food intake (Y_O), and average body weight during

the period (W), body-weight change from beginning to end of the period (ΔW), and egg mass produced (E) were recorded. An expected food consumption Y_T was then estimated for each bird from a multiple regression equation using W, ΔW and E, (a separate equation being estimated for each strain and for each year):

$$Y_T = aW^{\alpha} + b\Delta W + cE$$

This equation is similar to those of Byerly (1941) and others *e.g.* Leeson *et al.* (1973), Gous *et al.* (1978), McDonald (1978). For α the value 0·5 was used as this was close to the value which minimised variance in our conditions. Values for α of between 0·5 and 1·0, however, had little influence in this respect. A common estimate for the parameters of this equation in our experiments has been published previously, (Mérat *et al.*, 1979).

The difference between observed (Y_O) and expected (Y_T) food consumptions for each bird is termed " residual " food consumption (Y_R) : $Y_R = Y_O - Y_T$. We investigated the significance of its genetic variance and phenotypic correlations with the following variables: observed food intake (g), 8-week body weight (g), age at first egg (d), egg number (from first egg to 11 months of age), average egg weight over 2 weeks in the last 28-d period (g), albumen height—average of 2 eggs per hen in the last period (to within 0·1 mm), shell thickness—average of 2 to 4 eggs per hen in the last period (to within 0·01 mm), proportion of cracked eggs over the whole period of measurement (%), and ratio of yolk weight to albumen weight (one egg per hen, last period).

The following morphological or physiological measurements were taken at about the age of 10 months: haematocrit (%), thickness of dorsal plumage between the skin and outer limit of uncompressed plumage (mm), wattle length (mm), shank length (mm), shank width (mm), shank temperature—measured once, in the afternoon, rectal temperature—measured once, in the morning, and comb temperature—measured once, in the afternoon (°C).

Two behavioural measurements were taken in an " open field " on 2-d-old chicks, as described by Faure and Folmer (1975): latency time (s) before a recorded displacement, and number of tops (crossing a light beam) within 1 min.

Only females for which all of the traits measured in a given year were recorded are considered, but certain measurements were not done every year, in particular those on egg components (yolk : albumen) and open-field measurements. Data from a few hens which had not laid during the measurement periods were discarded. From 1974 the plumage of hens at the end of the period of measurement was noted as " deteriorated " (with bare patches) or as not, so as to evaluate the relationship of this criterion with food efficiency.

Except for laying and egg traits the same measurements were made on males.

Statistical analysis

The presence of a genetic variation in residual food intake was tested by variance analysis between sire families within years, pooling all years for each population and sex.

The phenotypic correlations of Y_R with other traits were estimated for each strain and sex over all years on a within-year basis after testing homogeneity between years.

For the binominal trait deteriorated or non-deteriorated plumage, Y_R was compared by a *t*-test on pairs of full or half-sisters, one with intact plumage, the other with deteriorated plumage.

RESULTS AND DISCUSSION

The main results are shown in Tables 1, 2 and 3. For plumage condition at the end of the period of measurement a total of 75 pairs (deteriorated or non-deteriorated) was available. The mean value of Y_R was greater by 93·4 g for the former ($P < 0{\cdot}02$).

TABLE 1

Variance analysis of residual food intake (Y_R) *in two populations of hens and cockerels by sire families within years*

Sex and strain	Degrees of freedom: Sires	Residual	*F*
Females			
Jouy	49	958	2·56***
M99	74	1073	2·41***
Males			
Jouy	16	97	1·60
M99	57	219	2·49***

*** Significant ($P < 0{\cdot}001$)

Variance between families

The highly significant *F* values in Table 1 (except for Jouy males, possibly due to their limited number) confirm our previous finding (Bordas and Mérat, 1974) of the existence of between-sire variation in residual consumption in both populations of adult birds, especially laying hens. This implies a significant heritability of this trait in these populations. This was not estimated, because of the small size of dam families in these data.

Phenotypic correlations of Y_R and suggested interpretations

Owing to the definition of Y_R, which makes it statistically independent of body weight, its correlations with body measurements cannot be attributed to a general effect of body size (Tables 2 and 3).

For the highly significant positive correlation of Y_R with wattle size in both sexes, the simplest explanation seems to be that an appreciable amount of the total heat dissipated by the adult bird (15% according to Sturkie, 1965) arises from the unfeathered head appendages. Thus the larger these appendages are, the greater will be the heat dissipation, and our results suggest that an increase in food intake is the mechanisn by which the concomitant increase in energy need is compensated. This is consistent with the reduction (about 2%) of food intake associated with the pea-comb gene which limits the size of these appendages (Mérat and Bordas, 1979).

A similar interpretation may be suggested for the positive correlation, also highly significant in both sexes, of Y_R with surface temperature of the shank, despite

TABLE 2

Phenotypic correlations of residual food intake (Y_R) *with measured traits for hens of two strains*

	Strain					
	Jouy		M99		Total[2]	
Trait[1]	DF	r	DF	r	DF	r
8-week body weight	735	+0·04	966	−0·06	1701	−0·02
Age at 1st egg	449	+0·05	1006	−0·04	1455	−0·01
Egg number	958	−0·02	1073	+0·04	2031	+0·01
Mean egg weight	958	+0·01	1073	−0·01	2031	+0·01
Albumen height	958	+0·05	1073	−0·10**	2031	(−0·03)
Shell thickness	958	−0·04	1073	−0·10**	2031	−0·07**
Broken eggs	636	+0·17***		...		...
Yolk : albumen	281	+0·22***		...		...
Plumage thickness	233	+0·01	369	−0·04	602	−0·02
Wattle length	905	+0·14***	977	+0·11**	1882	+0·12***
Shank length	905	+0·25***	977	−0·01	1882	(+0·11)
Shank width	415	+0·18***	639	+0·00	1054	(+0·07)
Shank temperature	607	+0·14***	452	+0·18**	1059	+0·16***
Comb temperature	50	+0·15	80	+0·08	130	+0·11
Rectal temperature	346	+0·03	372	−0·01	718	+0·01
Haematocrit	384	+0·07	452	−0·00	836	+0·03
Latency time	47	+0·29*	171	+0·06	218	+0·11
Number of tops	47	−0·24	171	+0·02	218	−0·04
Observed food intake (Y_O)	958	+0·52***	1073	+0·52***	2031	+0·52***

Significant * $P<0·05$, ** $P<0·01$, *** $P<0·001$.

[1] Traits represent those tested for correlation with Y_R, descriptions and units are given in the text. Correlations for body weight and body-weight gain were close to zero and have been omitted. For the correlation between Y_O and Y_R positive deviations from zero are expected, the values being the ratio of the standard deviation of Y_R to that of Y_O (assuming no correlation between Y_T and Y_R). These ratios were estimated from our data as 0·59 (Jouy) and 0·54 (M99).

[2] Values in parentheses showed significant heterogeneity for strains.

TABLE 3

Phenotypic correlations of residual food intake (Y_R) *with measured traits for cockerels of two strains*

	Strain					
	Jouy		M99		Total[2]	
Trait[1]	DF	r	DF	r	DF	r
8-week body weight	174	+0·02	258	−0·01	432	+0·00
Plumage thickness	...		113	+0·19*	...	
Wattle length	26	+0·25	258	+0·22***	284	+0·22***
Shank length	26	+0·26	258	+0·10	284	+0·11
Shank width	...		221	−0·02	...	
Shank temperature	26	+0·07	134	+0·27**	160	+0·24**
Rectal temperature	...		134	+0·15	...	
Haematocrit	26	−0·06	134	+0·06	160	+0·04
Latency time	...		36	−0·44**	...	
Number of tops	...		36	−0·03	...	
Observed food intake (Y_O)	118	+0·75***	258	+0·74***	376	+0·74***

Significant * $P<0·05$, ** $P<0·01$, *** $P<0·001$.

[1] As Table 2. The estimated ratio of standard deviations for Y_R and Y_O for the total was 0·75.

[2] No heterogeneity for strains was shown for any value.

the facts that this measurement was made only once per individual and that this temperature may reflect a response to stress (Duncan, 1973; Duncan and Filshie, 1976). Comb temperature, measured on females, seemed to show a similar trend, although this was not significant probably because of the limited numbers available. Conversely, there was no correlation with rectal temperature, which shows little variation and consequently is likely to be more dependent on variation associated with the conditions of the measurement itself. In accordance with the findings of others previously mentioned, we suggest, as an explanation of the difference in food consumption associated with plumage condition, the greater heat loss arising from poorer insulation in birds with deteriorated plumage. We have commented previously (Mérat *et al.* 1979) that the relatively low incidence and amount of this deterioration in our conditions may be attributed to the use of individual cages. On the contrary, it is interesting to note that back plumage thickness, where there is no involvement of bare patches, showed no correlation with Y_R (except a questionable positive correlation among M99 males).

The case for shank length and width is not clear. A physical interpretation of the highly significant positive correlation between Y_R and shank length in the Jouy population might be invoked (increased heat loss associated with the greater size of an unfeathered body part), but this remains doubtful because superficial blood supply to the shanks is not comparable with that to the comb and wattles, and the correlation was not found in the M99 population.

As expected, total food consumption (Y_O) showed a fairly high correlation with Y_R. This correlation was higher among males than among females, which also was predictable, because for the latter, egg production is an additional source of variation for Y_O independent of Y_R. It appears normal also that no appreciable correlation was found between Y_R and egg number and mean egg weight, as Y_R is defined as statistically independent of egg mass. On the other hand, the fact that no significant correlations existed between Y_R and growth rate and age at first egg is worthy of note, as well as the absence of any correlation with albumen height in the Jouy population (the slightly negative correlation of this trait with Y_R in the M99 population has no obvious explanation).

The significant positive association between residual food consumption and the ratio of yolk to albumen weight probably reflects the increase in dry matter percentage and energy content in the egg when the proportion of yolk increases (Romanoff and Romanoff, 1949). Heil (1976) observed a slight positive correlation between food conversion and yolk weight expressed as a proportion of egg weight, which seems to accord with the present result. Orlov and Tuchemski (1974) found a positive correlation between food conversion and the protein and lipid content of eggs. On the other hand, Hagger (1977*a*) explained part of the difference in food efficiency between two lines of laying hens by a difference in the dry matter content of eggs.

The slight negative association of Y_R, on the whole, with egg-shell thickness, and the highly significant positive correlation with porportion of cracked eggs can not yet be explained, but they may be compared with the observations of others. Arboleda *et al.* (1976) mentioned a negative correlation between specific gravity of eggs and observed (uncorrected) food intake. We also found with the same hens as those used in the present work that body-weight gain (ΔW) during the period of

measurement also showed a negative correlation with egg-shell thickness ($-0{\cdot}16$ with fixed body weight, 1559 DF, $P<0{\cdot}001$). It is known that this body-weight gain is associated with additional food consumption (Bordas and Mérat, 1976; Leclercq *et al.*, 1977) and increased fatness (Svenson, 1964; Leclercq *et al*, 1977). Moreover, Tierney (1979 and personal communication) concluded that selection for, or restriction of feeding to effect, control of fat deposition leads to an improvement in shell quality as measured by specific gravity.

Finally, neither haematocrit value nor measurements of activity in open-fields showed any definite correlations with Y_R: the significant positive correlation with latency time in Jouy females is of doubtful meaning, as it was not found in the other strain nor in the total. In males this variable was negatively correlated with Y_R, a finding that should be confirmed.

On the whole, taking account only of variables showing highly significant correlations with Y_R among females in the Jouy strain (wattle length, shank length, width and temperature, broken eggs, yolk : albumen ratio), the proportion of the total variance in Y_R which is explained by these correlations is about 20% (to which about 1% should be added for the association with plumage condition). Obviously, a large part remains unexplained and is likely to correspond to variations in physiological factors such as basal metabolic rate and regulation of appetite.

CONCLUSIONS

The consequences of these findings for breeding programmes are worthy of consideration. The residual food intake of laying hens, which is independent of egg laying and body size, is costly to improve directly by breeding. However, a limited improvement might be obtained by indirect selection for a few easily measured traits which are correlated with this variable. This may be the case for the size of unfeathered head appendages and for plumage condition after several months' production, if the heritabilities of these traits are appreciable and they reveal no unfavourable correlations with production traits. It seems likely that genetic correlations of these morphological traits with Y_R would be in the same direction as phenotypic ones, from the supposed physical basis underlying them. We have also suggested (Bordas and Mérat, 1974) elimination of hens with excessive water consumption. It might be more questionable however to consider surface temperatures (*e.g.* shank temperature) for indirect selection, because of their possible relation with response to stress factors (Duncan, 1973).

Genetic improvement of Y_R may become economically significant as gains on egg production tend to a plateau. A survey of phenotypic correlations of this variable with other variables of economic significance suggests, provided that genetic correlations are in the same direction, that its improvement by selection should have no harmful effects, except possibly a slight reduction in yolk proportion. An additional advantage might be a reduction in the incidence of broken eggs.

Note added

Since completing this manuscript we have become aware of the findings of Bentsen (1980). These concern genetic variation for residual food conversion

efficiency and correlations of this trait, most of which are fairly similar to those presented here.

REFERENCES

ARBOLEDA, C. R. (1971). Economic feasibility of measuring feed consumption in a poultry breeding operation. *Ph.D. thesis, Iowa State University.*

ARBOLEDA, C. R., HARRIS, D. L. AND NORDSKOG, A. W. (1976). Efficiency of selection in layer-type chickens by using supplementary information on feed consumption. *Theoretical and Applied Genetics*, **48**: 75-83.

BALNAVE, D. (1974). Biological factors affecting energy expenditure. In: *Energy Requirements of Poultry*. Edit. MORRIS, T. R. AND FREEMAN, B. M., pp. 25-46. Edinburgh, British Poultry Science Ltd.

BENTSEN, H. B. (1980). Inherited variance in feed efficiency by laying hens. *6th European Poultry Conference, Hamburg*, 2: 40-44.

BORDAS, A. AND MERAT, P. (1974). Variation génétique et correlations phénotypiques caractérisant la consommation alimentaire de poules pondeuses après correction pour le poids corporel et la ponte. *Annales de Génétique et de Sélection animale*, 6: 369-379.

BORDAS, A. AND MERAT, P. (1976). Gain de poids pendant la ponte et sélection sur l'efficacité alimentaire chez la poule. *Annales de Génétique et de Sélection animale*, **8**: 527-535.

BORDAS, A., OBEIDAH, A. AND MERAT, P. (1978). On the heredity of water intake and feed efficiency in the fowl. *Annales de Génétique et de Sélection animale*, **10**: 233-250.

BYERLY, T. C. (1941). Feeds and other costs of producing market eggs. *Bulletin. University of Maryland Agricultural Experimental Station*, **A1**.

CHARLES, D. R. (1976). The economic importance of feathering. *Gleadthorpe Poultry Booklet*, pp. 25-27. Mansfield, England. Gleadthorpe Experimental Husbandry Farm.

DUNCAN, I. J. H. (1973). Can the psychologist measure stress? *New Scientist*, **60**: 173-175.

DUNCAN, I. J. H. AND FILSHIE, J. H. (1976). Radiotelemetry of avian shank temperature using a thin film hybrid microcircuit. *Medical and Biological Engineering*, **13**: 544-550.

EMMANS, G. C. AND CHARLES, D. R. (1976). Climatic environment and poultry feeding in practice. In: *Nutrition and Climatic Environment*, pp. 31-49. Edit. HARESIGN, W., SWAN, H. AND LEWIS, D. London, Butterworths.

FARRELL, D. J. (1975). A comparison of the energy metabolism of two breeds of hens and their crosses using respiration calorimetry. *British Poultry Science*, **16**: 103-113.

FAURE, J. M. AND FOLMER, J. C. (1975). Etude génétique de l'activité précoce en open-field de jeunes poussins. *Annales de Génétique et Sélection animale*, **7**: 123-132.

FILMER, D. G. (1974). Factors influencing food intake in practice. The laying fowl. In: *Proceedings of the 7th Nutrition Conference for Feed Manufacturers, Nottingham*, pp. 141-182. Edit. SWAN, H. AND LEWIS, D. London, J. and A. Churchill.

GLAZENER, E. W. AND BLOW, W. L. (1954). Variation in hybrids for efficiency of egg production. *Poultry Science*, **33**: 1055.

GOUS, R. M., BYERLY, T. C., THOMAS, O. P. AND KESSLER, J. W. (1978). A partition equation to predict food and energy intake by laying hens. *16th World's Poultry Congress, Malta*, 1-8.

HAGGER, C. (1977*a*). Untersuchungen zur Futterverwertung von Legehennen. 2. Genetisch parameter. *Archiv für Geflügelkunde*, **42**: 10-15.

HAGGER, C. (1977*b*). Untersuchungen zur Futterverwertung von Legehennen. 1. Phänotypische Beziehungen. *Archiv für Geflügelkunde*, **41**: 266-270.

HEIL, G. (1976). Schätzung der Genetischen Parameter für die Futterverwertung von Legehennen in Gruppenkäfingen und ihre Beziehung zu anderen Leistung- und Eiqualitätsmarkmalen. *Thesis, University of München.*

HURNIK, J. F., SUMMERS, J. D., WALKER, J. P. AND SZKOTNICKI, W. (1977). Production traits influencing the individual feed conversion ratio. *Poultry Science*, **56**: 912-917.

LECLERCQ, B., BLUM, J. C. AND BOYER, J. P. (1977). Significance and genetic control of body weight change in the hen during the laying period. *Archiv für Geflügelkunde*, **41**: 121-124.

LEE, C. Y. AND NORDSKOG, A. W. (1975). Value of feed consumption records to predict net income in layer-type chickens. *Poultry Science*, **54**: 1912-1918.

LEESON, S., LEWIS, D. AND SHRIMPTON, D. H. (1973). Multiple linear regression equations for the prediction of food intake in the laying fowl. *British Poultry Science*, **14**: 595-608.

LEESON, S. AND MORRISON, W. D. (1978). Effects of feather cover on feed efficiency in laying birds. *Poultry Science*, **57**: 1094-1096.

MCDONALD, D. (1978). Feed intake of laying hens. *World's Poultry Science Journal*, **34**: 209-221.

MERAT, P. (1968). Consommation alimentaire de pondeuses Cc et cc. *Annales de Zootechnie*, **17**: 337-339.

MERAT, P. AND BORDAS, A. (1972). Consommation alimentaire de pondeuses blanches (Ii) et colorées (ii). *Annales de Génétique et de Sélection animale*, **4**: 375-384.

MERAT, P. AND BORDAS, A. (1974). Consommation alimentaire d'animaux à plumage réduit (gène Na) ou normal en la présence ou l'absence du gène de nanisme dw. *Annales de Génétique et de Sélection animale*, **6**: 17-28.

MERAT, P. AND BORDAS, A. (1979). Effects associated with the pea-comb gene on chick weight and body weight and food efficiency of adult hens. *British Poultry Science*, **20**: 463-472.

MERAT, P., BORDAS, A. AND COQUERELLE, G. (1979). The relationship of several genes suppressing plumage colour with body weight, food intake and feather loss of laying hens. *British Poultry Science*, **20**: 587-594.

NORDSKOG, A. W., FRENCH, H. L., ARBOLEDA, C. R. AND CASEY, D. W. (1972). Breeding for efficiency of egg production. *World's Poultry Science Journal*, **29**: 175-188.

ORLOV, M. B. AND TUCHEMSKI, I. (1974). Food conversion by laying hens of a productive type over three generations. *15th World's Poultry Congress, New Orleans*, 165-166.

ROMANOFF, A. L. AND ROMANOFF, A. J. (1949). *The Avian Egg*. New York, Wiley and Sons.

STURKIE, P. D. (1965). *Avian Physiology*, 2nd ed. New York, Comstock.

SVENSON, S. A. (1964). Composition and energy content of eggs, growing chicks and hens, with some notes on preparation and methods of analysis. *Lantbrukshögskolans annaler*, **30**: 405-422.

TIERNEY, M. L. (1979). Selection for efficiency of production in Australorp hens. *Proceedings of the Inaugural Conference of the Australian Association of Animal Breeding and Genetics, Armidale*, 311-312.

WATANABE, K., YAMADA, Y. AND EBISAWA, S. (1975). Breeding for improvement of feed efficiency in egg layers. 1. Estimation of genetic parameters. *Japanese Poultry Science*, **12**: 219-229.

WATHES, C. M. (1976). The effect of feathering on heat loss of the laying fowl. *Gleadthorpe Poultry Booklet*, pp. 22-24. Mansfield, England, Gleathorpe Experimental Husbandry Farm.

Heterosis in Egg-Laying Lines Under Divergent Selection for Residual Feed Consumption

A. BORDAS, P. MÉRAT, and F. MINVIELLE[1]

Institut National de la Recherche Agronomique Laboratoire de Génétique Factorielle, 78352 Jouy en Josas Cedex, France

ABSTRACT Two lines selected since 1976 for high (R^+) or low (R^-) residual feed consumption (RFC) from a common genetic base were compared with one another and with their F_1 reciprocal crosses for traits of egg production and quality, for morphological traits, body weights, and feed consumption. Heterosis was 11, –2.5, 8, and 2%, respectively, for egg number, age at first egg, egg laying rate, and egg weight, with marked differences between reciprocal crosses for all those traits but egg number. Heterosis for wattle length and shank length was 3.8 and 1.3%, respectively, essentially because $R^+ \times R^-$ crossbreds, with larger mean values, resembled the R^+ line for those traits, which may therefore be associated with the presence of genes linked to the Z chromosomes. On the other hand, heterosis for RFC (–3.6%) originated from similar crossbred advantage in both reciprocal crosses, thereby suggesting that RFC is not determined by sex-linked genes.

(*Key words*: residual feed consumption, feed intake, egg production, morphological traits, heterosis)

1996 Poultry Science 75:20–24

INTRODUCTION

In egg lines, feed efficiency depends mainly on body size and egg production of the hen. However, once individual differences of body weight, body weight changes, and egg production have been accounted for, some variation in feed efficiency still remains between birds. This variation may be characterized by residual feed consumption (RFC) (e.g., Byerly *et al.*, 1980). Differences in RFC clearly have a genetic basis, as a divergent selection on this trait, undertaken in 1976 in our laboratory, had led to a mean difference of about three phenotypic standard deviations in 1989 (Bordas *et al.*, 1992), and as significant genetic progress has also been obtained in other selection experiments on RFC (Katle and Kolstad, 1991; Schulman *et al.*, 1994). Moreover, heritability of RFC has been estimated by several workers, with recent reports of values between 0.4 and 0.5 (Luiting and Urff, 1991; Schulman *et al.*, 1994), but with earlier reports between 0.1 and 0.5 (Bentsen, 1983; Fairfull and Chambers, 1984; Wing and Nordskog, 1982). To investigate RFC further from a genetic standpoint, crossing high and low RFC lines remained to be done, as preliminary unpublished work with crosses after seven generations of divergent selection had yielded about 3% heterosis for RFC. The objective of the present work was to evaluate heterosis for RFC, feed consumption, and production traits through reciprocal crosses of the high and low RFC lines, after 16 generations of selection.

Received for publication April 18, 1995.
Accepted for publication October 10, 1995.
[1]To whom correspondence should be addressed.

MATERIALS AND METHODS

Lines and Crosses

Lines selected for low (R^-) and high (R^+) values of RFC were set up in 1975 from a Rhode Island Red population. Individual selection for RFC led to a large divergence between the two lines for both sexes, and to correlated responses in feed efficiency and in the size of several appendages (Bordas and Mérat, 1984; Bordas *et al.*, 1992).

In 1992, 9 cocks and 54 hens were sampled from all paternal families (1 cock and 6 hens per family), for each one of the two pure lines, R^- and R^+. Each cock was used to inseminate three half sisters from another family of its own line and three half sisters from the other line. Therefore, pure and crossbred experimental birds were sired by the same sets of R^+ or R^- cocks. Reproduction was by artificial insemination and birds were obtained in a single hatch from a 2-wk egg collection. Female chicks were group reared in a single unit under standard conditions. At 17 wk of age, 40 pullets per group were sampled about equally from the nine or eight paternal families (progeny from one cock was not used, as the mean RFC of his purebred progeny was markedly different from those of the other pure line RFC groups). They were transferred at random into individual cages in a single room (14 h light:10 h dark) with a temperature of 21 ± 2 C and were provided *ad libitum* access to a

commercial egg layer diet (15.5% total protein, 2600 kcal ME/kg, 3.4% calcium).

Measurements

The list of variables studied is shown in Table 1. Egg production data were collected up to the age of 43 wk. Egg laying rate was the average number of eggs laid by a hen since its first egg, expressed as a percentage. Clutch length was the number of consecutive days with an oviposition. A pause was an absence of oviposition for at least 2 d. The percentage days in pauses was calculated for each hen as the proportion of days of pauses in the total number of days on test since its first oviposition. Average egg weight was estimated from a 2-wk egg collection started at 31 wk of age. During the same period, egg albumen height and shell thickness were recorded by using two eggs per hen. Percentages of soft-shelled eggs, of broken eggs, and of eggs with double yolks (identified by visual inspection only) were calculated over the whole test period. Data associated with feed consumption were recorded individually for 28 d between 29 and 33 wk of age (after the peak of egg production). Body weight was the mean body weight during the feed consumption test period. Change of body weight (ΔW) between 29 and 33 wk of age, total egg mass (EM) produced in that period, along with total feed intake (FI) at this time were obtained also. The RFC was defined as the difference between FI and expected feed intake (EFI) estimated from a single multiple linear regression of FI on BW, ΔW, and EM (Byerly *et al.*, 1980). The regression equation adjusted to the data was: $EFI = 93.8\ BW^{0.5} + 1.50\ \Delta W + 1.02\ EM - 2250$, with $R^2 = 0.54$.

Statistical Analysis

Six R^- line hens and two R^+ line hens died during the test and corresponding data were discarded from the statistical analysis. Variables expressed as percentage were transformed with the arc sine square root transformation without changing significance levels of the effects. Therefore, only results on untransformed data are given.

For each variable, a one-way analysis of variance was performed to compare genetic types (R^-, R^+, $R^- \times R^+$, and $R^+ \times R^-$). Heterosis was estimated, overall, as percentage difference between crossbreds and average of purebreds and, for each reciprocal cross, from appropriate linear combinations of least squares means for genetic types. Significance of heterosis was evaluated by the Student's test with n – r degrees of freedom, with n being the total number of observations (152) and r the number of groups (4). All analyses were performed by using the General Linear Models procedure (SAS Institute, 1988).

RESULTS AND DISCUSSION

Means for all traits, with corresponding significance level for differences between genetic types are given in Table 1. Significant differences were obtained for all egg production traits but double-yolk eggs, for all three morphological characteristics and for all variables associated with feed consumption except BW and ΔW. Of course, among those, differences obtained for traits under selection (RFC) or associated with selection (FI, wattle length) resulted from the large divergence obtained by selection (Bordas *et al.*, 1992). Correspondingly, means of crossbred groups for these traits were between the extremes achieved in pure lines. On the other hand, significant differences between genetic types were also found for other traits, like egg production measurements, not affected by selection on RFC but showing the classical advantage of crossbreds over purebreds (Fairfull, 1990).

Effects of heterosis, overall and for each reciprocal cross, and associated significance levels are given for all variables in Table 2. Heterosis was present for most traits not associated with divergent selection. Indeed, percentage heterosis for egg production traits, early body weights, and egg quality traits, with values equal to 11% for egg number, 2.5% for age at first egg, 8% for egg laying rate, 2% for egg weight, and about 4% for 8- and 17-wk body weights, were well within the range of those compiled by Fairfull (1990). Heterosis was also found for traits not usually reported in the literature, like clutch length, days of pauses, and proportions of soft-shelled eggs and broken eggs.

Both lines had the same genetic base (Bordas and Mérat, 1984) and divergent selection had not induced any significant correlated responses for egg production and egg quality traits and BW (Bordas *et al.*, 1992). Therefore, R^+ and R^- lines should have remained rather similar genetically, concerning this set of characters, with, correspondingly, little heterosis in the F_1. However, significant heterosis was obtained, which indicates that divergence through single trait selection may have rapidly modified the genome in a way that did not show phenotypically on unselected traits in pure lines, but which was revealed through heterosis when crossing those lines.

Among the traits associated with the selective process, either through a correlated response, as in the case of wattle length (Bordas *et al.*, 1992), or through the definition of RFC, as in the case of EM, heterosis was small but it was significant for wattle length, shank length, and feed conversion. The significant value obtained for RFC itself (–3.6% of the average FI) is quite similar to the 3% heterosis for RFC estimated from a limited number of crosses (unpublished data) in Generation 8 of this selection experiment. This observation tends to confirm previous reports that in poultry lines selected for a single trait, heterosis for that trait was little changed as selection proceeded (Cole and Hutt, 1973; Gowe *et al.*, 1973). A similar value of heterosis, consistent with recent reports (e.g., Singh *et al.*, 1993), was obtained for feed conversion. It resulted from the marked advantage of the $R^+ \times R^-$ cross for egg mass, as no heterosis was found for FI, the numerator part of the feed conversion variable. The difference between FI and RFC for heterosis indicates that the two traits are under

TABLE 1. Means ± SEM of egg production traits, body weights, morphological traits, and feed consumption resulting from pureline and reciprocal cross progeny

Trait	Parental combinations[1]				Significance of effects
	R⁻ × R⁻	R⁻ × R⁺	R⁺ × R⁻	R⁺ × R⁺	
Number tested	34	40	40	38	
8-wk BW, g	712.8 ± 9.9	741.7 ± 9.6	715.2 ± 10.8	687.8 ± 14.4	*
17-wk BW, g	1,665.5 ± 24.7	1,760.0 ± 14.2	1,713.0 ± 17.2	1,688.3 ± 22.7	**
Age at first egg, d	150.0 ± 1.3	145.8 ± 0.9	149.1 ± 1.0	152.6 ± 1.1	***
Egg number	109.6 ± 4.5	121.6 ± 2.8	125.6 ± 2.1	112.3 ± 4.4	**
Egg laying rate, %	72.6 ± 2.9	78.9 ± 1.7	83.2 ± 1.2	76.2 ± 2.9	**
Clutch length, d	4.3 ± 0.3	5.2 ± 0.3	6.4 ± 0.5	5.3 ± 0.4	**
Days of pauses, %	12.3 ± 3.1	8.3 ± 1.6	4.7 ± 0.8	13.7 ± 3.1	*
Egg weight, g	52.5 ± 0.5	52.7 ± 0.4	52.4 ± 0.5	50.4 ± 0.5	**
Albumen height, 0.01 mm	56.9 ± 1.2	57.0 ± 1.3	51.5 ± 1.0	42.1 ± 1.3	***
Shell thickness, 0.01 mm	32.9 ± 0.4	33.0 ± 0.3	32.9 ± 0.3	31.7 ± 0.3	**
Soft-shelled eggs, %	5.4 ± 1.5	5.2 ± 0.7	3.2 ± 0.4	7.7 ± 1.5	*
Double-yolk egg number	1.7 ± 0.3	2.8 ± 0.5	3.0 ± 0.4	2.7 ± 0.6	
Broken eggs, %	21.7 ± 2.7	22.3 ± 2.6	13.5 ± 1.5	28.9 ± 2.7	***
Morphological traits at 43 wk					
Wattle length, mm	16.9 ± 0.7	23.6 ± 0.5	24.6 ± 0.6	28.5 ± 0.6	***
Shank length, mm	102.7 ± 0.5	106.6 ± 0.5	107.5 ± 0.7	108.7 ± 0.4	***
Shank diameter, mm	9.15 ± 0.07	9.27 ± 0.06	9.28 ± 0.06	9.14 ± 0.06	
Traits associated with feed intake (28-d test)					
BW, g	1,987.0 ± 36.7	2,041.5 ± 34.9	2,052.1 ± 29.4	2,003.8 ± 34.3	
Change of BW (ΔW), g	21.2 ± 10.3	32.0 ± 11.3	13.2 ± 9.9	19.7 ± 10.4	
Egg mass (EM), g	1,150.7 ± 26.6	1,181.2 ± 36.3	1,271.5 ± 20.6	1,199.0 ± 30.0	*
Feed intake (FI), g	2,786.7 ± 59.3	3,176.5 ± 65.0	3,252.5 ± 61.1	3,656.7 ± 67.3	***
Residual feed consumption (RFC), g	−329.7 ± 23.6	−60.2 ± 27.3	−61.2 ± 33.3	+437.2 ± 39.7	***
Feed conversion (FI/EM)	2.44 ± 0.05	2.69 ± 0.07	2.56 ± 0.03	3.09 ± 0.07	***

[1]R⁻ line selected for low RFC; R⁺ line selected for high RFC.

*$P < 0.05$.

**$P < 0.01$.

***$P < 0.001$.

TABLE 2. Crossbred superiority (percentage) evaluated for combined (heterosis) and separate reciprocal crosses of R⁻ and R⁺ lines

Trait	Overall	Mating source[1]	
		R⁻ × R⁺	R⁺ × R⁻
8-wk BW	4.0**	6.0**	2.2
17-wk BW	3.5**	4.9***	2.2
Age at first egg	-2.5***	-3.6***	-1.5
Egg number	11.4***	9.6*	13.8***
Egg laying rate	8.2**	6.0	11.8***
Clutch length	22.0**	9.1	35.0***
Days of pauses	-50.0**	-26.0	-63.8**
Egg weight	2.1*	2.5*	1.6
Albumen height	4.3	9.6***	-1.0
Shell thickness	2.0*	2.1*	2.0
Soft-shelled eggs	-36.6*	-10.3	-52.0**
Double-yolk eggs	22.0	21.0	23.0
Broken eggs	-29.2**	-11.9	-46.6***
Morphological traits at 43 wk			
Wattle length	3.8*	0.8	5.0**
Shank length	1.3*	0.9	1.6**
Shank diameter	1.4*	1.3	1.4
Traits associated with feed intake (28-d test)			
BW	2.6	2.3	2.8
Change of BW (ΔW)	11.0	56.2	-35.2
Egg mass (EM)	4.4	0.5	8.2**
Feed intake (FI)	-0.2	-1.4	1.0
Residual feed consumption (RFC),[2]	-3.6***	-3.5**	-3.7**
Feed conversion (FI/EM)	-3.7**	-2.5	-7.2*

[1]R⁻ line selected for low RFC; R⁺ line selected for high RFC.

[2]Expressed as percentage of average feed intake.

*$P < 0.05$.

**$P < 0.01$.

***$P < 0.001$.

somewhat different genetic control, despite a 0.4 genetic correlation (Tixier-Boichard *et al.*, 1995). Heterosis for RFC was also comparable to the value (-54.2 g that is -3.1% of average feed intake) found for Japanese quail at generation 9 of a selection experiment for egg production (Minvielle *et al.*, 1995).

Heterosis resulted from superiority for one of the two reciprocal crosses for most traits, exceptions being egg number and RFC. Indeed, R⁻ × R⁺ crossbreds were superior for characters of early growth and precocity and for egg quality, whereas R⁺ × R⁻ reciprocals were better for traits involved with improved body heat dissipation (wattle and shank lengths) and regularity of oviposition. Although some reciprocal differences could be expected that would result from sex-linked and maternal effects (Fairfull, 1990), interpretation of the varied differences between superiority of R⁻ × R⁺ and R⁺ × R⁻ crosses found in this work is not straightforward. However, for wattle length and shank length, higher figures for R⁺ × R⁻ females correspond to morphological characteristics of the R⁺ line and then, possibly, to the presence of specific alleles on the Z chromosome in this line. On the other hand, regularity of oviposition of R⁺ × R⁻ hens, as indicated by clutch length and pauses, corresponds more to characteristics of the R⁻ line, which would tend to implicate, for these traits, maternal effects or W-linked genes.

Finally the absence of reciprocal effects on heterosis for RFC suggests that genetic variation for this trait is not importantly affected by genes on sex-chromosomes nor by maternal effects.

ACKNOWLEDGMENTS

The technical assistance of B. Rivet at the poultry experimental unit, La Minière, is gratefully acknowledged. This work was supported by a grant from the INRA AIP program Appétit.

REFERENCES

Bentsen, H. B., 1983. Genetic variation in feed efficiency of laying hens at constant body weight and egg production. II- Sources of variation in feed consumption. Acta Agric. Scand. 33:305–320.

Bordas, A., and P. Mérat, 1984. Correlated responses in a selection experiment on residual feed intake of adult Rhode Island Red cocks and hens. Annales Agriculturae Fenniae 23:233–237.

Bordas, A., M. Tixier-Boichard, and P. Mérat, 1992. Direct and correlated responses to divergent selection for residual food intake in Rhode-Island Red laying hens. Br. Poult. Sci. 33:741–754.

Byerly, T. C., J. W. Kessler, R. M. Gous, and O. P. Thomas, 1980. Feed requirement for egg production. Poultry Sci. 59: 2500–2507.

Cole, R. K., and F. B. Hutt, 1973. Selection and heterosis in Cornell White Leghorns: a review, with special consideration of interstrain hybrids. Anim. Br. Abstr. 41:103–118.

Fairfull, R. W., 1990. Heterosis. Pages 913–933 *in:* Poultry Breeding and Genetics. R. D. Crawford, ed. Elsevier, Amsterdam, The Netherlands.

Fairfull, R. W., and J. R. Chambers, 1984. Breeding for feed efficiency. Can. J. Anim. Sci. 64:513–527.

Gowe, R. S., W. E. Lentz, and J. H. Strain, 1973. Long-term selection for egg production in several strains of White Leghorns: performance of selected and control strains including genetic parameters of two control strains. Pages 225–245 *in:* Proceedings 4th European Poultry Conference, London, UK.

Katle, J., and N. Kolstad, 1991. Selection for efficiency of food utilization in laying hens: direct response in residual food consumption and correlated responses in weight gain, egg production and body weight. Br. Poult. Sci. 32:939–953.

Luiting, P., and E. M. Urff, 1991. Residual feed consumption in laying hens. 2. Genetic variation and correlations. Poultry Sci. 70:1663–1672.

Minvielle, F., J. L. Monvoisin, A. Frénot, and J. Costa, 1995. Heterosis for early egg production and feed intake in Japanese quail. Arch. Geflügelkd. 59:89–93.

SAS Institute, 1988. SAS® User's Guide: Statistics. SAS Institute Inc., Cary, NC.

Schulman, N., M. Tuiskula-Haavisto, L. Sütonen, and E. A. Mäntysaari, 1994. Genetic variation of residual feed consumption in a selected Finnish egg-layer population. Poultry Sci. 73:1479–1484.

Singh, S. S., S. K. Verma, and A. K. Shrivastava, 1993. Evaluation of gene effects for efficiency of feed utilization for egg mass production in laying type chicken. Poult. Abstr. 19(2):47–48.

Tixier-Boichard, M., D. Boichard, E. Groeneveld, and A. Bordas, 1995. Restricted maximum likelihood estimates of genetic parameters of adult male and female Rhode Island Red chickens divergently selected for residual feed consumption. Poultry Sci. 74:1245–1252.

Wing, T. L., and A. W. Nordskog, 1982. Use of individual feed records in a selection program for egg production efficiency. I- Heritability of the residual component of feed efficiency. Poultry Sci. 61:226–230.

Asian-Aust. J. Anim. Sci.
Vol. 21, No. 5 : 624 - 627
May 2008

www.ajas.info

Isolation and Characterization of Microsatellite Markers in Tsaiya Duck

M. C. Hsiao, H. C. Liu*, Y. C. Hsu[1], Y. H. Hu, S. H. Li[2] and S. R. Lee

Ilan Branch, Livestock Research Institute, Council of Agriculture, Executive Yuan, Yilan 268 Taiwan

ABSTRACT : An enrichment library of GATA-repeats from genomic DNA was constructed in this study to isolate and characterize microsatellite loci in Tsaiya duck (*Anas platyrhynchos*). Thirty-three microsatellite markers were developed and used to detect polymorphisms in 30 Tsaiya ducks. A total of 177 alleles were observed and all loci except APT022 were polymorphic. The number of alleles ranged from 2 to 9 with an average of 5.5 per microsatellite locus. The observed and expected heterozygosity of these polymorphic markers ranged from 0.07 to 0.93 with an average number of 0.60 and 0.10 to 0.86 with an average number of 0.61, respectively. Among the polymorphic markers, the observed heterozygosities of 23 loci were higher than 0.50 (69.70%). The polymorphism information content (PIC) in the 32 loci ranged from 0.09 to 0.83 with an average of 0.57. Seven of the 33 duck microsatellite loci had orthologs in the chicken genome, but only APT004 had a similar core repeat to chickens. These microsatellite markers will be useful in constructing a genetic linkage map for the duck and a comparative mapping with the chicken can also provide a valuable tool for studies related to biodiversity and population genetics in this duck species. (**Key Words :** Microsatellite Marker, Polymorphism, Tsaiya Duck)

INTRODUCTION

Many important agricultural traits are quantitative traits controlled by multiple genes. The recent development of molecular genetic mapping tools has enabled the identification of quantitative trait loci (QTL) in the genome. The application of marker assisted selection for QTL has the potential to enhance the accuracy of animal breeding programs, particularly for traits that are difficult to improve through traditional selection methods (Meuwissen and Goddard, 1996). Microsatellites, also known as short tandem repeats (STR), are tandem repeated motifs of 1-6 bases. They are found abundantly and at random throughout most eukaryotic genomes (Stallings et al., 1991). Microsatellites are highly polymorphic and have become one of the most useful tools for population genetic studies, linkage mapping, parentage determination and QTL analysis. In chicken, swine and cattle populations a large number of microsatellites have been isolated and widely used for these purposes (Kong et al., 2006; Liu et al., 2006). In contrast, fewer genetic markers have been established in the duck and only a few articles have been published for some species including the Peking duck, eider duck, muscovy duck, mallard, white-headed duck, ruddy duck and musk duck (Maak et al., 2000; Maak et al., 2003; Paulus and Tiedemann, 2003; Stai and Hughes, 2003; Denk et al., 2004; Munoz-Fuentes et al., 2005; Guay and Mulder, 2005; Huang et al., 2005; Huang et al., 2006). Although the first duck genetic linkage map has been developed (Huang et al., 2006), it only spans 1,353.3 cM with an average interval distance of 15.04 cM. More microsatellites are needed to establish a complete duck genetic map. We attempted to isolate microsatellite markers for the Tsaiya duck (*Anas platyrhynchos*) and investigated its polymorphisms.

MATERIALS AND METHODS

Collection and extraction of Tsaiya duck genomic DNA

Thirty individuals (15 males and 15 females) were selected from a germplasm preservation population of Tsaiya ducks kept at the Ilan Branch since 1984. In the beginning, the population was reproduced via natural mating. At the 8th generation, we randomly divided this germplasm preservation duckling population into 15-sire families. The rotational crossbreeding system was then used

* Corresponding Author: H. C. Liu. Tel: +886-3-9503107, Fax: +886-3-9501950, E-mail: scliu@mail.tlri.gov.tw

[1]Institute of Natural Resources, National Dong Hwa University, Hualien 974 Taiwan.

[2] Department of Biology, National Taiwan Normal University, Taipei 116 Taiwan.

Received June 29, 2007; Accepted November 9, 2007

to avoid inbreeding. We randomly chose one male and one female as samples from each sire family from the 9th generation. Genomic DNA was extracted from fresh blood using the GenoMaker kit (Watson BioTech, Taiwan) following the manufacturer's instructions. Briefly, 40 μl blood was lysed in 1 ml GenoMaker reagent and extracted with chloroform. DNA was precipitated with isopropanol and quantified using a spectrophotometer.

Construction of genomic DNA libraries enriched for microsatellites

The library was enriched for GATA repeats following a combination of modified procedures according to Hamilton et al. (1999) and Hsu et al. (2003). A pooled genomic DNA of 3 Tsaiya ducks was digested with *Alu*I, *Hae*III and *Rsa*I. The fragments were then ligated with SNX linkers (Hamilton et al., 1999). The ligated products were amplified using PCR at 94°C for 5 min, 30 cycles of 94°C for 1 min,

Table 1. Characteristics of 33 novel microsatellite loci in the Tsaiya duck

Locus	Repeat motif in clone	Primer sequences (5'-3')	Size from clone (bp)	T_a (°C)	$MgCl_2$ (mM)	Fragment (bp)	No. of alleles	H_O	H_E	PIC	GenBank accession no.
APT001	$(GATA)_{12}$	F: GTC CCA CTG GTT TGC TGT CC R: ACT ACG CAT GGC AGT GAG GTT	206	55	2.5	178-206	3	0.23	0.52	0.46	DQ884881
APT002	$(GATA)_7GACA(GATA)_3$	F: ACC CTC CCA CAG ATT AAA GAG AAG T R: GGA AGG ATG CCC TGA TTT ACA C	133	55	2.5	129-145	5	0.70	0.66	0.61	DQ884882
APT003	$(GATA)_{11}$	F: GAT CAT TGC ACT TGA AAT TAT TGT TAT TT R: TGT GCA TTA CTG TGG CAG ATC TG	228	55	2.5	220-236	4	0.57	0.54	0.49	DQ884883
APT004	$GATAGAT(GATA)_{15}$	F: GGG CAG GAA AAT CTC CTG AAT R: TCT CAG TGG CTG AGC GGT C	306	55	2.5	294-322	9	0.80	0.86	0.83	DQ884884
APT005	$(GATA)_{17}$	F: TCC GTA CAG ACC AAC ATC GG R: AGG TCT TTA CAG CCC ACT CCC	311	55	2.5	283-319	9	0.83	0.80	0.77	DQ884885
APT006	$(GATA)_{12}$	F: CTT CCC ATT GCA GTG TTG GTC R: TTG GCA TCT TTG TTC TGC AGA	326	50	2.5	318-342	6	0.57	0.64	0.59	DQ884886
APT007	$(GATA)_{14}$	F: TCT TAA ATG TTA GTG ATA CCA GCA TCT TT R: TTC GGG ATG AGA AGG AAG GA	218	50	2.5	194-230	6	0.80	0.72	0.67	DQ884887
APT008	$(GATA)_{12}$	F: CAA AGA AAT CCT AGA ACA TCA TTC AAA T R: TCT TCT GGC TTT TCA CCT TAG TTT AGT A	188	50	2.5	184-208	7	0.87	0.79	0.75	DQ884888
APT009	$(GATA)_2GAT(GATA)_{15}$	F: CCA GGC AGT TGC TGT GTA ACA R: GGC GCT TTC TTC TAT GAT CGA	334	55	2.5	330-354	7	0.50	0.81	0.77	DQ884889
APT010	$(GATA)_9GAT(GATA)_3$	F: CAC TCA GGC TTT TAG GTC CAT TAA TA R: CAT CTG AGA ATG CAC TTA CTG TCA AA	204	55	2.5	192-215	6	0.70	0.63	0.56	DQ884890
APT011	$(GATA)_8$	F: CAT ACA GGG AGT CTG AGA TGA TCA A R: TTA TGT TCC ATT CAG GGC TTT CTT	160	50	2.5	152-164	3	0.07	0.13	0.12	DQ884891
APT012	$(GATA)_{16}$	F: TTG AGC CTC AGG TTC TAA ACT CCT A R: TCA TAA CAT TTC AGA CCA GTT TTC AGA	205	50	2.5	185-205	6	0.87	0.73	0.67	DQ884892
APT013	$(GATA)_{10}$	F: CCA ACC ACC AGG AAG TAC TGT AAA TA R: AGG AAA GTT CAG ACA CAT GGA TTG	131	50	2.5	127-171	9	0.77	0.81	0.78	DQ884893
APT014	$(GATA)_{11}$	F: GCA CCA GGT AAT TTA TGT CAG AAA TAA T R: GAA GTG CAA AAC ATG GTT CAG G	321	50	2.5	317-325	3	0.10	0.16	0.15	DQ884894
APT015	$(GATA)_{13}$	F: CTG TTA TGA CAC CAT GTT TGG ATT TA R: CGT GCT CTG CAA CAA CTG AAA	138	55	2.5	126-150	7	0.77	0.75	0.71	DQ884895
APT016	$(GATA)_{10}$	F: TCT TAA ATG GGA CTG ATG GAG AGA G R: ACC TAT TTT ATC TCA GGA TGC AAT TAT G	112	50	2.5	112-120	3	0.53	0.54	0.45	DQ884896
APT017	$(GGAT)_6(GATA)_{12}$	F: TGG ATG GAC AGA CGG GTG A R: TGG AAG TTT TGA TTT CTA GTG CTT ACA	185	55	2.5	161-189	5	0.77	0.74	0.69	DQ884897
APT018	$(GATA)_9(GAAA)_{14}(GA)_2(GAAA)_2(GA)_6(GAAA)_2$	F: GTG GCA GTT TAA TGA AAG CGA AA R: TGG AGG TAC CCA AAG GAG AAT TC	267	50	2.5	267-295	8	0.20	0.86	0.83	DQ884898
APT019	$(GATA)_{11}$	F: CCA AGA TTA GGG CTA TGT GGT GAT R: AAG GAT TGA GAC AGG AGA TGG G	218	50	2.5	206-218	2	0.37	0.35	0.28	DQ884899
APT020	$(GATA)_{14}$	F: TTC CAA GTT TGT CAT GCC AAT AGA R: CTG ACC ATG TTA GGG CGT TTT AG	201	50	2.5	177-205	8	0.93	0.83	0.79	DQ884900
APT021	$(GATA)_{10}$	F: GCA CTC CCT AAC TAG TAG CGC TCT R: GAA GCA TTG TCA TAC TTG CCT GA	133	50	2.5	133-169	6	0.80	0.79	0.75	DQ884901
APT022	$(GATA)_{12}$	F: TCA GTG AAA GCC ACA GTC AGA TC R: TTT AGG CAC TGA AGC CCA ACA	120	50	2.5	120	1	0.00	0.00	0.00	DQ884902
APT023	$(GATA)_4AATA(GATA)_7$	F: CCA AAC AAG AGA AGA TGA TAG AGA GAC A R: GAA TCA ATA AAC TGC TTT GAT CCT GAC	117	50	2.5	113-121	3	0.57	0.51	0.43	DQ884903
APT024	$GATAGACA(GATA)_7$	F: TGT GGG CAG TTC CTC AAC AA R: GCC CAC CCT CTC TTT CTG AAG	102	55	2.5	102-118	4	0.50	0.51	0.44	DQ884904
APT025	$(GATA)_{13}$	F: TCC TAA GAA ACG TTG CTT CAT AGA CC R: GAG TTA AGC TTC ATC ACT CTG TGA CTG	121	50	3	105-133	7	0.70	0.67	0.63	DQ884905
APT026	$(GATA)_{10}$	F: CCC TGA AAG GCT GTT TTA TAT ATC CA R: ATG TAA ATA AAG TAG CCT TGC ACG GT	138	55	2.5	138-142	2	0.60	0.51	0.38	DQ884906
APT027	$(GATA)_{10}$	F: ATT TCC AAA ATC TTG TGC TTT AAG C R: TTT TTG TTC TTT CTC TCT CTC CCT CT	151	50	3	151-155	2	0.10	0.10	0.09	DQ884907
APT028	$(GATA)_{10}$	F: CAT TCA TGT TTA TTT CTT CTG GTA TGT G R: GTT AAA ATG GGA AGG CTT CAC TAG A	167	55	2.5	131-187	7	0.83	0.82	0.78	DQ884908
APT029	$(GATA)_{14}$	F: TCT GCA AGG TAT TCT CAT TCT TAT TCT T R: GAT ACG TAG AGT GGA TGC TGG AGA T	171	50	3	143-179	6	0.73	0.66	0.60	DQ884909
APT030	$(GATA)_{13}$	F: TGG ATA TAC CAT GCC AGT GCA R: TGG CTT GTG GGA GAG ATG ATG	206	50	2.5	190-226	8	0.67	0.54	0.52	DQ884910
APT031	$(GATA)_{12}$	F: GCT GGA AGA AAG GAG AAG GAG G R: AGA AAA ACA GTA TGA GCG AAC AGG T	210	50	2.5	194-234	8	0.90	0.84	0.81	DQ884911
APT032	$GATA(GACA)_2GACT(GATA)_4$	F: TCA CTT TCT TGA CTC TCC TTG GTT T R: TGA CTT GAA TTC TGT TCA GGA TAA ATG	259	55	2.5	207-259	4	0.70	0.67	0.60	DQ884912
APT033	$(GATA)_{13}$	F: CTT CAC CCT ACC TCA TAA GGA ACT G R: ATT CCA AAT CTG CAA GGT GAG TAT TA	266	55	2.5	266-274	3	0.13	0.19	0.17	DQ884913

T_a = Annealing temperature, H_O = Observed heterozygosity, H_E = Expected heterozygosity, PIC = Polymorphism information content.

Table 2. The orthologous microsatellite duck DNA in the chicken genome

Locus	Score (bits)	E value	Chromosome*	Physical position[1]	Chicken repeat	Duck repeat
APT004	154	2e-35	3	6514982-6515130	GATAAATA$(GATA)_{10}(CATA)_6$	GATAGAT$(GATA)_{15}$
APT005	119	5e-25	1	8485714-8485875	T_{21}	$(TATC)_{17}$
APT006	117	1e-24	4	27975792-27975899	No	$(GATA)_{12}$
APT012	196	3e-48	2	185809-185999	A_{13}	$(GATA)_{16}$
APT014	106	4e-21	un	5986-6145	No	$(GATA)_{11}$
APT016	167	1e-39	4	8793108-8793307	No	$(GATA)_{10}$
APT031	204	1e-50	4	77150-77368	No	$(GATA)_{12}$

* Chromosome is the pair of chromosome in the chicken. 'un' means that the position in the chicken genome is uncertain.
[1] The physical position in the chicken genome.

55°C for 1 min and 72°C for 2 min, followed by a final cycle at 72°C for 7 min. The amplified products were used for subtractive hybridization with 3'-biotinylated $(GATA)_{10}$ oligonucleotides to select the microsatellite-containing DNA fragments. The biotin-labeled oligonucleotides were eluted using Dynabeads MyOne Streptavidin (Dynal, Norway) according to the manufacturer's protocol. Repeat-enriched DNA was made double-stranded and amplified under the above described PCR conditions and then cloned in the pGEM-T Easy Vector (Promega, USA). After transformation into JM109 competent cells, 800 colonies containing inserts were lifted to Nylon membranes (Roche, Germany) and hybridized with 3'-DIG-labeled $(GATA)_8$ oligonucleotides. The positive colonies were cultured to extract their plasmids and then sequenced using the BigDye Terminator Kit on a 3730xl DNA Analyzer (Applied Biosystems).

Genotyping

Sequences were aligned with SeqWeb Version 2.1 (Wisconsin Package). The Primer Express software (Applied Biosystems) was used to design PCR primers. The primer pairs showing the expected PCR products were selected for polymorphism screening. The forward primers of these primer pairs were labeled with FAM or HEX fluorescent dye. PCR reaction mixtures (10 μl) containing 10 ng genomic DNA, 2.5 mM or 3 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl pH 8.3, 200 μM of each dNTP, 0.4 μM of forward and reverse primer and 0.25 U AmpliTaq Gold DNA polymerase (Applied Biosystems/Roche) were prepared. After initial incubation at 95°C for 10 min, each PCR amplification was performed for 30 cycles of denaturing at 95°C for 20 s, annealing for 30 s at an appropriate annealing temperature (50 or 55°C) for each primer pair and extension at 72°C for 30 s. This was followed by a final cycle at 72°C for 1 h. The PCR products were analyzed in a MegaBACE 1000 auto sequencer (Amersham Biosciences). The sizes of the DNA fragments were investigated using the Genetic Profiler Version 2.2 software (Amersham Biosciences).

Statistics and similarity searching

The observed and expected heterozygosities and polymorphism information content (PIC) were calculated using the CERVUS 2.0 program (Marshall et al., 1998). Hardy-Weinberg expectation and linkage disequilibrium tests were performed with the FSTAT 2.9.3.2 software (http://www2.unil.ch/popgen/softwares/fstat.htm). The sequences were analyzed using the BLAST program (NCBI) to identify the orthologous microsatellite duck DNA in the chicken genome. The unique match sequences with E-values smaller than e-20 from the chicken were used as orthologs to the duck microsatellite DNA.

RESULTS AND DISCUSSION

Eighty positive clones out of 800 colonies screened from the GATA-enriched genomic library were sequenced. There were 47 different loci from 75 sequences containing GATA repeats and 33 of these loci were chosen for further polymorphism tests in 30 Tsaiya ducks. After searching further with the BLAST program, we found that these 33 microsatellite Tsaiya duck loci were novel. The characteristics of the 33 microsatellite loci are summarized in Table 1. A total of 177 alleles were observed and all loci except APT022 were polymorphic. The number of alleles ranged from 2 to 9 with an average of 5.5 per microsatellite locus. The observed and expected heterozygosity of these polymorphic markers ranged from 0.07 to 0.93, with an average number of 0.60, and 0.10 to 0.86, with an average number of 0.61, respectively. Among the polymorphic markers, the observed heterozygosities of 23 loci were higher than 0.50 (69.70%). The polymorphism information content (PIC) of the 32 loci ranged from 0.09 to 0.83, with an average number of 0.57. Based on the classification of Botstein et al. (1980), twenty-one (65.63%) polymorphic markers were highly informative (PIC>0.50), seven (21.88%) were reasonably informative (0.50>PIC>0.25), and four (12.50%) were slightly informative (PIC<0.25). No locus showed significant deviation from Hardy-Weinberg equilibrium and there was no linkage disequilibrium among the loci ($p>0.05$). The orthologous microsatellite duck DNA in the chicken genome search results are presented in Table 2. Only APT004 out of the seven duck microsatellite loci with orthologs in the chicken genome had a similar core repeat. The others had different

core repeats or were absent from the orthologous loci in the chicken genome.

Values of 94.44% (Maak et al., 2003), 80.00% (Huang et al., 2005) and 77.89% (Huang et al., 2006) polymorphisms have been reported for duck-specific microsatellite markers tested in the duck genome. The higher (96.97%) polymorphism seen in this study could be a reflection of the test population genetic constitution, which was derived from a germplasm preservation population without artificial selection. Based on the PIC values, most polymorphic markers were highly or reasonably informative and only a few were slightly informative. Therefore, these markers will be very useful for mapping the duck genome. As the Chicken Genome Project moves toward functional genomics studies, the availability of the chicken genome sequence has proven to be an invaluable tool in studying the genomes of other avian species, including the duck. A good comparative genetic map based on the orthologous microsatellite markers will provide the substrates for major gene identification (Reed et al., 2005). After a similar BLAST search, 21.21% of the microsatellite loci were conserved between the duck and the chicken. This result was similar to a previous report (20.42%, Huang et al., 2006). In conclusion, these microsatellite markers will be useful in construction of a duck genetic linkage map. Comparative mapping with the chicken can provide a valuable tool for studies related to duck biodiversity and population genetics.

REFERENCES

Botstein, D., R. L. White, M. Skolnick and R. W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

Denk, A. G., B. Gautschi, K. Carter and B. Kempenaers. 2004. Seven polymorphic microsatellite loci for paternity assessment in the mallard (*Anas platyrhynchos*). Mol. Ecol. Notes 4:506-508.

Guay, P. -J. and R. A. Mulder. 2005. Isolation and characterization of microsatellite markers in musk duck (*Biziura lobata*: Aves), and their application to other waterfowl species. Mol. Ecol. Notes 5:249-252.

Hamilton, M. B., E. L. Pincus, A. Di Fiore and R. C. Fleischer. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. Biotechniques 27:500-507.

Hsu, Y. C., L. L. Severinghaus, Y. S. Lin and S. H. Li. 2003. Isolation and characterization of microsatellite DNA markers from the Lanyu scops owl (*Otus elegans botelensis*). Mol. Ecol. Notes 3:595-597.

Huang, Y. H., J. F. Tu, X. B. Cheng, B. Tang, X. X. Hu, Z. L. Liu, J. D. Feng, Y. K. Lou, L. Lin, K. Xu, Y. L. Zhao and N. Li. 2005. Characterization of 35 novel microsatellite DNA markers from the duck (*Anas platyrhynchos*) genome and cross-amplification in other birds. Genet. Sel. Evol. 37:455-472.

Huang, Y. H., Y. H. Zhao, C. S. Haley, S. P. Hu, J. P. Hao, C. X. Wu and N. Li. 2006. A genetic and cytogenetic map for the duck *(Anas platyrhynchos*). Genetics 173:287-296.

Kong, H. S., J. D. Oh, J. H. Lee, K. J. Jo, B. D. Sang, C. H. Choi, S. D. Kim, S. J. Lee, S. H. Yeon, G. J. Jeon and H. K. Lee. 2006. Genetic variation and relationships of korean native chickens and foreign breeds using 15 microsatellite markers. Asian-Aust. J. Anim. Sci. 19:1546-1550.

Liu, G. Q., X. P. Jiang, J. Y. Wang and Z. Y. Wang. 2006. Correlations between heterozygosity at microsatellite loci, mean d2 and body weight in a chinese native chicken. Asian-Aust. J. Anim. Sci. 19:1671-1677.

Maak, S., K. Neumann, G. von Lengerken and R. Gattermann. 2000. First seven microsatellites developed for the Peking duck (*Anas platyrhynchos*). Anim. Genet. 31:228-241.

Maak, S., K. Wimmers, S. Weigend and K. Neumann. 2003. Isolation and characterization of 18 microsatellites in the Peking duck (*Anas platyrhynchos*) and their application in other waterfowl species. Mol. Ecol. Notes 3:224-227.

Marshall, T. C., J. Slate, L. E. B. Kruuk and J. M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7:639-655.

Meuwissen, T. H. E. and M. E. Goddard. 1996. The use of marker haplotypes in animal breeding schemes. Genet. Sel. Evol. 28:161-176.

Munoz-Fuentes, V., N. Gyllenstrand, J. J. Negro, A. J. Green and C. Vila. 2005. Microsatellite markers for two stifftail ducks: the white-headed duck, *Oxyura leucocephala*, and the ruddy duck, *O. jamaicensis*. Mol. Ecol. Notes 5:263-265.

Paulus, K. B. and R. Tiedemann. 2003. Ten polymorphic autosomal microsatellite loci for Eider duck *Somateria mollissima* and their cross-species applicability among waterfowl species (Anatidae). Mol. Ecol. Notes 3:250-252.

Reed, K. M., J. Holm, M. Morrisson, S. Leroux and A. Vignal. 2005. Assignment of non-informative turkey genetic markers through comparative approaches. Cytogenet. Genome Res. 109:527-532.

Stai, S. M. and C. R. Hughes. 2003. Characterization of microsatellite loci in wild and domestic Muscovy ducks (*Cairina moschata*). Anim. Genet. 34:387-389.

Stallings, R. L., A. F. Ford, D. Nelson, D. C. Torney, C. E. Hildebrand and R. K. Moyzis. 1991. Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. Genomics 10:807-815.

EFFICIENCY OF FEED USE IN BEEF CATTLE [1]

ROBERT M. KOCH,[2] L. A. SWIGER,[2] DOYLE CHAMBERS [3] AND K. E. GREGORY [4,5]

University of Nebraska, Oklahoma State University and United States Department of Agriculture

DIFFERENCES among animals in converting feed into body tissue are important in determining net income from beef cattle operations. However, measuring individual feed consumption is costly because of increased equipment and labor requirements. The heritability of efficiency of feed use and its genetic relationship with other measurable traits need to be examined carefully before recommendations concerning individual feeding can be made. The problem of measuring efficiency of feed use is discussed in this paper and various measures are evaluated. The genetic and phenotypic variation and covariation among efficiency, gain and feed consumption are examined.

Data

Data were collected on 1324 individually-fed bull and heifer calves from the experimental breeding herds located at the Fort Robinson Beef Cattle Research Station, Crawford, Nebraska, the Nebraska Agricultural Experiment Station, Lincoln, and the Fort Reno Livestock Research Station, El Reno, Oklahoma. All male calves were maintained as bulls for the entire performance period studied. The calves in this study were born in the years 1951 to 1955 at Lincoln and Fort Robinson and 1954 to 1959 at Fort Reno. Heifers from the 1955 calf crop at Fort Robinson were not individually fed. Only bulls were individually fed at Fort Reno. Two closed Hereford lines and one each of the Angus and Shorthorn breeds were located at Lincoln. Six Hereford lines and one Angus line were in production at the Fort Robinson Station in 1951. By 1955, ten Hereford and two Angus lines were being carried at Fort Robinson. Three Hereford lines and one Angus line were located at Fort Reno. The calves were the progeny of 61 sires at Fort Robinson, 25 sires at Lincoln, and 34 sires at Fort Reno.

At Lincoln and Fort Robinson a 168-day feeding period was started 28 days following weaning. During the post-weaning period, bulls and heifers were separated and kept in large lots when they were not tied to the individual feeders. The calves born in 1951 were fed by tying the calves to individual self-feeders for 2 hours in the morning and for a comparable period in the afternoon. In the 4 succeeding years the calves were tied to their feeders in the evening and turned out in the morning. In all cases, feed consumption was limited only by the appetite or capacity of the animals and the time allotted for eating. Gain in weight and feed consumption were recorded for each calf and summarized at 28-day intervals. Data from the calves born at Lincoln were adjusted for inbreeding of the calf and inbreeding of the dam using the adjustments derived in an earlier analysis (Swiger *et al.*, 1961). Data from the Fort Robinson herd were not adjusted for inbreeding effects because the average coefficient of inbreeding was low (0.03 for dams and 0.05 for calves). At Fort Reno, a 154-day feeding period was started at weaning. The bulls were assigned to three lots on the basis of age and were assigned at random to feed pens in those lots. The calves were put in individual feeding pens in the evening and turned out in the morning. In the Fort Reno data the average coefficients of inbreeding were low (0.03 for dams and 0.04 for calves). The inbreeding effects accounted for less than 2% of the variation in each of the traits studied. Therefore, the traits were not adjusted for inbreeding.

The rations fed are shown in table 1. The average TDN intake per day was 5.7 lb. at Fort Robinson, 9.6 lb. at Lincoln and 11.4 lb. at Fort Reno. The stations were considered separately throughout the analysis to examine the effect that level of feed consumption and

[1] Published with approval of the Director as Paper No. 1288, Journal Series, Nebraska Agriculture Experiment Station. Contribution from the North Central Regional Project NC-1, Improvement of Beef Cattle through Breeding Methods.

[2] University of Nebraska, Lincoln.

[3] Oklahoma State University, Stillwater.

[4] Beef Cattle Research Branch, Animal Husbandry Research Division, Agricultural Research Service, United States Department of Agriculture, Lincoln, Nebraska.

[5] The authors express their appreciation to V. H. Arthaud, R. L. Davis and E. W. Schleicher for assistance in collecting the data.

TABLE 1. COMPOSITION OF RATION FED[a] EACH YEAR AT LINCOLN, FORT ROBINSON AND FORT RENO

	Lincoln		Fort Robinson
1951	45.5% ground shelled corn 14.0% ground oats 7.0% wheat bran 3.5% soybean oil meal 30.0% alfalfa hay	1951	Alfalfa hay
		1952	50% alfalfa hay and 50% oat hay
		1953	Same as 1952
1952	40.0% ground shelled corn 17.0% ground oats 15.0% dried molasses beet pulp 3.0% soybean oil meal 25.0% alfalfa hay	1954 Bulls	Hay same as 1952, plus 3.5 lb. barley and soybean meal mixture per day
		Heifers	50% alfalfa hay and 50% oat hay
1953	38.0% ground shelled corn 15.0% ground oats 15.0% dried molasses beet pulp 7.0% soybean oil meal 18.8% alfalfa hay 6.2% brome hay	1955	Hay same as 1954, plus 4.2 lb. oats per day
			Fort Reno
		1954–59	35.0% ground ear corn 10.0% whole oats 10.0% wheat bran 5.0% blackstrap molasses 10.0% cottonseed meal 20.0% cottonseed hulls 10.0% alfalfa hay
1954	38.0% ground shelled corn 15.0% ground oats 15.0% dried molasses beet pulp 7.0% soybean oil meal 12.5% alfalfa hay 12.5% brome hay		
1955	Same as 1954		

[a] Fed *ad libitum* unless otherwise indicated; hay was chopped at all stations.

composition of the ration might have on the relationship studied. The approximate amount of total digestible nutrients consumed (TDN) were computed using the average values for feeds presented by the National Research Council (N.R.C., 1950).

Measuring Feed Efficiency

Feed efficiency is defined in this study as the gain in body weight resulting from the consumption of a given amount of feed or its inverse. Variation in the composition of gains (fat, lean, or bone) and in maintenance requirements prevents this measure from being a precise estimate of energy conversion rate. The most useful criterion for evaluting efficiency of feed use in beef animals may be the amount of edible product produced for a given energy intake rather than the fraction of energy in the feed which was converted to animal tissue. In this analysis only live weight gains were available to measure efficiency of feed use.

Efficiency of feed use is not a directly measurable trait, but must be computed as a function of feed consumed, gain in body weight and time. Sometimes one or more of the traits are held constant. In experiments based on a fixed amount of gain or feed consumption, animals would be on test at different times. Variations in efficiency could result from animals being evaluated under differing environmental conditions. In experiments conducted over the same time interval, variations in body weight, and thus maintenance requirements, differences in composition of gain, and differences in feed consumption influence variations in measured efficiency. The average effects of the measurable traits can be evaluated statistically, but differences in environmental conditions where each animal is tested over a different time period cannot be evaluated.

Animals of different weights have different requirements for maintenance. Since each feeding test reported here began at a fixed time each year (about 1 month following weaning), the calves were evaluated through different ranges of body weight, such as 300 to 600, 400 to 700, or 500 to 800 lb. For fair comparisons the measure of efficiency must take into account differences in the weight at which various animals were evaluated. In these data, feed efficiency was considered a function of gain, feed consumption and average weight while on test.

Feed efficiency can be computed by holding either gain or feed consumption constant statistically. That is, feed efficiency can be expressed as: (1) feed consumption adjusted for differences in gain or (2) gain adjusted for differences in feed consumption. The gain or feed consumption observed is partitioned into a portion expected on the basis of average performance of animals consuming a given amount of feed or making a given amount of gain depending on whether (1) or (2) is being considered. The deviation from the expected value based on regression is attributed to differences in efficiency of feed use. In each case differences in weight should be accounted for since they affect average maintenance requirements.

Animal husbandmen have almost universally computed efficiency either as the ratio of gain to feed or feed to gain. The ratio of gain to feed assumes the biological relationship is represented by the linear function, Gain=rate of feed conversion x feed+error, where the variance of the error for gain is proportional to the amount of feed consumed and the rate of feed conversion is the increase in body weight per unit of feed consumed. The least squares estimator of this regression is the average ratio of gain to feed and is a line passing through the origin of the gain and feed axis. Since some feed is required for maintenance even when there is no gain, a measure which forces the regression through the origin is biologically unrealistic. The regression of gain on feed should allow for estimation of the intercept of the feed-axis caused by maintenance requirements. It seems unlikely that the relationship between gain and feed consumed would be linear over a wide range of gain and feed consumption although linear regression may be descriptive over a limited range. Plotting the data used in this study indicated that quadratic regressions would describe the relationship between gain, feed consumption and weight over the ranges observed at each station with almost as much precision as more complex mathematical curves.

It is stressed that only gain and feed consumption can be measured directly. Feed efficiency can only be inferred from the arithmetic manipulation of gain and feed consumption. As a result only two degrees of freedom are available for studying the relationships between gain, feed consumption and feed efficiency. As will be seen later, certain automatic relationships must occur when these three variables are considered simultaneously.

In other studies the ratio of gain and feed has been used as a measure of efficiency and then correlated with gain or feed consumption. Interpretation of such relationships can lead to faulty conclusions because of spurious correlations (Pearson, 1897) and the failure of the index to describe the biological function accurately (Tanner, 1949).

The methods used in this study maximize the variation in efficiency. For example, adjusting gain for feed consumption removes that part of the variation in gain attributed to differences in feed consumption. The remaining variation in gain is attributed to differences in efficiency of feed use. Since the phenotypic measure of efficiency is a residual, it is a maximum estimate of the variation associated with efficiency and includes any errors of measurement of gain in body weight. The gain-feed ratio, commonly used by animal husbandmen, is also a residual measure and similarly maximizes the variation attributed to efficiency.

For the first measure of efficiency, feed consumption was adjusted for gain and mid-weight (initial weight plus one-half of the gain). Four partial regressions, i.e., both first and second degree partial regression coefficients of feed consumption on gain and feed consumption on mid-weight, were computed simultaneously. The data were classified by sex, year of birth, feeding group within year (Lincoln had two weaning groups each year), line and sire. The sums of squares and cross-products within subclasses were pooled to compute the partial regressions for each station. The four partial regressions jointly accounted for 48, 60 and 60% of the sums of squares of feed consumption at Fort Robinson, Lincoln and Fort Reno, respectively. Feed consumption adjusted in this manner represents feed consumption through the same range of live weight. Calves with low values of adjusted feed consumption would be more efficient than those with high values.

The second measure of efficiency was computed by adjusting gain for differences in feed consumption after first adjusting feed consumption for differences in mid-weight to account for average differences in maintenance. Feed consumption was adjusted for mid-weight using the two partial regressions (i.e., the quadratic regression) of feed consump-

tion on mid-weight which were fitted simultaneously with the regressions of feed consumption on gain. An alternative to adjusting for mid-weight would be to adjust only for differences in initial weight. However, if two calves with the same initial weight gained different amounts during the feeding period, they would receive the same adjustment when in reality the calves would have maintained different weights through the feeding period. Another alternative would be the use of a weight taken near the middle of the feeding period. This weight would be a more accurate measure of weight maintained through the feeding period than initial weight and the error due to differences in fill would not be associated with the weighing errors in measuring gain from initial and final weights. If the regression were markedly curvilinear and extended over a wide range, it would be advantageous to make adjustments for differences in weight at several intervals instead of a single adjustment as was made here.

Regression equations for gain on feed consumption adjusted for mid-weight were computed separately for bulls and heifers. The differences between the regressions of gain on feed consumption for bulls and heifers at Lincoln and Fort Robinson were small and a test of significance of pooling the data for both sexes vs. computing separate regressions yielded an F value of less than one. Therefore, the data from the two sexes were pooled. The quadratic regression was significant for the Fort Reno and Fort Robinson data, but not for the Lincoln data. In the Fort Robinson data the quadratic equation accounted for 36.3% of the variation in gain while fitting a linear regression accounted for 33.7%. At Fort Reno these values were 39.4 and 36.4% for the quadratic and linear regression, respectively. In the Lincoln data the quadratic regression accounted for 24.8% and the linear regression 24.7% of the variation in gain. Figure 1 shows the relationship between gain and feed consumption (TDN) when the quadratic regressions were plotted for Fort Robinson and Fort Reno and the linear regression was plotted for Lincoln.

The linear regression coefficients of gain on pounds of TDN consumed for Fort Robinson, Lincoln and Fort Reno were $0.15 \pm .009$, $0.17 \pm .020$ and $0.19 \pm .018$ (expressed as lb. of gain per lb. of TDN), respectively. The similarity of the coefficients is striking considering the differences in the environments and feeding levels at the three stations. The slight increase in size of the coefficients, 0.15, 0.17 and 0.19, corresponds to daily fed intakes of 5.7, 9.6 and 11.4 lb. of TDN at Fort Robinson, Lincoln and Fort Reno, respectively.

Differences between the three stations in average gain per pound of feed (TDN) are also apparent in figure 1 when the respective ranges of feed consumption overlap. This difference between stations in converting TDN into gain could be due to real differences in average feed efficiency, differences in maintenance requirements or inaccuracies involved in comparing high roughage with more concentrated rations on a TDN basis —perhaps an over estimation of the net energy in roughages (Blaxter, 1956).

Once maintenance requirements are met, increased increments of feed consumption would be expected to result in proportionate increases in gain, provided there are no changes

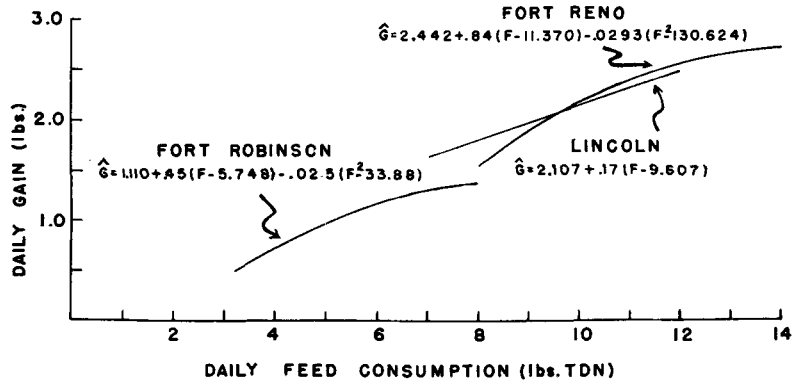


Figure 1. Regression of gain on feed consumption (adjusted for weight) for Fort Robinson, Lincoln and Fort Reno.

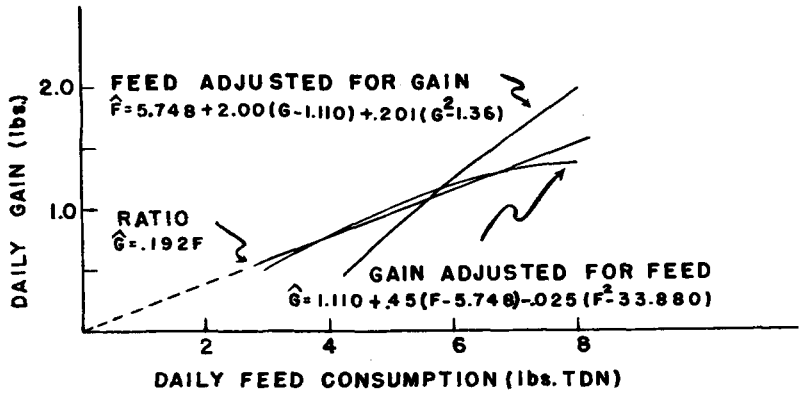


Figure 2. Regression of feed consumption on gain, regression of gain on feed consumption, and the ratio of gain to feed consumed for Fort Robinson (adjusted for weight).

in composition of gain and/or efficiency of feed use. Figure 1 shows that increases in feed consumtion did not result in proportional increases in gain. This curvilinear effect could result from increased losses through heat increment of feeding (Brody, 1945) and/or changes in composition of gain at the higher levels of intake.

The third measure of efficiency of feed use studied was the ratio of gain to feed consumed. Feed consumption adjusted for differences in mid-weight was used as the denominator. This measure of feed efficiency was computed to study its relation to the other measures described above since it is the most commonly reported measure of efficiency.

A comparison of the three regressions: (1) feed consumption adjusted for differences in gain and weight ($Feed_A$), (2) gain adjusted for differences in feed consumption and weight ($Gain_A$), and (3) the ratio of gain to feed consumed (adjusted for weight) are shown in figure 2 for the Fort Robinson data only. Correlations were computed among these three measures of efficiency and are given in table 2.

Consider the correlations between the gain-to-feed ratio (Ratio) and adjusted gain ($Gain_A$). Values of the ratio are measured as deviations from a line fitting the data as closely as possible with the restriction of passing through the origin. Figure 2 suggests that the correlation between the ratio and adjusted gain should be high for the Fort Robinson data since such a line would lie close to the regressions actually computed between gain and feed consumption for the second measure of efficiency. The ratio may be highly correlated with adjusted gain over narrow ranges of feed consumption and when gain in weight is fairly large. The ratio of gain to feed consumed would be a relatively poor measure of efficiency if the animals had been fed near maintenance or if feed consumption and gain varied over a wide enough range for the curvilinearity to be of more importance.

The correlation between $Gain_A$ and $Feed_A$ depends on the correlation between gain and feed consumption and can only be high when differences in efficiency of feed use and errors of measurement are small. From a functional

TABLE 2. CORRELATIONS BETWEEN THREE MEASURES OF EFFICIENCY OF FEED USE, (1) FEED CONSUMPTION ADJUSTED FOR GAIN AND WEIGHT ($FEED_A$) (2) GAIN ADJUSTED FOR CONSUMPTION AND WEIGHT ($GAIN_A$) AND (3) THE RATIO OF GAIN TO FEED CONSUMPTION ADJUSTED FOR WEIGHT (RATIO)

	Fort Robinson		Lincoln		Fort Reno	
Traits	Range[a]	Correlation	Range	Correlation	Range	Correlation
$Feed_A$, $gain_A$		—.79		—.52		—.55
$Feed_A$, ratio	3 to 9	—.69	7 to 12	—.74	8 to 14	—.64
$Gain_A$, ratio		0.88		0.90		0.96

[a] Range in lb. of feed consumption (TDN) of calves included in computing the correlations.

view the choice of dependent and independent variables favors the regression of gain on feed ($Gain_A$).

Genetic Analysis

The data were analyzed by Method 2 of Henderson (1953). Least-square effects were fitted for years and sexes simultaneously with sire effects. The data were adjusted for year and sex effects and sire components were estimated from the mean squares for sires within lines. The variances and covariances of half-sibs were computed within subclasses prior to adjusting the data. The estimates of the phenotypic, genetic and environmental components of variance and covariance were computed from the sire and half-sib components using the average inbreeding coefficients of calves and their parents. Heritabilities and phenotypic, genetic and environmental correlations were computed using these components. The degrees of freedom for sires within lines were 21, 49 and 30, respectively, for Lincoln, Fort Robinson and Fort Reno. The corresponding degrees of freedom for the half-sib components were 227, 602 and 193, respectively.

Heritabilities. The heritabilities and the phenotypic components of variance for the various traits are shown in table 3. The average level of TDN consumption at Fort Robinson was 5.7 lb.; at Lincoln, 9.6 lb.; and at Fort Reno, 11.4 lb. The heritability estimates varied from values less than 0 to values greater than 1.0, exceeding the theoretical limits for heritability. The magnitude of the estimates did not seem to be related to the level of feeding. The combined estimate was obtained by weighting the heritabilities from the separate stations according to the inverse of their approximate sampling variances.

The combined estimate of 0.65 for heritability of gain on test compares with 0.50, the average value of other workers reporting heritability from half-sib data (Brown and Gifford, 1962; Dawson *et al.*, 1955; Knapp and Clark, 1950, 1951; Knapp and Nordskog, 1946; McCormick *et al.*, 1956; Patterson *et al*,. 1955; Shelby *et al.*, 1955, 1960; Swiger, 1961; Urick *et al.*, 1957; Warwick and Cartwright, 1955). An average heritability value of 0.66 for gain in feedlot by sire-offspring regression methods was reported by Carter and Kincaid, (1959), Chambers *et al.* (1961) and Knapp and Nordskog (1946).

The heritabilities of 0.62 for gain adjusted for feed consumed and 0.28 for feed consumed adjusted for differences in gain have no counterparts in the literature. The heritability of 0.36 for ratio of gain to feed consumed compares with an average value of 0.51 reported from half-sib data by Brown and Gifford (1962), Carter and Kincaid (1959), Dawson *et al.* (1955), Knapp and Nordskog (1946), and Shelby *et al.* (1955, 1960). An average value of 0.35 for heritability of feed gain ratio estimated by sire-offspring regression was reported by Carter and Kincaid (1959) and Knapp and Nordskog (1946).

The heritability of 0.64 for feed consumed compares with 0.76 reported by Brown and Gifford (1962).

Relationships among Gain, Feed Consumption and Efficiency. The relationships among

TABLE 3. HERITABILITY ESTIMATES AND THEIR APPROXIMATE STANDARD ERRORS AND THE PHENOTYPIC COMPONENTS OF VARIANCE

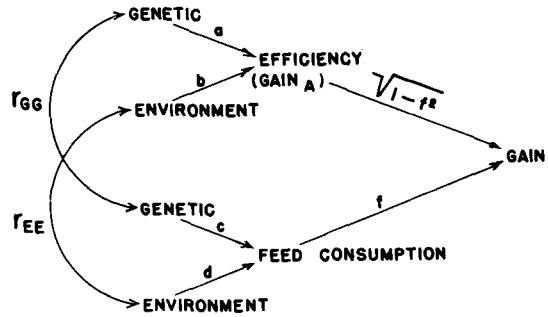
	Station			
Heritability estimates	Fort Robinson	Lincoln	Fort Reno	Combined
Gain on test	0.56±0.17	1.32±0.33	0.43±0.26	0.65±0.13
Feed consumed	0.44±0.15	0.87±0.28	1.24±0.33	0.64±0.12
$Feed_A$	0.52±0.16	—.21±0.16	1.37±0.33	0.28±0.11
$Gain_A$	0.53±0.16	0.72±0.25	0.77±0.29	0.62±0.12
Ratio	0.59±0.17	0.14±0.13	0.82±0.30	0.36±0.10
Phenotypic variance (values based on average daily gain and TDN per day)				
Gain on test	0.0412	0.0958	0.0934	
Feed consumed	0.5637	0.6692	1.3057	
$Feed_A$	0.3912	0.3287	0.8798	
$Gain_A$	0.0261	0.0525	0.0643	
Ratio	0.000850	0.000488	0.000525	

gain, feed consumption and efficiency have certain automatic properties because only gain and feed consumption are measured directly. Efficiency is computed as a function of gain and feed consumption. In this study the phenotypic correlation between feed consumption and gain was used to partition the variation in gain into two parts. One part was attributed to differences in feed consumption, $f^2\ \sigma^2_{Gain}$, where f is the phenotypic correlation between feed consumption and gain. The other part was the residual, $(1\text{-}f^2)\ \sigma^2_{Gain}$, not accounted for by differences in feed consumption. The residual portion is our measure of efficiency. The two portions of gain (the expected gain according to feed consumption and the remainder) were then partitioned into genetic and environmental components. These relationships are shown in figure 3 Treating efficiency as a residual means the phenotypic correlation between feed consumption and efficiency is automatically zero. The correlations between the genetic and the environmental factors affecting feed consumption and efficiency were determined by the relationships among sire means with the restriction that the sum of the genetic and the environmental influences must equal zero.

Sampling errors and/or station differences prevent any clear interpretation concerning differences in the relationships among gain, feed consumption, and efficiency associated with level of energy intake.

All sets of data indicate a sizable direct influence of genes affecting gain through efficiency. For the combined estimate $a^2(1\text{-}f^2)=0.38$. All sets of data indicate a sizeable direct effect of genes affecting gain through feed consumption. For the combined estimate $c^2f^2=0.25$. Although the relationship between genes which affect efficiency and those which affect feed consumption (r_{GG}) is not consistent between stations, two sets of data and the combined estimate indicate the relationship is positive. The combined estimates indicate that efficiency and feed consumption jointly account for 0.02 of the genetic variation in gain, $2r_{GG}f\sqrt{1\text{-}f^2}$ ac.

The genetic value for gain is the sum of the genetic values for efficiency and feed con-



f = PHENOTYPIC CORRELATION BETWEEN FEED CONSUMPTION AND GAIN

$\sqrt{1-f^2}$ = RESIDUAL CORRELATION BETWEEN GAIN AND FACTORS AFFECTING GAIN OTHER THAN FEED CONSUMPTION

	VALUES OF PATHS FOR THE THREE LOCATIONS			
PATHS	FORT ROBINSON	LINCOLN	FORT RENO	COMBINED
f	.58	.72	.58	.62
$\sqrt{1-f^2}$	.81	.69	.81	.78
a	.73	.85	.88	.79
b	.69	.53	.48	.62
c	.66	.93	1.11	.80
d	.75	.36	—	.60
r_{GG}	.13	.67	-.54	.04
r_{EE}	-.12	-2.79	—	-.07

Figure 3. Path coefficient diagram for the genetic and environmental relationships among feed efficiency, feed consumption and gain.

sumption. The correlation between the genetic value for efficiency and the phenotype of gain is $a\sqrt{1-f^2}+cf\ r_{GG}=0.64$ in the combined data. The correlation between the genetic value for feed consumption and the phenotype of gain is $cf+a\sqrt{1-f^2}\ r_{GG}=0.52$. The genetic correlation between efficiency and gain is the correlation between the genetic value of efficiency and the phenotype of gain (0.64) divided by the square root of the heritability of gain=0.64/0.81=0.79. Similarly the genetic correlation between feed consumption and gain is 0.52/0.81=0.64. The genetic correlation between efficiency and feed consumption is $r_{GG}=0.04$.

These data suggest that selection would be effective for gain, feed consumption or any of the measures of efficiency. Selection for gain will result in increased efficiency and greater feed consumption with about 60% of the genetic increase in gain being due to increased efficiency and 40% due to increased feed consumption. Selection for efficiency of gain alone would do little toward increased feed consumption and vice versa, where differences in maintenance are discounted.

If improved efficiency were the main objective then direct selection for efficiency would lead to rapid improvement with the high heritability (0.62) found in these data. However, where it is not possible to measure feed consumption and compute efficiency, selection for gain would lead to 81% as much genetic improvement as selecting directly for efficiency (genetic correlation between efficiency and gain times the square root of the ratio of the heritability of gain over the heritability of efficiency).

Studies involving carcass composition are needed to determine feed efficiency measures for energy conversion or for edible portion instead of increase in body weight without regard to composition of gain as was studied here.

Summary

Feed efficiency measured as a function of gain in body weight and feed consumed was studied for 1,324 bull and heifer calves at three experiment stations. Three measures were computed: (1) feed consumption adjusted for differences in gain; (2) gain adjusted for differences in feed consumption; and (3) the ratio of gain to feed consumed. In all three measures, mid-weight was considered simultaneously in an attempt to remove differences in maintenance requirements.

Efficiency expressed as gain adjusted for differences in feed consumption (i.e., ±deviation from the regression of gain on consumption) was considered the most accurate mathematical description of the cause and effect relationships and resulted in the highest heritability of the three measures studied.

No trends in the heritabilities calculated for each experiment station were noted. The combined heritabilities were 0.65 for gain on test, 0.64 for feed consumed, 0.62 for gain adjusted for differences in feed consumption, 0.28 for feed consumption adjusted for differences in gain, and 0.36 for the ratio of gain to feed consumed.

A path analysis of feed efficiency (gain adjusted for feed consumption), feed consumption and gain was made. The analysis indicated that 38% of the variation in gain could be attributed directly to genetic differences in feed efficiency. Genetic differences in feed consumption accounted for 25% of the variation in gain. The remaining 37% of the variation in gain was accounted for by variations in environmental influences.

The genetic correlation between feed efficiency and gain was 0.79, between feed consumption and gain was 0.64, and between feed efficiency and feed consumption was 0.04.

These results indicate that selecting for gain should be effective and lead to both increased feed efficiency and increased feed consumption. Selecting for feed efficiency would increase feed efficiency and result in increased daily gain, but feed consumption would not be affected. Selection for feed consumption would increase feed consumption and daily gain, but would lead to no improvement in feed efficiency other than that attributable to a smaller portion of the intake being used for body maintenance.

Literature Cited

Blaxter, K. L. 1956. The nutritive value of feeds as sources of energy. J. Dairy Sci. 39:1396.

Brody, Samuel. 1945. Bioenergetics and Growth. Reinhold Publ. Co. pp. 76–99.

Brown, C. J. and Warren Gifford. 1962. Estimates of heritability and genetic correlations among certain traits of performance-tested beef bulls. Ark. Agr. Exp. Sta. Bul. 653.

Carter, R. C. and C. M. Kincaid. 1959. Estimates of genetic and phenotypic parameters in beef cattle. II. Heritability estimates from parent offspring and half-sib resemblances. J. Animal Sci. 18:323.

Chambers, Doyle, Joe B. Armstrong and Dwight F. Stephens. 1961. Progeny testing beef bulls for growth. Okla. Agr. Exp. Sta. Misc. Publ. MP-64.

Dawson, W. M., T. S. Yao and A. C. Cook. 1955.

Heritability of growth, beef characters and body measurements in milking Shorthorn steers. J. Animal. Sci. 14:208.
Henderson, C. R. 1953. Estimation of variance and covariance components. Biometrics 9:226.
Knapp, Bradford, Jr. and R. T. Clark. 1950. Revised estimates of heritability of economic characters in beef cattle. J. Animal Sci. 9:582.
Knapp, Bradford, Jr. and R. T. Clark. 1951. Genetic and environmental correlations between weaning scores and subsequent gains in the feedlot with record of performance steers. J. Animal Sci. 10:365.
Knapp, Bradford, Jr. and A. W. Nordskog. 1946. Heritability of growth and efficiency in beef cattle. J. Animal Sci. 5:62.
McCormick, W. C., B. L. Southwell and E. J. Warwick. 1956. Factors affecting performance in herds of purebred and grade polled Hereford cattle. Ga. Agr. Exp. Sta. Tech. Bul. N.S. 5.
N.R.C. 1950. Nutrient requirements of domestic animals, Number IV. Nutrient requirements of beef cattle. National Research Council, Washington, D. C.
Patterson, R. E., J. T. Cartwright, J. H. Jones and J. J. Bayles. 1955. Performance testing of beef breeding stock. J. Animal Sci. 14:1034.
Pearson, Karl. 1897. Mathematical contributions to the theory of evolution—on a form of spurious correlation which may arise when indices are used in the measurement of organs. Proc. Royal Soc. (London) 60:489.
Shelby, C. E., R. T. Clark, J. R. Quesenberry and R. R. Woodward. 1960. Heritability of some economic characteristics in record of performance bulls. J. Animal Sci. 19:450.
Shelby, C. E., R. T. Clark and R. R. Woodward. 1955. The heritability of some economic characteristics of beef cattle. J. Animal Sci. 14:372.
Swiger, L. A. 1961. Genetic and environmental influences on gain of beef cattle during various periods of life. J. Animal Sci. 20:183.
Swiger, L. A., K. E. Gregory, Robert M. Koch and V. H. Arthaud. 1961. Effect of inbreeding on performance traits of beef cattle. J. Animal Sci. 20:626.
Tanner, J. M. 1949. Fallacy of per-weight and per-surface area standards, and their relation to spurious correlation. J. Appl. Physiol. 2(1):1.
Urick, J., A. E. Flower, F. S. Willson and C. E. Shelby. 1957. A genetic study in steer progeny groups during successive growth periods. J. Animal Sci. 16:217.
Warwick, Bruce L. and T. C. Cartwright. 1955. Heritability of rate of gain in young growing beef cattle. J. Animal Sci. 14:363.

Residual Feed Consumption in Laying Hens. 2. Genetic Variation and Correlations

P. LUITING and E. M. URFF

Department of Animal Breeding, Wageningen Agricultural University, P.O. Box 338, NL-6700 AH Wageningen, The Netherlands

(Received for publication November 20, 1990)

ABSTRACT The study described here deals with the quantification of the genetic variation of "residual feed consumption" (RFC) of hens of a White Leghorn population during a 44-wk laying period (20 to 64 wk of age) in 11 time segments of 4 wk each, fed either a commercial or a low-energy diet (11.7 and 10.0 MJ ME/kg, respectively, where 1 MJ = .239 Mcal). The RFC is defined operationally as the difference between the observed feed consumption of a laying hen and its consumption as predicted from a model with metabolic body weight, egg mass production, and body weight gain as independent variables.

The RFC was found to be highly heritable in all periods. The heritability of RFC accumulated over the whole laying period (RFC-T) was estimated as .42 to .62. For each time segment between 32 and 56 wk of age, genetic correlations between RFC and RFC-T were estimated to be larger than .91. The genetic sources causing variation in RFC during the first part of lay seem to differ from those causing variation later on, and to be of less importance during the rest of the laying period. It was concluded that RFC shows a considerable systematic and permanent additive genetic variance, and that RFC measurements for selection can be limited to one to three time segments between 32 and 56 wk. Furthermore, less environmental variance and therefore higher heritabilities and genetic correlations seemed to exist for birds fed the low energy diet in comparison with those fed the commercial diet. No clear differences could be found between genetic and phenotypic correlation estimates of RFC with feed consumption, metabolic body weight, egg mass production, and body weight gain.

(*Key words*: residual feed consumption, genetic parameters, maintenance, part-record, dietary energy)

1991 Poultry Science 70:1663–1672

INTRODUCTION

Luiting and Urff (1991b) concluded from multiple regression analyses in a population of White Leghorn (WL) laying hens, that 10 to 30% of the variance of daily feed consumption (FCD) among hens remained unaccounted for by metabolic body weight ($MBW^{.75}$), daily egg mass production (adjusted for abnormal eggs, EMDc), and body weight gain (BWG). This unexplained term is referred to as "residual feed consumption" (RFC), defined operationally as the difference between the observed feed consumption of a hen and its consumption as predicted from $MBW^{.75}$, EMDc, and BWG. In other words, hens showing equal production levels and body weights may differ considerably with regard to feed consumption, and therefore feed efficiency.

Differences among hens in maintenance requirements per metabolic kilogram (kilogram$^{.75}$ of body weight) seem primarily responsible for this and a large amount of systematic and permanent variation in RFC appears to exist (Luiting and Urff, 1991b). The repeatability of RFC was estimated as .52 to .58. If a large part of this systematic and permanent variation is of genetic origin, it might be exploited in a breeding program.

The study described here deals with the quantification of the genetic variation of RFC of hens of a WL population during a 44-wk laying period in 11 time periods of 4 wk. This is done by estimating 1) heritabilities of RFC in different time segments, 2) genetic correlations between RFC measurements in different time segments, and 3) genetic correlations between RFC and $MBW^{.75}$, EMDc, BWG, and FCD per period of 4 wk. Morris (1972) suggested that differences in partial energetic efficiencies between laying hens are expressed more clearly at a suboptimum energy consumption. To verify this, the hens in the present study were fed with either a commercial diet or a low-energy diet (11.7 and 10.0 MJ ME/kg, respectively, where 1 MJ = .239 Mcal).

MATERIALS AND METHODS

Data

The population of hens described by Luiting and Urff (1991a) was used for the present study. It consisted of 704 hens in 94 half sib groups of 3 or 4 full sib pairs each (for a total of 352 full sib pairs). From 18 wk of age, when the hens were transferred to individual battery cages, one member of each full sib pair was fed a commercial diet (calculated content 11.7 MJ ME/kg and 155 g CP/kg) and the other one a low energy diet (10.0 MJ ME/kg and 152 g CP/kg), both provided for *ad libitum* intake. Details were given by Luiting and Urff (1991a).

According to Luiting and Urff (1991a), multiple regressions of FCD (in grams per day) on $MBW^{.75}$ (in kilograms$^{.75}$), EMDc (in grams per day), and BWG (in grams per day) were performed within each of 11 4-wk periods (starting at 20 wk of age) and within each diet; furthermore, the residual feed consumption (RFC in kiloJoules of ME per day) was calculated per hen and per period by multiplying with the calculated ME contents (Luiting and Urff, 1991b). To investigate the overall performance per hen, RFC was averaged over all 4-wk periods (RFC-T). Following Manson (1972) an averaged RFC per hen was also calculated for the following time segments: 24 to 32, 20 to 32, 32 to 44, 20 to 44, 44 to 64, and 32 to 64 wk of age (RFC_{24-32}, RFC_{20-32}, RFC_{32-44}, RFC_{20-44}, RFC_{44-64}, and RFC_{32-64}, respectively). To investigate the inheritance of individual stability of RFC per hen, the standard deviation of RFC among periods for each hen (sdRFC) was calculated after standardization and log-transformation (Luiting and Urff, 1991b).

Heritabilities of Residual Feed Consumption

A univariate restricted maximum likelihood procedure (the EM algorithm by Meyer, 1986) was used to estimate variance components of RFC within diet, for each single period, for cumulative periods and for the defined time segments, from the following random model:

$$RFC_{ij} = \mu + s_i + e_{ij} \qquad [1]$$

where i = 1, . . . , s; s = number of sires in the analysis; j = 1, . . . , n; n = number of hens per sire; RFC_{ij} = residual feed consumption of hen ij (kJ ME/day); μ = overall mean; s_i = random effect of sire i; e_{ij} = random effect of hen j within sire i.

Estimates from Henderson's Method III were used as starting values for variance components. Iteration was stopped when differences in variance components between successive rounds were less than .01%. Genetic variances (σ^2_A) were calculated from between-sire variances (σ^2_S) as $\sigma^2_A = 4\sigma^2_S$, phenotypic variances (σ^2_P) as $\sigma^2_P = \sigma^2_S + \sigma^2_e$, where σ^2_e is the error variance, and environmental variances (σ^2_E) as $\sigma^2_E = \sigma^2_P - \sigma^2_A$. Heritabilities ($h^2$) were derived as $h^2 = \sigma^2_A/\sigma^2_P$. Sampling errors of genetic parameters were approximated according to Meyer (1986); no formal significance tests of parameter diet-contrasts and time trends have been performed in view of the non-normal distributions of the parameters or the dependence of estimates. The same procedure was followed for sdRFC.

Genetic Correlations of Residual Feed Consumption Over Time and with Economic Traits

Covariance components of RFC during each single period with RFC-T, of RFC accumulated over successively longer periods with RFC-T, and among RFC in the defined time segments, were estimated per diet by the same procedure as described for variance components, but now from a bivariate random model with the same structure as [1], using a multivariate restricted maximum likelihood procedure with equal designs. Genetic covariances (σ_{Aij}, where the subscripts i and j denote the traits) were computed from the between-sire covariances (σ_{Sij}) as $\sigma_{Aij} = 4\sigma_{Sij}$, phenotypic covariances (σ_{Pij}) as $\sigma_{Pij} = \sigma_{Sij} + \sigma_{eij}$, where σ_{eij} is the error covariance, and environmental covariances (σ_{Eij}) as $\sigma_{Eij} = \sigma_{Pij} - \sigma_{Aij}$. Genetic ($r_A$) and environmental ($r_E$) correlations were derived as $r_A = \sigma_{Sij}/(\sigma_{Si}\,\sigma_{Sj})$ and $r_E = \sigma_{Eij}/(\sigma_{Ei}\,\sigma_{Ej})$. The same procedure was followed for covariances of RFC-T with sdRFC. The same procedure was followed for (co)variances of RFC with $MBW^{.75}$, EMDc, BWG, and FCD per diet and per period, and per diet averaged over periods.

RESULTS

Heritabilities of Residual Feed Consumption Over Time

The left side of Figure 1 shows the estimated σ^2_A, σ^2_E, and h^2 of RFC for single periods. Sampling errors of h^2 estimates were within the range of .16 to .24. The h^2 estimates varied considerably among periods, and had a tendency to decrease with time (about 1.5 to 2 times the sampling error). The estimates of σ^2_A showed a large amount of variation among periods but there was no clear time trend. In contrast, σ^2_E increased gradually over time. Also h^2 estimates of the low-energy diet were on average slightly higher than those of the commercial diet. For σ^2_A no clear difference between the diets was observed, in contrast to σ^2_E, which had lower values on the low-energy diet.

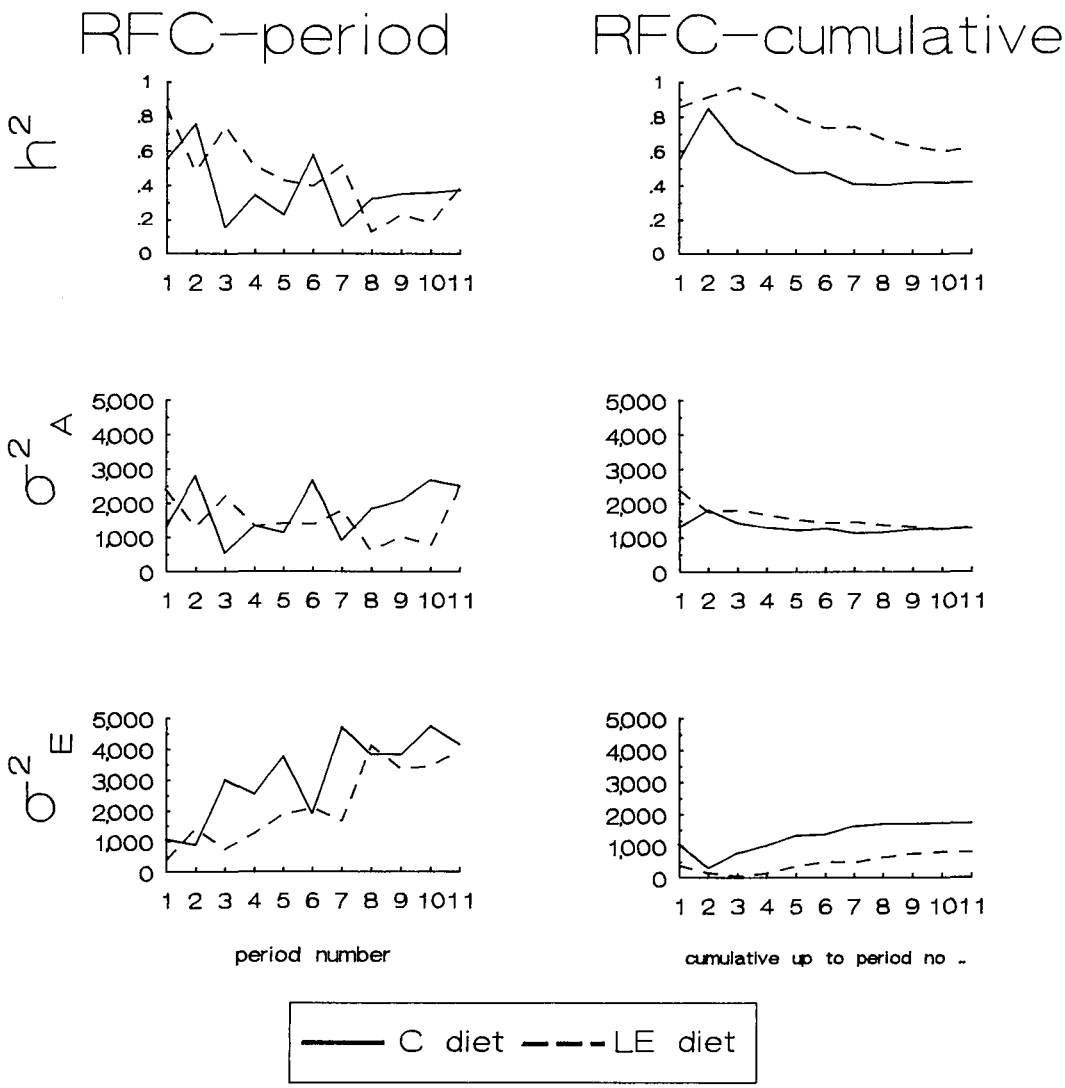


FIGURE 1. Heritabilities (h^2) and genetic (σ^2_A) and environmental (σ^2_E) variances of residual feed consumption on the commercial diet (C) and on the low-energy (LE) diet (11.7 and 10.0 MJ ME/kg, respectively, where 1 MJ = .239 Mcal) during each single period (RFC-period in kiloJoules of ME per day) and during successively longer accumulated periods (RFC-cumulative in kiloJoules of ME per day).

The right side of Figure 1 shows the estimated σ^2_A, σ^2_E, and h^2 of RFC over accumulated periods. Sampling errors of h^2 estimates ranged from .21 to .23. The same time trends and diet differences were noted as was observed for RFC for single periods, but the graphs were more stable, the h^2 estimates were higher and the σ^2_E estimates were lower. Cumulative h^2 estimates over the first three periods were very high (.68 and .93 for the commercial and the low-energy diet, respectively), whereas σ^2_A estimates were somewhat larger and σ^2_E estimates much smaller than in later periods. After that, cumulative h^2 decreased gradually to .42 and .62 for the commercial and low-energy diet, respectively; σ^2_A was stable and σ^2_E increased to a stable level after cumulation to six or eight periods.

Genetic Correlations of Residual Feed Consumption Over Time

The left side of Figure 2 shows the σ_{Aij}, σ_{Eij}, and correlations of RFC measurements in each single period with the averaged RFC-T. Sampling errors of r_A estimates ranged from .06 to .28. Whereas σ^2_E (Figure 1) increased continuously, σ_{Eij} increased primarily in the beginning. Therefore, r_E was lower at the beginning of the laying period (especially for the low-energy diet). Just as for σ^2_A (Figure 1), the estimates of σ_{Aij} did not show a clear time trend. Thus, r_A estimates show little of a pattern in time (they were all very close to unity), except for a small increase in the first few periods.

The r_A estimates did not differ clearly between the two diets, with the possible exceptions of the first two periods and at the end of the laying period, when the low energy diet had the highest values (Figure 2). The r_E estimates were generally a little lower on the low-energy diet, which was the result of lower σ_{Eij} and somewhat lower σ^2_E (Figure 1).

The right side of Figure 2 shows the σ_{Aij}, σ_{Eij}, and correlations of RFC accumulated over successively longer periods with RFC-T. Sampling errors of r_A estimates ranged from .00 (for estimates close to unity) to .24. Again, the same time trends and diet differences were observed for RFC accumulated over successively longer periods as for RFC in each single period: very high r_A estimates (up to unity), especially after the first two periods; higher σ_{Eij} estimates and, only at the beginning of the laying period, somewhat smaller r_A on the commercial diet than on the low-energy diet.

Heritabilities and Genetic Correlations of Residual Feed Consumption Time Segments

Table 1 gives the estimated h^2 of and r_A among the defined time segments. These parameters followed the same trends as observed for the single and accumulated periods. Up to 32 wk of age h^2 estimates were higher than after that age (especially for the low-energy diet) but estimates of r_A with RFC-T were lower (especially for the commercial diet). Estimates of accumulated h^2 and r_A remained higher and lower, respectively, when the first periods up to 32 wk were included. All genetic parameter estimates were higher on the low-energy diet than on the commercial diet.

The h^2 estimates for sdRFC were .06 and .15 for the low-energy and the commercial diet, respectively (sampling errors .05 to .26). Respective estimates of r_A for sdRFC with RFC-T were .29 and .04 (sampling errors .57 to .82).

Genetic Correlations of Residual Feed Consumption with Economic Traits

Estimates of h^2 for FCD, $MBW^{.75}$, EMDc, and BWG, as presented at the left side of Figure 3, resulted in 16 out of 66 (diets by traits by periods) cases in either nonconvergent (four cases) or impermissible results ($\sigma^2_A<0$; 12 cases); most of these cases apply to EMDc and BWG. The estimates of h^2 for FCD and $MBW^{.75}$ indicate additive genetic variation during the whole laying period. For $MBW^{.75}$ higher h^2 values were obtained for the commercial than for the low-energy diet; for FCD no such diet difference was found. For EMDc and BWG h^2 estimates declined over time and reached zero values.

Estimation of genetic correlation coefficients per period between RFC and FCD, $MBW^{.75}$, EMDc, and BWG resulted in 15 more cases of impermissible results ($|r_A|>1$); most of these cases apply to EMDc and BWG in periods with very low h^2. When the traits were averaged over all periods, two r_A estimates were $|r_A|>1$. The r_A estimates obtained often had very high sampling errors (.16 to .94), especially when h^2 estimates

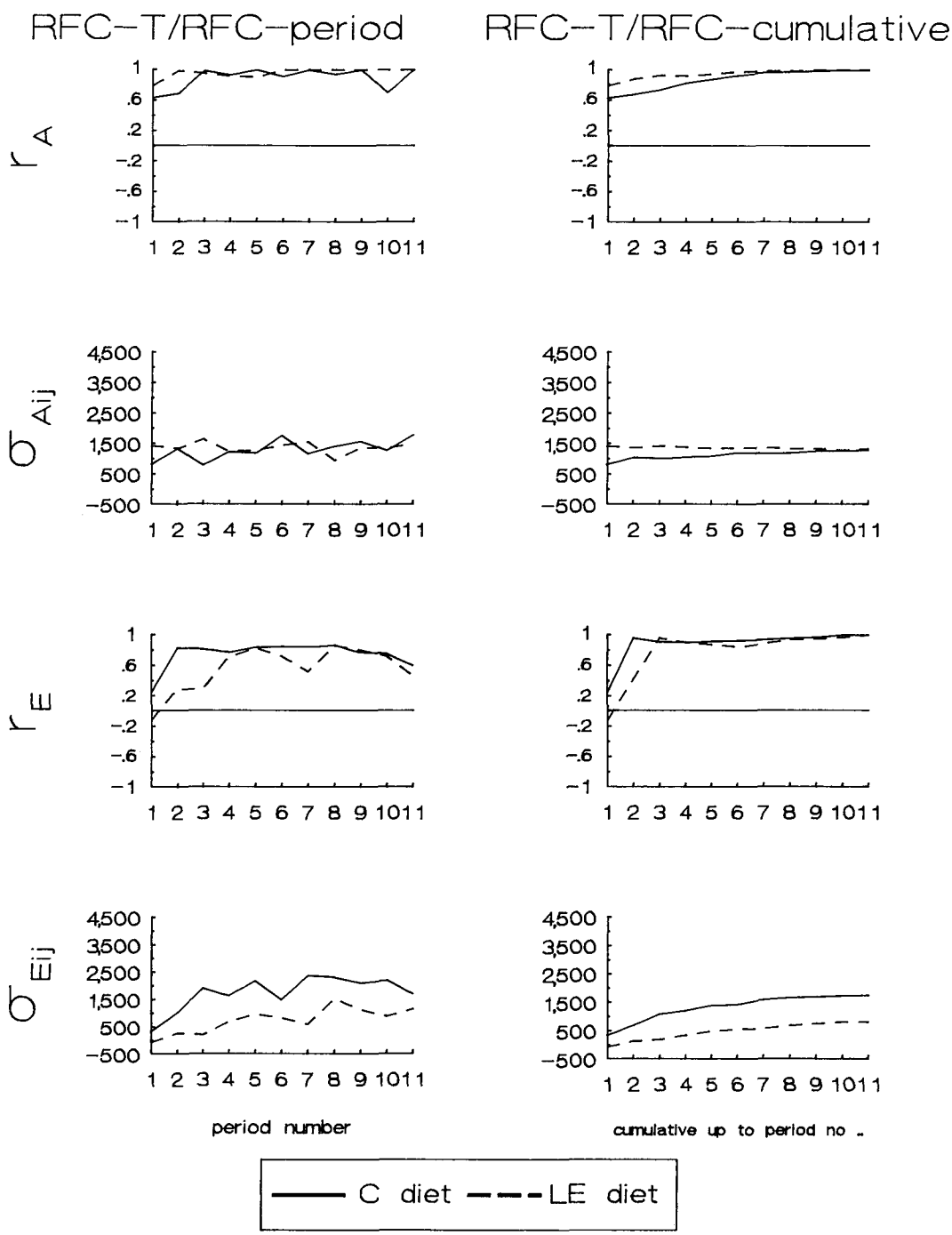


FIGURE 2. Genetic and environmental covariances (σ_{Aij} and σ_{Eij}) and correlations (r_A and r_E) of residual feed consumption (RFC) during each single period (RFC-period in kiloJoules of ME per day) with the total RFC (RFC-T in kiloJoules of ME per day); and, of the RFC accumulated over successively longer periods (RFC-cumulative in kiloJoules of ME per day) with RFC-T, on the commercial diet (C) and on the low-energy (LE) diet (11.7 and 10.0 MJ ME/kg, respectively, where 1 MJ = .239 Mcal).

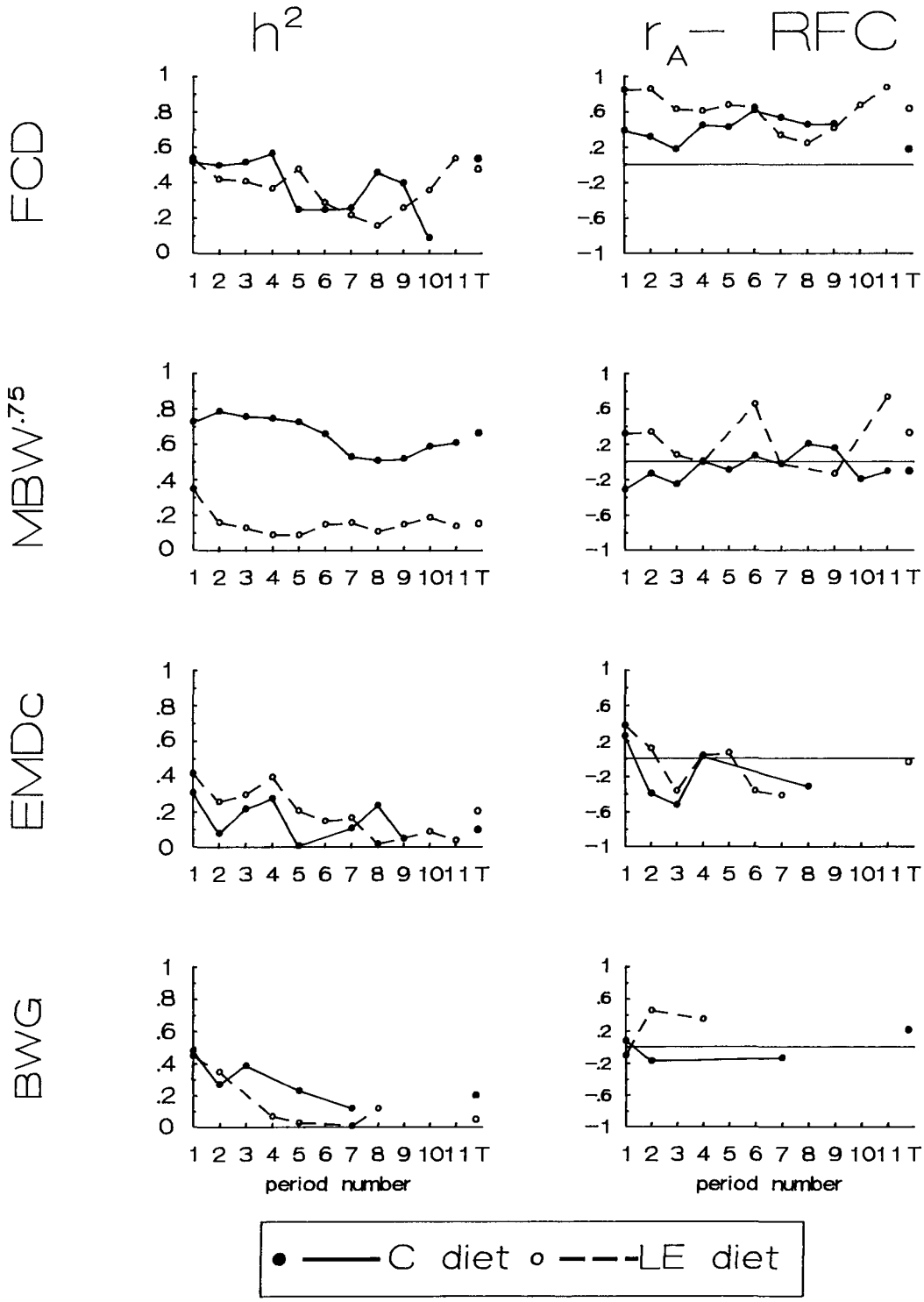


FIGURE 3. Heritabilities (h^2) of feed consumption (FCD), metabolic body weight ($MBW^{.75}$), egg mass production (EMDc), and body weight gain (BWG), and genetic correlations of residual feed consumption (r_A-RFC) with FCD, $MBW^{.75}$, EMDc, and BWG during each single period and accumulated over all periods (T), on the commercial diet (C) and on the low-energy (LE) diet (11.7 and 10.0 MJ ME/kg, respectively, where 1 MJ = .239 Mcal).

TABLE 1. *Heritabilities (h^2) of and genetic correlations (r_A) between residual feed consumption of different time segments of the two diets (commercial diet[1]: r_A and h^2 above the diagonal, with h^2 on the diagonal; low-energy diet[1]: r_A and h^2 below the diagonal)*

Variable	RFC_{20-24}[2]	RFC_{24-32}	RFC_{20-32}	RFC_{32-44}	RFC_{20-44}	RFC_{44-64}	RFC_{32-64}	RFC_{20-64}–T[3]
RFC_{20-24}[2]	.86 .55	.77	.88	.36	.69	.43	.42	.63
RFC_{24-32}	.84 .80	.57	.98	.71	.92	.43	.55	.71
RFC_{20-32}	.94 .98	.97	.65	.63	.90	.46	.54	.74
RFC_{32-44}	.71 1.00	.93	.45	.43	.91	.90	.97	.95
RFC_{20-44}	.86 1.00	.99	.98	.74	.48	.74	.83	.93
RFC_{44-64}	.66 .86	.80	1.00	.92	.43	.41	.98	.93
RFC_{32-64}	.67 .92	.86	1.00	.94	1.00	.48	.41	.97
RFC_{20-64}–T[3]	.79 .95	.93	1.00	.98	.98	.99	.62	.42

[1]Commercial and low-energy diet containing 11.7 and 10.0 MJ ME/kg, respectively, where 1 MJ = .239 Mcal.

[2]RFC_{i-j} = residual feed consumption of the time segment between i and j weeks of age.

[3]T = total recording period.

were almost zero. The r_A estimates between RFC and FCD seemed to be positive, whereas no clear values could be obtained for the ones for RFC with $MBW^{.75}$, EMDc, and BWG (at the right side of Figure 3).

DISCUSSION

The limited size of the two populations (i.e., diet groups) causes difficulties when interpreting the results. In view of the large sampling errors for the genetic parameter estimates, the derivation of variance components due to additive genetic and environmental effects is not very likely to result in consistent and meaningful figures for the two diets. Nevertheless, some tentative conclusions will be drawn by generalizing the contrasts among the time periods over the two diets, and the contrasts between diets over the various periods, so that combination with present (and future) other estimates may provide more reliable information.

Heritabilities of Residual Feed Consumption

The h^2 estimates of RFC were very high (>.5) up to about 28 wk of age in the commercial diet and up to about 32 wk in the low-energy diet. After that the h^2 estimates stabilize around .3 to .4 for both diets. Similar estimates and age trends in h^2 were found for RFC in WL populations by Bentsen (1983) from 16 to 66 wk of age (starting at about .5 and stabilizing around .2 to .3) and by Katle (1987) from 16 to 40 wk of age (starting at about .6 and stabilizing around .4), whereas the estimates from Wing and Nordskog (1982) for RFC in WL between 32 and 36 wk of age (.15 to .29) are rather low but fall within the range of the present study. In contrast, Arboleda *et al.* (1976) obtained in WL for two 4-wk periods with 8 wk in between (before and after the peak of lay) estimates of h^2 for RFC that were near zero.

A decrease in h^2 with age has often been reported for traits related to egg production, egg quality, and body weight (Clayton and Robertson, 1966; Tawfik *et al.*, 1976; Flock, 1977; Liljedahl *et al.*, 1984; Engström *et al.*, 1990). In most cases this involves an increasing σ_E^2 with age. The most probable explanation for this phenomenon is that the hens have increasing difficulties in coping with cumulative environmental stress (see Orgel, 1963; Gavora *et al.*,

1980). Liljedahl *et al.* (1984) and Engström *et al.* (1990) also found increasing σ^2_A with age for egg production. In the present study, σ^2_E increased, but σ^2_A did not increase. The steady increase in σ^2_P of RFC with age in both diets, as described by Luiting and Urff (1991b), is primarily the result of this increase in σ^2_E, because no clear pattern can be seen in σ^2_A (see Figure 1).

Accumulation of periods gave more stable and higher h^2 estimates than per period, presumably by the fact that random error on the estimates was diminishing with increasing accumulation, which also became evident from the lower σ^2_E and the similar σ^2_A levels. Accumulated h^2 estimates for the whole laying period from 20 to 64 wk of age (RFC-T) were .42 and .62 for the commercial diet and the low-energy diet, respectively. Furthermore, cumulative h^2 estimates remained high because of inclusion of the first periods up to 32 wk of age. Accumulation of periods after 32 wk of age (Table 1) gave h^2 estimates of about .4 for both diets. In most reports concerning light laying hens, similar estimates for accumulated whole laying periods were observed (Hagger and Abplanalp, 1978, .30 to .61 from 20 to 60 wk; Bentsen, 1983, .53 from 16 to 66 wk; Katle, 1987, .77 from 16 to 40 wk; Pauw, 1987, .4 from 20 to 72 wk), and accumulation after the first weeks often gave somewhat lower estimates (Hagger and Abplanalp, 1978, .22 to .64 from 40 to 60 wk; Bentsen, 1983, .35 from 34 to 66 wk).

Although the differences in h^2 estimates between diets were usually not much larger than the respective sampling errors, the general pattern seems to point to higher h^2 estimates on the low-energy diet than on the commercial diet, often as a result of a lower σ^2_E and a similar σ^2_A. This means a lower σ^2_P, which has been discussed already by Luiting and Urff (1991b). In that study it was concluded that the low energy diet results in a reduction of BWG and RFC in order to maintain a high egg production level. This effect must be largely environmental: it seems that hens that spend surplus energy for RFC-related processes simply adapt when being fed on a low-energy diet, but σ^2_A appears to be unaffected. However, the difference in h^2 estimates that was observed in the cumulative figures seemed to be caused only by differences in h^2 estimates before 32 wk of age (see also Table 1). The latter estimates on the low-energy diet were probably overestimates due to sampling errors. It was suggested by Luiting and Urff (1991b) to estimate h^2 for sdRFC as a measure of individual stability over time. This hypothesis was supported by the small positive phenotypic correlation (r_P) between sdRFC and RFC-T (Luiting and Urff, 1991b). The genetic parameter estimates from the present study were rather low and not very conclusive because of the high sampling errors.

Genetic Correlations of Residual Feed Consumption Over Time

On both diets, the time trend in r_P of RFC during each single period with RFC-T, as described by Luiting and Urff (1991b), was similar to the time trend in r_E given in the present study. The values of r_E and r_P were very similar after the first few periods. It follows that r_A estimates were very close to unity with the exception of the first periods. Also it seemed that the small difference in r_P between the two diets, as described by Luiting and Urff (1991b), was primarily the result of a small difference in r_E. Furthermore, the conclusion of diminishing random error on the estimates of cumulative variances (Figure 1), leading to more stable-looking graphs, holds here as well for the cumulative covariances and correlations. On the commercial diet, r_A estimates between the first three periods and the total accumulated RFC (.6 to .8) were somewhat lower than in later periods (almost unity). For birds fed the low-energy diet, this lower value was only observed for the first period. Table 1 also shows that r_A estimates between RFC measured up to 32 wk of age and later measurements were relatively low. Together with the time trend in h^2, this indicated that the genetic sources causing variation in RFC during the first part of lay differ from those causing variation later on, and were of less importance during the rest of the laying period. Clayton and Robertson (1966) drew the same conclusions with regard to body weight (gain) and egg weight. Flock (1977) observed the same phenomenon with egg production, both on part-time (8-wk periods) and cumulative records. Liljedahl (1989) interpreted low r_A estimates

(.58 to .74) between early and later egg production data as an indication of genotype by environment interaction. In this interpretation, age is considered as an internal environment. Following Liljedahl's reasoning, in hens used in the present experiment about 29% of gene action responsible for RFC would be common to the two ages (RFC_{20-32} versus RFC_{32-64}, Table 1) on the commercial diet.

Genetic Correlations of Residual Feed Consumption with Economic Traits

Just as r_P of RFC with $MBW^{.75}$, EMDc, and BWG per period, estimates of r_A did not clearly differ from zero. The r_A between RFC and FCD per period seem to be positive just as r_P (square root of $1 - R^2$; see Luiting and Urff, 1991b); on the commercial diet they have the same magnitude as r_P, whereas on the low-energy diet they were somewhat higher. For breeding purposes it would be better to estimate RFC by using genetic rather than phenotypic parameters. Of course, this would require estimates of r_A per period, which demand data from large numbers of related animals. This is the reason why the estimates are scarce in the literature or have large sampling errors. The fact that no clear differences could be found in the present study between the estimates of r_P and r_A of RFC with $MBW^{.75}$, EMDc, and BWG per period, suggests that calculation of RFC by phenotypic multiple regression analyses is an acceptable alternative in a breeding program if no reliable estimates of r_A are available.

In conclusion, the h^2 of the total accumulated RFC-T was .42, and the σ_P was 4.7 g/day (Luiting and Urff, 1991b) on the commercial diet. The expected response from direct selection of nucleus females for a low RFC, with selection intensity i = 1, would be about 2.0 g/day on the nucleus level. Because of the high h^2 and r_A with RFC-T, recording during a short time segment would still result in a high expected selection response: one period of 4 wk after 32 wk of age (h^2 about .3 and r_A with RFC-T about .95) would result in 1.6 g/day, three successive periods from 32 to 44 wk of age (h^2 .43 and r_A with RFC-T .95) would result in 1.9 g/day. Measuring during one or a few periods would improve the ratio between savings and costs of RFC selection. A similar measurement period was recommended for phenotypic studies, based on the estimated repeatabilities (Luiting and Urff, 1991b).

REFERENCES

Arboleda, C. R., D. L. Harris, and A. W. Nordskog, 1976. Efficiency of selection in layer type chickens by using supplementary information on feed consumption. 2. Application to net income. Theor. Appl. Genet. 48: 75–83.

Bentsen, H. B., 1983. Genetic variation in feed efficiency of laying hens at constant body weight and egg production. 2. Sources of variation in feed consumption. Acta Agric. Scand. 33:305–320.

Clayton, G. A., and A. Robertson, 1966. Genetics of changes in economic traits during the laying year. Br. Poult. Sci. 7:143–151.

Engström, G., L. E. Liljedahl, M. Wilhelmson, and K. Johansson, 1990. Genetic and environmental variation in relation to aging. Pages 77–80 *in*: Proc. 4th World Congress on Genetics Applied to Livestock Production. Vol. XVI. Edinburgh, Scotland.

Flock, D. K., 1977. Prediction of full-year egg production from part records. J. Anim. Breed. Genet. 94:89–103.

Gavora, J. S., J. L. Spencer, R. S. Gowe, and D. L. Harris, 1980. Lymphoid leukosis virus infection: effects on production and mortality and consequences in selection for high egg production. Poultry Sci. 59: 2165–2178.

Hagger, C., and H. Abplanalp, 1978. Food consumption records for the genetic improvement of income over feed costs in laying flocks of White Leghorn. Br. Poult. Sci. 19:651–667.

Katle, J., 1987. Results and informations from the Norwegian selection experiment on feed efficiency in laying hens. Pages 81–92 *in*: Proc. 103th Nordiske Jordbrugsforskeres Forening seminar "Genetiske forskelle i foderudnyttelse hos høns". Nordiske Jordbrugsforskeres Forening Utredninger-Rapporter No. 35. NJF, Oslo, Norway.

Liljedahl, L. E., 1989. Genotype × environmental interaction with age as an internal environmental factor. Pages 109–116 *in*: Genotype × Environment Interactions in Poultry Production. Les Colloques de l'Institut National de Recherches Agronomique No. 50. P. Mérat, ed. INRA, Paris, France.

Liljedahl, L. E., J. S. Gavora, R. W. Fairfull, and R. S. Gowe, 1984. Age changes in genetic and environmental variation in laying hens. Theor. Appl. Genet. 67: 391–401.

Luiting, P., and E. M. Urff, 1991a. Optimization of a model to estimate residual feed consumption in the laying hen. Livest. Prod. Sci. 27:321–338.

Luiting, P., and E. M. Urff, 1991b. Residual feed consumption in laying hens. 1. Quantification of phenotypic variation and repeatabilities. Poultry Sci. 70:1655–1662.

Manson, J. M., 1972. Body weight, egg weight and egg production. Pages 115–137 *in*: Egg Formation and Production. B. M. Freeman and P. E. Lake, ed. Longman, Br. Poult. Sci., Edinburgh, Scotland.

Meyer, K., 1986. Restricted maximum likelihood estimation of genetic parameters—in practice. Pages 454–459 *in*: Proc. 3rd World Congress on Genetics Applied to Livestock Production. Vol. XII. University of Nebraska, Lincoln, NB.

Morris, T. R., 1972. Prospects for improving the efficiency of nutrient utilisation. Pages 139–159 *in*: Egg Formation and Production. B. M. Freeman and P. E. Lake, ed. Longman, Br. Poult. Sci., Edinburgh, Scotland.

Orgel, L. E., 1963. The maintenance of the accuracy of

protein synthesis and its relevance to aging. Proc. Natl. Acad. Sci. USA 49:517–521.

Pauw, R., 1987. Analyse verschiedener Selektionskriterien zur Effizienzbewertung bei Legehennen. Ph.D. dissertation, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany.

Tawfik, E. S., P. Horst, and J. Petersen, 1976. Untersuchungen an Legehennen über genetische Fundierung und Beziehungen von Legeleistung, Legereife, Körpergewicht und Kriterien der Eibeschaffenheit. 2. Mitteilung: Phänotypische und genetische Beziehungen. Arch. Geflügelkd. 45:166–175.

Wing, T. L., and A. W. Nordskog, 1982. Use of individual records in a selection programme for egg production efficiency. 1. Heritability of the residual component of feed efficiency. Poultry Sci. 61:226–230.

Inference of Population Structure Using Multilocus Genotype Data

Jonathan K. Pritchard, Matthew Stephens and Peter Donnelly

Department of Statistics, University of Oxford, Oxford OX1 3TG, United Kingdom

Manuscript received September 23, 1999
Accepted for publication February 18, 2000

ABSTRACT

We describe a model-based clustering method for using multilocus genotype data to infer population structure and assign individuals to populations. We assume a model in which there are K populations (where K may be unknown), each of which is characterized by a set of allele frequencies at each locus. Individuals in the sample are assigned (probabilistically) to populations, or jointly to two or more populations if their genotypes indicate that they are admixed. Our model does not assume a particular mutation process, and it can be applied to most of the commonly used genetic markers, provided that they are not closely linked. Applications of our method include demonstrating the presence of population structure, assigning individuals to populations, studying hybrid zones, and identifying migrants and admixed individuals. We show that the method can produce highly accurate assignments using modest numbers of loci—*e.g.*, seven microsatellite loci in an example using genotype data from an endangered bird species. The software used for this article is available from http://www.stats.ox.ac.uk/~pritch/home.html.

IN applications of population genetics, it is often useful to classify individuals in a sample into populations. In one scenario, the investigator begins with a sample of individuals and wants to say something about the properties of populations. For example, in studies of human evolution, the population is often considered to be the unit of interest, and a great deal of work has focused on learning about the evolutionary relationships of modern populations (*e.g.*, Cavalli *et al.* 1994). In a second scenario, the investigator begins with a set of predefined populations and wishes to classify individuals of unknown origin. This type of problem arises in many contexts (reviewed by Davies *et al.* 1999). A standard approach involves sampling DNA from members of a number of potential source populations and using these samples to estimate allele frequencies in each population at a series of unlinked loci. Using the estimated allele frequencies, it is then possible to compute the likelihood that a given genotype originated in each population. Individuals of unknown origin can be assigned to populations according to these likelihoods Paetkau *et al.* 1995; Rannala and Mountain 1997).

In both situations described above, a crucial first step is to define a set of populations. The definition of populations is typically subjective, based, for example, on linguistic, cultural, or physical characters, as well as the geographic location of sampled individuals. This subjective approach is usually a sensible way of incorporating diverse types of information. However, it may be difficult to know whether a given assignment of individuals to populations based on these subjective criteria represents a natural assignment in genetic terms, and it would be useful to be able to confirm that subjective classifications are consistent with genetic information and hence appropriate for studying the questions of interest. Further, there are situations where one is interested in "cryptic" population structure—*i.e.*, population structure that is difficult to detect using visible characters, but may be significant in genetic terms. For example, when association mapping is used to find disease genes, the presence of undetected population structure can lead to spurious associations and thus invalidate standard tests (Ewens and Spielman 1995). The problem of cryptic population structure also arises in the context of DNA fingerprinting for forensics, where it is important to assess the degree of population structure to estimate the probability of false matches (Balding and Nichols 1994, 1995; Foreman *et al.* 1997; Roeder *et al.* 1998).

Pritchard and Rosenberg (1999) considered how genetic information might be used to detect the presence of cryptic population structure in the association mapping context. More generally, one would like to be able to identify the actual subpopulations and assign individuals (probabilistically) to these populations. In this article we use a Bayesian clustering approach to tackle this problem. We assume a model in which there are K populations (where K may be unknown), each of which is characterized by a set of allele frequencies at each locus. Our method attempts to assign individuals to populations on the basis of their genotypes, while simultaneously estimating population allele frequencies. The method can be applied to various types of markers [*e.g.*, microsatellites, restriction fragment length polymorphisms (RFLPs), or single nucleotide polymorphisms (SNPs)], but it assumes that the marker

Corresponding author: Jonathan Pritchard, Department of Statistics, University of Oxford, 1 S. Parks Rd., Oxford OX1 3TG, United Kingdom. E-mail: pritch@stats.ox.ac.uk

loci are unlinked and at linkage equilibrium with one another within populations. It also assumes Hardy-Weinberg equilibrium within populations. (We discuss these assumptions further in background on clustering methods and the discussion.)

Our approach is reminiscent of that taken by Smouse *et al.* (1990), who used the EM algorithm to learn about the contribution of different breeding populations to a sample of salmon collected in the open ocean. It is also closely related to the methods of Foreman *et al.* (1997) and Roeder *et al.* (1998), who were concerned with estimating the degree of cryptic population structure to assess the probability of obtaining a false match at DNA fingerprint loci. Consequently they focused on estimating the amount of genetic differentiation among the unobserved populations. In contrast, our primary interest lies in the assignment of individuals to populations. Our approach also differs in that it allows for the presence of admixed individuals in the sample, whose genetic makeup is drawn from more than one of the K populations.

In the next section we provide a brief description of clustering methods in general and describe some advantages of the model-based approach we take. The details of the models and algorithms used are given in models and methods. We illustrate our method with several examples in applications to data: both on simulated data and on sets of genotype data from an endangered bird species and from humans. incorporating population information describes how our method can be extended to incorporate geographic information into the inference process. This may be useful for testing whether particular individuals are migrants or to assist in classifying individuals of unknown origin (as in Rannala and Mountain 1997, for example). Background on the computational methods used in this article is provided in the appendix.

BACKGROUND ON CLUSTERING METHODS

Consider a situation where we have genetic data from a sample of individuals, each of whom is assumed to have originated from a *single* unknown population (no admixture). Suppose we wish to cluster together individuals who are genetically similar, identify distinct clusters, and perhaps see how these clusters relate to geographical or phenotypic data on the individuals. There are broadly two types of clustering methods we might use:

1. *Distance-based methods.* These proceed by calculating a pairwise distance matrix, whose entries give the distance (suitably defined) between every pair of individuals. This matrix may then be represented using some convenient graphical representation (such as a tree or a multidimensional scaling plot) and clusters may be identified by eye.
2. *Model-based methods.* These proceed by assuming that observations from each cluster are random draws from some parametric model. Inference for the parameters corresponding to each cluster is then done jointly with inference for the cluster membership of each individual, using standard statistical methods (for example, maximum-likelihood or Bayesian methods).

Distance-based methods are usually easy to apply and are often visually appealing. In the genetics literature, it has been common to adapt distance-based phylogenetic algorithms, such as neighbor-joining, to clustering multilocus genotype data (*e.g.*, Bowcock *et al.* 1994). However, these methods suffer from many disadvantages: the clusters identified may be heavily dependent on both the distance measure and graphical representation chosen; it is difficult to assess how confident we should be that the clusters obtained in this way are meaningful; and it is difficult to incorporate additional information such as the geographic sampling locations of individuals. Distance-based methods are thus more suited to exploratory data analysis than to fine statistical inference, and we have chosen to take a model-based approach here.

The first challenge when applying model-based methods is to specify a suitable model for observations from each cluster. To make our discussion more concrete we introduce very briefly some of our model and notation here; a fuller treatment is given later. Assume that each cluster (population) is modeled by a characteristic set of allele frequencies. Let X denote the genotypes of the sampled individuals, Z denote the (unknown) populations of origin of the individuals, and P denote the (unknown) allele frequencies in all populations. (Note that X, Z, and P actually represent multidimensional vectors.) Our main modeling assumptions are Hardy-Weinberg equilibrium within populations and complete linkage equilibrium between loci within populations. Under these assumptions each allele at each locus in each genotype is an independent draw from the appropriate frequency distribution, and this completely specifies the probability distribution $\Pr(X|Z, P)$ (given later in Equation 2). Loosely speaking, the idea here is that the model accounts for the presence of Hardy-Weinberg or linkage disequilibrium by introducing population structure and attempts to find population groupings that (as far as possible) are not in disequilibrium. While inference may depend heavily on these modeling assumptions, we feel that it is easier to assess the validity of explicit modeling assumptions than to compare the relative merits of more abstract quantities such as distance measures and graphical representations. In situations where these assumptions are deemed unreasonable then alternative models should be built.

Having specified our model, we must decide how to perform inference for the quantities of interest (Z and P). Here, we have chosen to adopt a Bayesian approach,

by specifying models (priors) Pr(Z) and Pr(P), for both Z and P. The Bayesian approach provides a coherent framework for incorporating the inherent uncertainty of parameter estimates into the inference procedure and for evaluating the strength of evidence for the inferred clustering. It also eases the incorporation of various sorts of prior information that may be available, such as information about the geographic sampling location of individuals.

Having observed the genotypes, X, our knowledge about Z and P is then given by the posterior distribution

$$\Pr(Z, P|X) \propto \Pr(Z)\Pr(P)\Pr(X|Z, P). \quad (1)$$

While it is not usually possible to compute this distribution exactly, it *is* possible to obtain an approximate sample $(Z^{(1)}, P^{(1)}), (Z^{(2)}, P^{(2)}), \ldots, (Z^{(M)}, P^{(M)})$ from Pr(Z, $P|X$) using Markov chain Monte Carlo (MCMC) methods described below (see Gilks *et al.* 1996b, for more general background). Inference for Z and P may then be based on summary statistics obtained from this sample (see *Inference for Z, P, and Q* below). A brief introduction to MCMC methods and Gibbs sampling may be found in the appendix.

MODELS AND METHODS

We now provide a more detailed description of our modeling assumptions and the algorithms used to perform inference, beginning with the simpler case where each individual is assumed to have originated in a single population (no admixture).

The model without admixture: Suppose we genotype N diploid individuals at L loci. In the case without admixture, each individual is assumed to originate in one of K populations, each with its own characteristic set of allele frequencies. Let the vector X denote the observed genotypes, Z the (unknown) populations of origin of the individuals, and P the (unknown) allele frequencies in the populations. These vectors consist of the following elements,

$(x_l^{\{i,1\}}, x_l^{\{i,2\}})$ = genotype of the ith individual at the lth locus, where $i = 1, 2, \ldots, N$ and $l = 1, 2, \ldots, L$;
$z^{(i)}$ = population from which individual i originated;
p_{klj} = frequency of allele j at locus l in population k, where $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, J_l$,

where J_l is the number of distinct alleles observed at locus l, and these alleles are labeled $1, 2, \ldots, J_l$.

Given the population of origin of each individual, the genotypes are assumed to be generated by drawing alleles independently from the appropriate population frequency distributions,

$$\Pr(x_l^{\{i,a\}} = j|Z, P) = p_{z(i)lj} \quad (2)$$

independently for each $x_l^{\{i,a\}}$. (Note that $p_{z(i)lj}$ is the frequency of allele j at locus l in the population of origin of individual i.)

Assume that before observing the genotypes we have no information about the population of origin of each individual and that the probability that individual i originated in population k is the same for all k,

$$\Pr(z^{(i)} = k) = 1/K, \quad (3)$$

independently for all individuals. (In cases where some populations may be more heavily represented in the sample than others, this assumption is inappropriate; it would be straightforward to extend our model to deal with such situations.)

We follow the suggestion of Balding and Nichols (1995) (see also Foreman *et al.* 1997 and Rannala and Mountain 1997) in using the Dirichlet distribution to model the allele frequencies at each locus within each population. The Dirichlet distribution $\mathcal{D}(\lambda_1, \lambda_2, \ldots, \lambda_J)$ is a distribution on allele frequencies $p = (p_1, p_2, \ldots, p_J)$ with the property that these frequencies sum to 1. We use this distribution to specify the probability of a particular set of allele frequencies $p_{kl\cdot}$ for population k at locus l,

$$p_{kl\cdot} \sim \mathcal{D}(\lambda_1, \lambda_2, \ldots, \lambda_{J_l}), \quad (4)$$

independently for each k,l. The expected frequency of allele j is proportional to λ_j, and the variance of this frequency decreases as the sum of the λ_j increases. We take $\lambda_1 = \lambda_2 = \cdots = \lambda_{J_l} = 1.0$, which gives a uniform distribution on the allele frequencies; alternatives are discussed in the discussion.

MCMC algorithm (without admixture): Equations 2, 3, and 4 define the quantities Pr($X|Z$, P), Pr(Z), and Pr(P), respectively. By setting $\theta = (\theta_1, \theta_2) = (Z, P)$ and letting $\pi(Z, P) = \Pr(Z, P|X)$ we can use the approach outlined in *Algorithm A1* to construct a Markov chain with stationary distribution Pr(Z, $P|X$) as follows:

Algorithm 1: *Starting with initial values $Z^{(0)}$ for Z (by drawing $Z^{(0)}$ at random using* (3) *for example), iterate the following steps for $m = 1, 2, \ldots$.*

Step 1. *Sample $P^{(m)}$ from* Pr($P|X$, $Z^{(m-1)}$).
Step 2. *Sample $Z^{(m)}$ from* Pr($Z|X$, $P^{(m)}$).

Informally, step 1 corresponds to estimating the allele frequencies for each population assuming that the population of origin of each individual is known; step 2 corresponds to estimating the population of origin of each individual, assuming that the population allele frequencies are known. For sufficiently large m and c, $(Z^{(m)}, P^{(m)})$, $(Z^{(m+c)}, P^{(m+c)})$, $(Z^{(m+2c)}, P^{(m+2c)})$, ... will be approximately independent random samples from Pr(Z, $P|X$). The distributions required to perform each step are given in the appendix.

The model with admixture: We now expand our model to allow for admixed individuals by introducing a vector Q to denote the admixture proportions for each individual. The elements of Q are

$q_k^{(i)}$ = proportion of individual i's genome that originated from population k.

It is also necessary to modify the vector Z to replace the assumption that each individual i originated in some unknown population $z^{(i)}$ with the assumption that each observed allele copy $x_l^{(i,a)}$ originated in some unknown population $z_l^{(i,a)}$:

$z_l^{(i,a)}$ = population of origin of allele copy $x_l^{(i,a)}$.

We use the term "allele copy" to refer to an allele carried at a particular locus by a particular individual.

Our primary interest now lies in estimating Q. We proceed in a manner similar to the case without admixture, beginning by specifying a probability model for (X, Z, P, Q). Analogues of (2) and (3) are

$$\Pr(x_l^{(i,a)} = j|Z, P, Q) = p_{z_l^{(i,a)}lj} \tag{5}$$

and

$$\Pr(z_l^{(i,a)} = k|P, Q) = q_k^{(i)}, \tag{6}$$

with (4) being used to model P as before. To complete our model we need to specify a distribution for Q, which in general will depend on the type and amount of admixture we expect to see. Here we model the admixture proportions $q^{(i)} = (q_1^{(i)}, \ldots, q_K^{(i)})$ of individual i using the Dirichlet distribution

$$q^{(i)} \sim \mathcal{D}(\alpha, \alpha, \ldots, \alpha) \tag{7}$$

independently for each individual. For large values of α ($\gg 1$), this models each individual as having allele copies originating from all K populations in equal proportions. For very small values of α ($\ll 1$), it models each individual as originating mostly from a single population, with each population being equally likely. As $\alpha \to 0$ this model becomes the same as our model without admixture (although the implementation of the MCMC algorithm is somewhat different). We allow α to range from 0.0 to 10.0 and attempt to learn about α from the data (specifically we put a uniform prior on $\alpha \in [0, 10]$ and use a Metropolis-Hastings update step to integrate out our uncertainty in α). This model may be considered suitable for situations where little is known about admixture; alternatives are discussed in the discussion.

MCMC algorithm (with admixture): The following algorithm may be used to sample from $\Pr(Z, P, Q|X)$.

Algorithm 2: *Starting with initial values $Z^{(0)}$ for Z (by drawing $Z^{(0)}$ at random using* (3) *for example), iterate the following steps for $m = 1, 2, \ldots$.*

Step 1. *Sample $P^{(m)}$, $Q^{(m)}$ from* $\Pr(P, Q|X, Z^{(m-1)})$.
Step 2. *Sample $Z^{(m)}$ from* $\Pr(Z|X, P^{(m)}, Q^{(m)})$.
Step 3. *Update α using a Metropolis-Hastings step.*

Informally, step 1 corresponds to estimating the allele frequencies for each population and the admixture proportions of each individual, assuming that the population of origin of each allele copy in each individual is known; step 2 corresponds to estimating the population of origin of each allele copy, assuming that the population allele frequencies and the admixture proportions are known. As before, for sufficiently large m and c, $(Z^{(m)}, P^{(m)}, Q^{(m)})$, $(Z^{(m+c)}, P^{(m+c)}, Q^{(m+c)})$, $(Z^{(m+2c)}, P^{(m+2c)}, Q^{(m+2c)})$, ... will be approximately independent random samples from $\Pr(Z, P, Q|X)$. The distributions required to perform each step are given in the appendix.

Inference: *Inference for Z, P, and Q:* We now discuss how the MCMC output can be used to perform inference on Z, P, and Q. For simplicity, we focus our attention on Q; inference for Z or P is similar.

Having obtained a sample $Q^{(1)}, \ldots, Q^{(M)}$ (using suitably large burn-in m and thinning interval c) from the posterior distribution of $Q = (q_1, \ldots, q_N)$ given X using the MCMC method, it is desirable to summarize the information contained, perhaps by a point estimate of Q. A seemingly obvious estimate is the posterior mean

$$E(q_i|X) \approx \frac{1}{M}\sum_{m=1}^{M} q_i^{(m)}. \tag{8}$$

However, the symmetry of our model implies that the posterior mean of q_i is $(1/K, 1/K, \ldots, 1/K)$ for all i, whatever the value of X. For example, suppose that there are just two populations and 10 individuals and that the genotypes of these individuals contain strong information that the first 5 are in one population and the second 5 are in the other population. Then either

$$q_1 \ldots q_5 \approx (1, 0) \quad \text{and} \quad q_6 \ldots q_{10} \approx (0, 1) \tag{9}$$

or

$$q_1 \ldots q_5 \approx (0, 1) \quad \text{and} \quad q_6 \ldots q_{10} \approx (1, 0), \tag{10}$$

with these two "symmetric modes" being equally likely, leading to the expectation of any given q_i being (0.5, 0.5). This is essentially a problem of nonidentifiability caused by the symmetry of the model [see Stephens (2000b) for more discussion].

In general, if there are K populations then there will be $K!$ sets of symmetric modes. Typically, MCMC schemes find it rather difficult to move between such modes, and the algorithms we describe will usually explore only one of the symmetric modes, even when run for a very large number of iterations. Fortunately this does not bother us greatly, since from the point of view of clustering all the symmetric modes are the same [compare the clusterings corresponding to (9) and (10)]. If our sampler explores only one symmetric mode then the sample means (8) will be very poor estimates of the posterior means for the q_i, but will be much better estimates of the *modes* of the q_i, which in this case turn out to be a much better summary of the information in the data. Ironically then, the poor mixing of the MCMC sampler between the symmetric modes gives the asymptotically useless estimator (8) some practical

value. Where the MCMC sampler succeeds in moving between symmetric modes, or where it is desired to combine results from samples obtained using different starting points (which may involve combining results corresponding to different modes), more sophisticated methods [such as those described by Stephens (2000b)] may be required.

Inference for the number of populations: The problem of inferring the number of clusters, K, present in a data set is notoriously difficult. In the Bayesian paradigm the way to proceed is theoretically straightforward: place a prior distribution on K and base inference for K on the posterior distribution

$$\Pr(K|X) \propto \Pr(X|K)\Pr(K). \tag{11}$$

However, this posterior distribution can be peculiarly dependent on the modeling assumptions made, even where the posterior distributions of other quantities (Q, Z, and P, say) are relatively robust to these assumptions. Moreover, there are typically severe computational challenges in estimating $\Pr(X|K)$. We therefore describe an alternative approach, which is motivated by approximating (11) in an *ad hoc* and computationally convenient way.

Arguments given in the appendix (*Inference on K, the number of populations*) suggest estimating $\Pr(X|K)$ using

$$\Pr(X|K) \approx \exp(-\hat{\mu}/2 - \hat{\sigma}^2/8), \tag{12}$$

where

$$\hat{\mu} = \frac{1}{M}\sum_{m=1}^{M} -2 \log \Pr(X|Z^{(m)}, P^{(m)}, Q^{(m)}) \tag{13}$$

and

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{m=1}^{M} (-2 \log \Pr(X|Z^{(m)}, P^{(m)}, Q^{(m)}) - \hat{\mu})^2. \tag{14}$$

We use (12) to estimate $\Pr(X|K)$ for each K and substitute these estimates into (11) to approximate the posterior distribution $\Pr(K|X)$.

In fact, the assumptions underlying (12) are dubious at best, and we do not claim (or believe) that our procedure provides a quantitatively accurate estimate of the posterior distribution of K. We see it merely as an *ad hoc* guide to which models are most consistent with the data, with the main justification being that it seems to give sensible answers in practice (see next section for examples). Notwithstanding this, for convenience we continue to refer to "estimating" $\Pr(K|X)$ and $\Pr(X|K)$.

APPLICATIONS TO DATA

We now illustrate the performance of our method on both simulated data and real data (from an endangered bird species and from humans). The analyses make use of the methods described in *The model with admixture.*

Simulated data: To test the performance of the clustering method in cases where the "answers" are known, we simulated data from three population models, using standard coalescent techniques (Hudson 1990). We assumed that sampled individuals were genotyped at a series of unlinked microsatellite loci. Data were simulated under the following models.

Model 1: A single random-mating population of constant size.

Model 2: Two random-mating populations of constant effective population size 2N. These were assumed to have split from a single ancestral population, also of size $2N$ at a time N generations in the past, with no subsequent migration.

Model 3: Admixture of populations. Two discrete populations of equal size, related as in model 2, were fused to produce a single random-mating population. Samples were collected after two generations of random mating in the merged population. Thus, individuals have i grandparents from population 1, and $4 - i$ grandparents from population 2 with probability $\binom{4}{i}/16$, where $i \in \{0, 4\}$. All loci were simulated independently.

We present results from analyzing data sets simulated under each model. Data set 1 was simulated under model 1, with 5 microsatellite loci. Data sets 2A and 2B were simulated under model 2, with 5 and 15 microsatellite loci, respectively. Data set 3 was simulated under model 3, with 60 loci (preliminary analyses with fewer loci showed this to be a much harder problem than models 1 and 2). Microsatellite mutation was modeled by a simple stepwise mutation process, with the mutation parameter $4N\mu$ set at 16.0 per locus (*i.e.*, the expected variance in repeat scores within populations was 8.0). We did not make use of the assumed mutation model in analyzing the simulated data.

Our analysis consists of two phases. First, we consider the issue of model choice—*i.e.*, how many populations are most appropriate for interpreting the data. Then, we examine the clustering of individuals for the inferred number of populations.

Choice of K for simulated data: For each model, we ran a series of independent runs of the Gibbs sampler for each value of K (the number of populations) between 1 and 5. The results presented are based on runs of 10^6 iterations or more, following a burn-in period of at least 30,000 iterations. To choose the length of the burn-in period, we printed out $\log(\Pr(X|P^{(m)}, Q^{(m)}))$, and several other summary statistics during the course of a series of trial runs, to estimate how long it took to reach (approximate) stationarity. To check for possible problems with mixing, we compared the estimates of $P(X|K)$ and other summary statistics obtained over several independent runs of the Gibbs sampler, starting from different initial points. In general, substantial differences between runs can indicate that either the runs should

TABLE 1

Estimated posterior probabilities of K, for simulated data sets 1, 2A, 2B, and 3 (denoted X_1, X_{2A}, X_{2B}, and X_3, respectively)

K	log $P(K\|X_1)$	$P(K\|X_{2A})$	$P(K\|X_{2B})$	$P(K\|X_3)$
1	~1.0	~0.0	~0.0	~0.0
2	~0.0	0.21	0.999	~1.0
3	~0.0	0.58	0.0009	~0.0
4	~0.0	0.21	~0.0	~0.0
5	~0.0	~0.0	~0.0	~0.0

The numbers should be regarded as a rough guide to which models are consistent with the data, rather than accurate estimates of posterior probabilities.

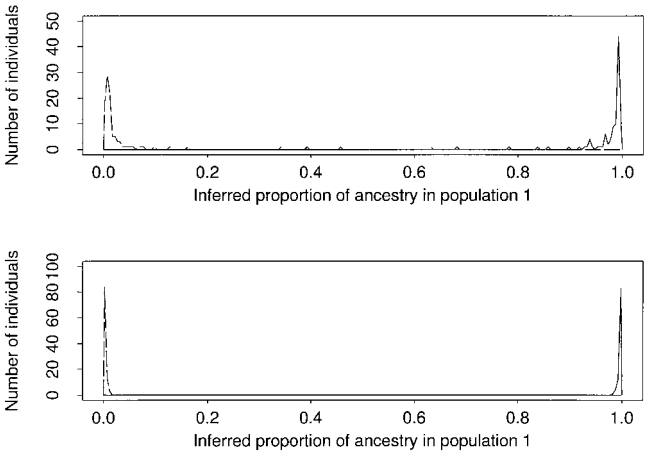


Figure 1.—Summary of the clustering results for simulated data sets 2A and 2B, respectively. For each individual, we computed the mean value of $q_1^{(i)}$ (the proportion of ancestry in population 1), over a single run of the Gibbs sampler. The dashed line is a histogram of mean values of $q_1^{(i)}$ for individuals from population 0; the solid line is for individuals from population 1.

be longer to obtain more accurate estimates or that independent runs are getting stuck in different modes in the parameter space. (Here, we consider the $K!$ modes that arise from the nonidentifiability of the K populations to be equivalent, since they arise from permuting the K population labels.)

We found that in most cases we obtained consistent estimates of $P(X|K)$ across independent runs. However, when analyzing data set 2A with $K = 3$, the Gibbs sampler found two different modes. This data set actually contains two populations, and when K is set to 3, one of the populations expands to fill two of the three clusters. It is somewhat arbitrary which of the two populations expands to fill the extra cluster: this leads to two modes of slightly different heights. The Gibbs sampler did not manage to move between the two modes in any of our runs.

In Table 1 we report estimates of the posterior probabilities of values of K, assuming a uniform prior on K between 1 and 5, obtained as described in *Inference for the number of populations*. We repeat the warning given there that these numbers should be regarded as rough guides to which models are consistent with the data, rather than accurate estimates of the posterior probabilities. In the case where we found two modes (data set 2A, $K = 3$), we present results based on the mode that gave the higher estimate of $\Pr(X|K)$.

With all four simulated data sets we were able to correctly infer whether or not there was population structure ($K = 1$ for data set 1 and $K > 1$ otherwise). In the case of data set 2A, which consisted of just 5 loci, there is not a clear estimate of K, as the posterior probability is consistent with both the correct value, $K = 2$, and also with $K = 3$ or 4. However, when the number of loci was increased to 15 (data set 2B), virtually all of the posterior probability was on the correct number of populations, $K = 2$.

Data set 3 was simulated under a more complicated model, where most individuals have mixed ancestry. In this case, the population was formed by admixture of two populations, so the "true" clustering is with $K = 2$, and Q estimating the number of grandparents from each of the two original populations, for each individual. Intuitively it seems that another plausible clustering would be with $K = 5$, individuals being assigned to clusters according to how many grandparents they have from each population. In biological terms, the solution with $K = 2$ is more natural and is indeed the inferred value of K for this data set using our *ad hoc* guide [the estimated value of $\Pr(X|K)$ was higher for $K = 5$ than for $K = 3$, 4, or 6, but much lower than for $K = 2$]. However, this raises an important point: the inferred value of K may not always have a clear biological interpretation (an issue that we return to in the discussion).

Clustering of simulated data: Having considered the problem of estimating the number of populations, we now examine the performance of the clustering algorithm in assigning particular individuals to the appropriate populations. In the case where the populations are discrete, the clustering performs very well (Figure 1), even with just 5 loci (data set 2A), and essentially perfectly with 15 loci (data set 2B).

The case with admixture (Figure 2) appears to be more difficult, even using many more loci. However, the clustering algorithm did manage to identify the population structure appropriately and estimated the ancestry of individuals with reasonable accuracy. Part of the reason that this problem is difficult is that it is hard to estimate the original allele frequencies (before admixture) when almost all the individuals (7/8) are admixed. A more fundamental problem is that it is difficult to get accurate estimates of $q^{(i)}$ for particular individuals because (as can be seen from the y-axis of Figure 2) for any given individual, the variance of how many of its alleles are *actually* derived from each population

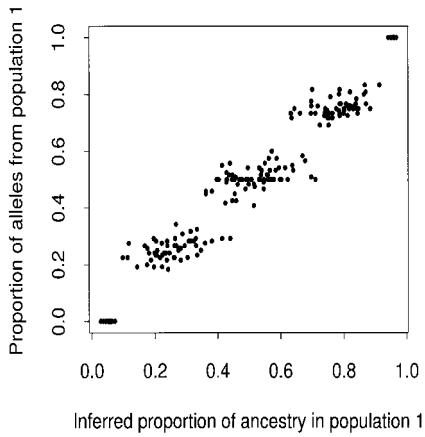


Figure 2.—Summary of the clustering results for simulated data set 3. Each point plots the estimated value of $q_1^{(i)}$ (the proportion of ancestry in population 1) for a particular individual against the fraction of their alleles that were actually derived from population 1 (across the 60 loci genotyped). The five clusters (from left to right) are for individuals with 0, 1, . . . , 4 grandparents in population 1, respectively.

can be substantial (for intermediate q). This property means that even if the allele frequencies were known, it would still be necessary to use a considerable number of loci to get accurate estimates of q for admixed individuals.

Data from the Taita thrush: We now present results from applying our method to genotype data from an endangered bird species, the Taita thrush, *Turdus helleri*. Individuals were sampled at four locations in southeast Kenya [Chawia (17 individuals), Ngangao (54), Mbololo (80), and Yale (4)]. Each individual was genotyped at seven microsatellite loci (Galbusera *et al.* 2000).

This data set is a useful test for our clustering method, because the geographic samples are likely to represent distinct populations. These locations represent fragments of indigenous cloud forest, separated from each other by human settlements and cultivated areas. Yale, which is a very small fragment, is quite close to Ngangao. Extensive data on ringed and radio-tagged birds over a 3-year period indicate low migration rates (Galbusera *et al.* 2000).

As discussed in background on clustering methods, it is currently common to use distance-based clustering methods to visualize genotype data of this kind. To permit a comparison between that type of approach and our own method, we begin by showing a neighbor-joining tree of the bird data (Figure 3). Inspection of the tree reveals that the Chawia and Mbololo individuals represent (somewhat) distinct clusters. Several individuals (marked by asterisks) appear to be classified with other groups. The four Yale individuals appear to fall within the Ngangao group [a view that is supported by summary statistics of divergence showing the Yale and Ngangao to be very closely related (Table 2)].

The tree illustrates several shortcomings of distance-

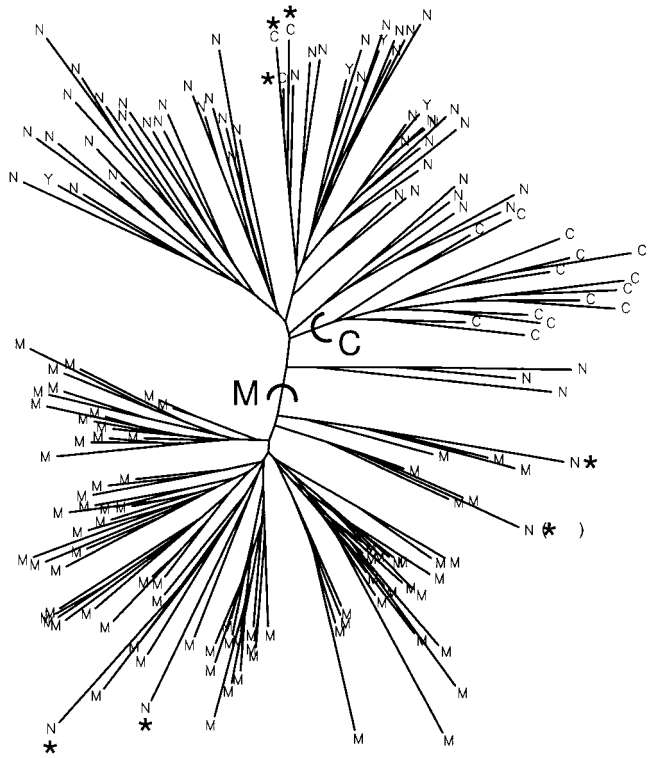


Figure 3.—Neighbor-joining tree of individuals in the *T. helleri* data set. Each tip represents a single individual. C, M, N, and Y indicate the populations of origin (Chawia, Mbololo, Ngangao, and Yale, respectively). Using the labels, it is possible to group the Chawia and Mbololo individuals into (somewhat) distinct clusters, as marked. However, it would not be possible to identify these clusters if the population labels were not available. Individuals who appear to be misclassified are marked *. One of these individuals [marked (*)] was also identified by our own algorithm as a possible migrant. The tree was constructed using the program *Neighbor* included in Phylip (Felsenstein 1993). The pairwise distance matrix was computed as follows (Mountain and Cavalli-Sforza 1997). For each pair of individuals, we added $1/L$ for each locus at which they had no alleles in common, $1/2L$ for each locus at which they had one allele in common (*e.g.*, AA:AB or AB:AC), and 0 for each locus at which they had two alleles in common (*e.g.*, AA:AA or AB:AB), where L is the number of loci compared.

TABLE 2

Summary statistics of variation within and between geographic groups

	Chawia	Mbololo	Ngangao	Yale
Chawia	5.1			
Mbololo	7.1	5.6		
Ngangao	3.1	1.6	5.5	
Yale	1.9	2.3	0.1	6.0

Diagonal, variance in repeat scores within groups; below diagonal, square of mean difference in repeat scores between populations [$(\delta\mu)^2$; Goldstein and Pollock 1997, Equation C3)].

based clustering methods. First, it would not be possible (in this case) to identify the appropriate clusters if the labels were missing. Second, since the tree does not use a formal probability model, it is difficult to ask statistical questions about features of the tree, for example: Are the individuals marked with asterisks actually migrants, or are they simply misclassified by chance? Is there evidence of population structure *within* the Ngangao group (which appears from the tree to be quite diverse)?

We now apply our clustering method to these data.

Choice of *K*, for Taita thrush data: To choose an appropriate value of K for modeling the data, we ran a series of independent runs of the Gibbs sampler at a range of values of K. After running numerous medium-length runs to investigate the behavior of the Gibbs sampler (using the diagnostics described in *Choice of K for simulated data*), we again chose to use a burn-in period of 30,000 iterations and to collect data for 10^6 iterations. We ran three to five independent simulations of this length for each K between 1 and 5 and found that the independent runs produced highly consistent results. At $K = 5$, a run of 10^6 steps takes ~70 min on our desktop machine.

Using the approach described in *Inference for the number of populations*, we estimated $\Pr(X|K)$ for $K = 1, 2, \ldots, 5$ and corresponding values of $\Pr(K|X)$ for a uniform prior on $K = 1, 2, \ldots, 5$. (In fact, this data set contains a lot of information about K, so that inference is relatively robust to choice of prior on K, and other priors, such as taking $\Pr(K)$ proportional to Poisson(1) for $K > 0$, would give virtually indistinguishable results.) From the estimates of $\Pr(K|X)$, shown in the last column of Table 3, it is clear that the models with $K = 1$ or 2 are completely insufficient to model the data and that the model with $K = 3$ is substantially better than models with larger K. Given these results, we now focus our subsequent analysis on the model with three populations.

Clustering results for Taita thrush data: Figure 4 shows a plot of the clustering results for the individuals in the sample, assuming that there are three populations (as inferred above). We did not use (and indeed, did not know) the sampling locations of individuals when we obtained these results. Our clustering algorithm seems to have performed very well, with just a few individuals (labeled 1–4) falling somewhat outside the obvious clusters. All of the points in the extreme corners (some of which may be difficult to resolve on the picture) are correctly assigned. The four Yale individuals were assigned to the Ngangao cluster, consistent with the neighbor-joining tree and the $(\delta\mu)^2$ distances. We return to this data set in incorporating population information to consider the question of whether the individuals that seem not to cluster tightly with others sampled from the same location are the product of migration.

TABLE 3

Inferring the value of *K*, the number of populations, for the *T. helleri* data

K	$\log P(X\|K)$	$P(K\|X)$
1	−3144	~0
2	−2769	~0
3	−2678	0.993
4	−2683	0.007
5	−2688	0.00005

The values in the last column assume a uniform prior for K ($K \in \{1, \ldots, 5\}$).

Application to human data: The next data set, taken from Jorde *et al.* (1995), includes data from 30 biallelic restriction site polymorphisms, genotyped in 72 Africans (Sotho, Tsonga, Nguni, Biaka and Mbuti Pygmies, and San) and 90 Europeans (British and French).

Application of our MCMC scheme with $K = 2$ indicates the presence of two very distinct clusters, corresponding to the Africans and Europeans in the sample (Figure 5). The model with $K = 2$ has vastly higher posterior probability than the model with $K = 1$.

Additional runs of the MCMC scheme with the models $K = 3$, 4, and 5 suggest that those models may be somewhat better than $K = 2$. This may reflect the presence of population structure within the continental groupings, although in this case the additional populations do not form discrete clusters and so are difficult to interpret.

Again it is interesting to contrast our clustering results with the neighbor-joining tree of these data (Figure 6). While our method finds it quite easy to separate the two continental groups into the correct clusters, it would not be possible to use the neighbor-joining tree to detect distinct clusters if the labels were not present. The data set of Jorde also contains a set of individuals of Asian origin (which are more closely related to Europeans than are Africans). Neither the neighbor-joining method nor our method differentiates between the Europeans and Asians with great accuracy using this data set.

INCORPORATING POPULATION INFORMATION

The results presented so far have focused on testing how well our method works. We now turn our attention to some further applications of this method.

Our clustering results (Figure 4) confirm that the three main geographic groupings in the thrush data set (Chawia, Mbololo, and Ngangao) represent three genetically distinct populations. The geographic labels correspond very closely to the genetic clustering in all but a handful of cases (1–4 in Figure 4). Individual 2 is also identified as a possible outlier on the neighbor-joining tree (Figure 3). Given this, it is natural to ask whether these apparent outliers are immigrants (or de-

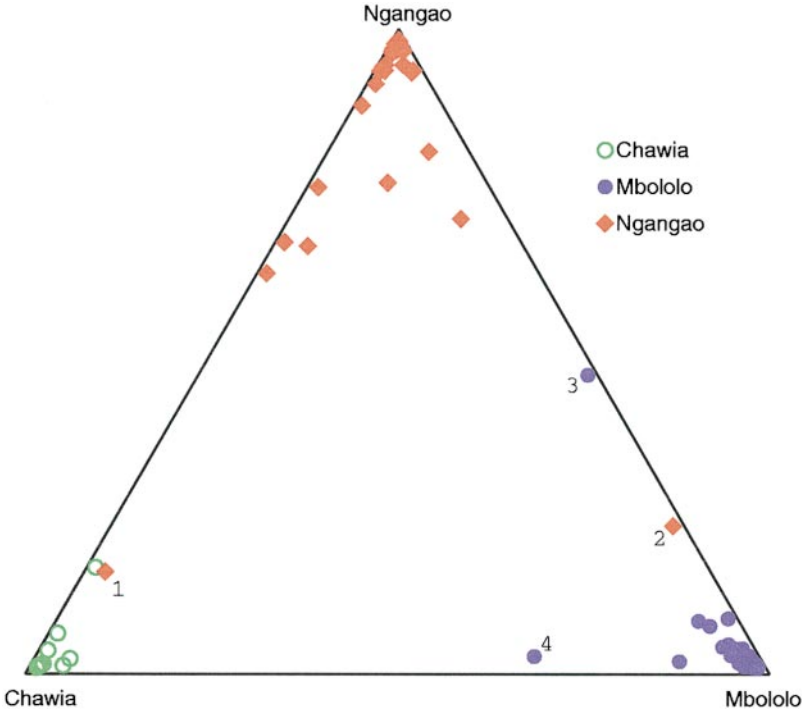


Figure 4.—Summary of the clustering results for the *T. helleri* data assuming three populations. Each point shows the mean estimated ancestry for an individual in the sample. For a given individual, the values of the three coefficients in the ancestry vector $q^{(i)}$ are given by the distances to each of the three sides of the equilateral triangle. After the clustering was performed, the points were labeled according to sampling location. Numbers 1–4 are individuals who appear to be possible outliers (see text). For clarity, the four Yale individuals (who fall into the Ngangao cluster) are not plotted. We were not told the sampling locations of individuals until after we obtained these results.

scendants of recent immigrants) from other populations. For example, given the genetic data, how probable is it that individual 1 is actually an immigrant from Chawia?

To answer this sort of question, we need to extend our algorithm to incorporate the geographic labels. By doing this, we break the symmetry of the labels, and we can ask specifically whether a particular individual is a migrant from Chawia (say). In essence our approach (described more formally in the next section) is to assume that each individual originated, with high probability, in the geographical region in which it was sampled, but to allow some small probability that it is an immigrant (or has immigrant ancestry). Note that this model is also suitable for situations in which individuals are classified according to some characteristic other than sampling location (physical appearance, for example). "Immigrants" in this situation would be individuals whose genetic makeup suggests they were misclassified. Thus, while we speak of "immigrants" and "immigrant ancestry," in some contexts these terms may relate to something other than changes in physical location.

Provided that geographic labels *usually* correspond to population membership, using the geographic infor-

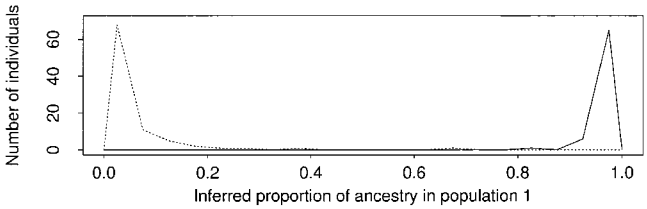


Figure 5.—Summary of the clustering results for the data set of Africans and Europeans taken from Jorde *et al.* (1995). For each individual, we computed the mean value of $q_1^{(i)}$ (the proportion of ancestry in population 1), over a single run of the Gibbs sampler. The dashed line is a histogram of mean values of $q_1^{(i)}$ for individuals of European origin; the solid line is for individuals of African origin.

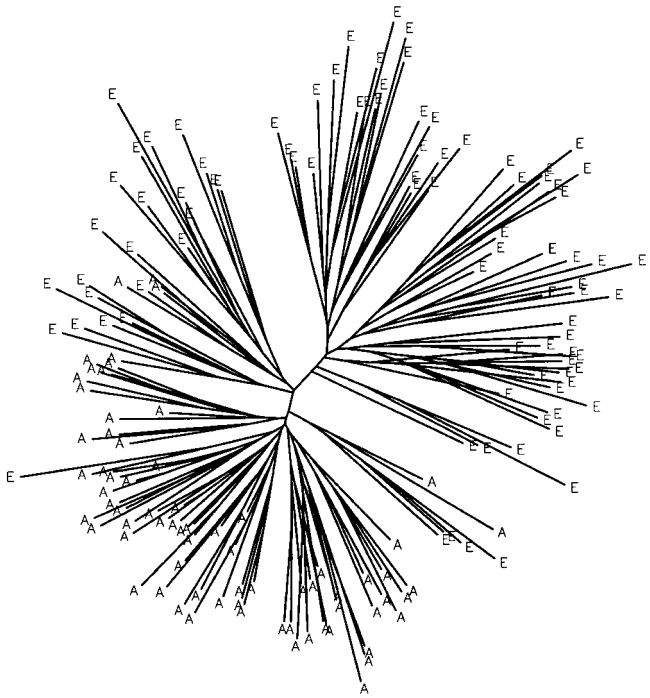


Figure 6.—Neighbor-joining tree of individuals in the data set of Jorde *et al.* (1995). Each tip represents a single individual. A and E indicate that individuals were African or European, respectively. The tree was constructed as in Figure 3.

mation will clearly improve our accuracy at assigning individuals to clusters; it will also improve our estimates of P, thus also giving us greater precision in assignment of individuals who do not have geographic information. However, in practice we suggest that before making use of such information, users of our method should first cluster the data without using the geographic labels, to check that the genetically defined clusters do in fact agree with geographic labels. We return to this issue in the discussion.

Rannala and Mountain (1997) also considered the problem of detecting immigrants and individuals with recent immigrant ancestors, taking a somewhat similar approach to that used here. However, rather than considering all individuals simultaneously, as we do here, they test each individual in the sample, one at a time, as a possible immigrant, assuming that all the other individuals are not immigrants. This approach will have reduced power to detect immigrants if the sample contains several immigrants from one population to another. In contrast, our approach can cope well with this kind of situation.

Model with prior population information: To incorporate geographic information, we use the following model. Our primary goal is to identify individuals who are immigrants, or who have recent immigrant ancestry, in the last G generations, say, where $G = 0$ is the present generation. [In practice there will only be substantial power to detect immigration for small G; *cf.* Rannala and Mountain (1997).]

First, we code each of the geographic locations by a (unique) integer between 1 and K, where K would usually be set equal to the number of locations. Using this coding, let $g^{(i)}$ represent the geographic sampling location of individual i. Now, let ν be the probability that an individual is an immigrant to population $g^{(i)}$ or has an immigrant ancestor in the last G generations. Otherwise, with probability $1 - \nu$, the individual is considered to be purely from population $g^{(i)}$. While in principle one could place a prior on ν and learn about it from the data as part of the MCMC scheme, in our current implementation the user must specify a fixed value for ν; we give some guidelines in the next section.

Assuming that migration is rare, we can use the approximation that each individual has at most one immigrant ancestor in the last G generations (where G is suitably small). Then, assuming a constant migration rate, the probability of an immigrant ancestor in generation t ($0 \le t \le G$) is proportional to 2^t, where $t = 0$ indicates that the individual migrated in the present generation. Thus, we set the prior on $q^{(i)}$ to be

$$q^{(i)}_{g^{(i)}} = 1, \quad q^{(i)}_k = 0 \quad (k \neq g^{(i)}) \tag{15}$$

with probability $1 - \nu$ and

$$q^{(i)}_{g^{(i)}} = 1 - 2^{-t},\ q^{(i)}_j = 2^{-t},\ q^{(i)}_k = 0 \quad (k \neq g^{(i)}, j) \tag{16}$$

for each $j \neq g^{(i)}$ with probability

$$\frac{2^t \nu}{(K-1)\Sigma_{T=0}^{G} 2^T}, \tag{17}$$

where $t \in \{0, \ldots, G\}$. As before, $q^{(i)}_l \ge 0$ for $l \in \{1, \ldots, K\}$, and $\Sigma q^{(i)}_l = 1$.

Again, we can sample from $\Pr(Q|X)$ using *Algorithm 2*. In this case, however, since there are a small number of possible values of $q^{(i)}$, we update $q^{(i)}$ by sampling directly from the posterior probability of $q^{(i)}|X,P$, rather than conditional on Z.

Note that in this framework, it is easy to include individuals for whom there is no geographic information by using the same prior and update steps as before (Equations 7 and A10).

Testing for migrants in the Taita thrush data: To apply our method, we must first specify a value for ν. In this case, based on mark-release-recapture data from these populations (Galbusera *et al.* 2000), migration seems relatively rare, and so ν is likely to be small. We performed analyses for $\nu = 0.05$ and $\nu = 0.1$; a summary of the results is shown in Table 4. Individuals 2 and 3 have moderate posterior probabilities of having migrant ancestry, but these probabilities are perhaps smaller than might be expected from examining Figure 4. This is due to a combination of the low prior probability for migration (from the mark-release-recapture data) and, perhaps more importantly, the fact that there is a limited amount of information in seven loci, so that the uncertainty associated with the position of the points marked 1, 2, 3, and 4 in Figure 4 may be quite large. A more definite conclusion could be obtained by typing more loci.

It is interesting to note that our conclusions here differ from those obtained on this data set using the package IMMANC (Rannala and Mountain 1997). IMMANC indicates that three individuals (1, 2, and 3 here) show significant evidence of immigrant ancestry at the 0.01 significance level (Galbusera *et al.* 2000). However, IMMANC does not make a multiple comparisons correction; such a correction would bring those results into line with ours.

We anticipate that our method might also be applied in situations where there is little data to help make an informed choice of ν. In such situations we suggest analyzing the data using several different values of ν, to see whether the conclusions are robust to choice of ν. The range of sensible values for ν will depend on the context, but typically we suggest values in the range 0.001–0.1 might be appropriate. Sensitivity to choice of ν indicates that the amount of information in the data is insufficient to draw strong conclusions.

DISCUSSION

We have described a method for using multilocus genotype data to learn about population structure and assign individuals (probabilistically) to populations.

TABLE 4

Testing whether particular individuals are immigrants or have recent immigrant ancestors

Individual	Geographic origin	Possible source	ν	No immigrant ancestry	Immigrant	Immigrant parent	Immigrant grandparent
1	Ngangao	Chawia	0.05	0.869	0.008	0.052	0.063
			0.10	0.739	0.019	0.106	0.123
2	Ngangao	Mbololo	0.05	0.673	0.029	0.126	0.168
			0.10	0.472	0.046	0.203	0.273
3	Mbololo	Ngangao	0.05	0.649	0.002	0.179	0.165
			0.10	0.464	0.003	0.271	0.253
4	Mbololo	Chawia	0.05	0.891	0.000	0.007	0.082
			0.10	0.791	0.000	0.014	0.157

The individuals are labeled as shown in Figure 4. "No immigrant ancestry" gives the probability that the ancestry of each individual is exclusively in the geographic origin population; the following columns show the probabilities that each individual has the given amount of ancestry in the possible source population. The rows do not add to 1 because there are small probabilities associated with individuals having ancestry in the third population.

Our method also provides a novel approach to testing for the presence of population structure ($K > 1$).

Our examples demonstrate that the method can accurately cluster individuals into their appropriate populations, even using only a modest number of loci. In practice, the accuracy of the assignments depends on a number of factors, including the number of individuals (which affects the accuracy of the estimate for P), the number of loci (which affects the accuracy of the estimate for Q), the amount of admixture, and the extent of allele-frequency differences among populations.

We anticipate that our method will be useful for identifying populations and assigning individuals in situations where there is little information about population structure. It should also be useful in problems where cryptic population structure is a concern, as a way of identifying subpopulations. Even in situations where there is nongenetic information that can be used to define populations, it may be useful to use the approach developed here to ensure that populations defined on an extrinsic basis reflect the underlying genetic structure.

As described in incorporating population information we have also developed a framework that makes it possible to combine genetic information with prior information about the geographic sampling location of individuals. Besides being used to detect migrants, this could also be used in situations where there is strong prior population information for some individuals, but not for others. For example, in hybrid zones it may be possible to identify some individuals who do not have mixed ancestry and then to estimate q for the rest (M. Beaumont, D. Gotelli, E. M. Barett, A. C. Kitchener, M. J. Daniels, J. K. Pritchard and M. W. Bruford, unpublished results). The advantage of using a clustering approach in such cases is that it makes the method more robust to the presence of misclassified individuals and should be more accurate than if only preclassified individuals are used to estimate allele frequencies (*cf.* Smouse *et al.* 1990).

Another type of application where the geographic information might be of value is in evolutionary studies of population relationships. Such analyses frequently make use of summary statistics based on population allele frequencies [*e.g.*, F_{ST} and $(\delta\mu)^2$]. In situations where the population allele frequencies might be affected by recent immigration or where population classifications are unclear, such summary statistics could be calculated directly from the population allele frequencies P estimated by the Gibbs sampler.

There are several ways in which the basic model that we have described here might be modified to produce better performance in particular cases. For example, in models and methods and applications to data we assumed relatively noninformative priors for q. However, in some situations, there might be quite a bit of information about likely values of q, and the estimation procedure could be improved by using that information. For example, in estimating admixture proportions for African Americans, it would be possible to improve the estimation procedure by making use of existing information about the extent of European admixture (*e.g.*, Parra *et al.* 1998).

A second way in which the basic model can be modified involves changing the way in which the allele frequencies P are estimated. Throughout this article, we have assumed that the allele frequencies in different populations are uncorrelated with one another. This is a convenient approximation for populations that are not extremely closely related and, as we have seen, can produce accurate clustering. However, loosely speaking, the model of uncorrelated allele frequencies says that we do not normally expect to see populations with very similar allele frequencies. This property has the result that the clustering algorithm may tend to merge subpopulations that share similar frequencies. An alternative,

which we have implemented in our software package, is to permit allele frequencies to be correlated across populations (appendix, *Model with correlated allele frequencies*). In a series of additional simulations, we have found that this allows us to perform accurate assignments of individuals in very closely related populations, though possibly at the cost of making us likely to overestimate *K*.

Our basic model might also be modified to allow for linkage among marker loci. Normally, we would not expect to see linkage disequilibrium within subpopulations, except between markers that are extremely close together. This means that in situations where there is little admixture, our assumption of independence among loci will be quite accurate. However, we might expect to see strong correlations among linked loci when there is recent admixture. This occurs because an individual who is admixed will inherit large chromosomal segments from one population or another. Thus, when the map order of marker loci is known, it should be possible to improve the accuracy of the estimation for such individuals by modeling the inheritance of these segments.

In this article we have devoted considerable attention to the problem of inferring *K*. This is an important practical problem from the standpoint of model choice. We need to have some way of deciding which clustering model is most appropriate for interpreting the data. However, we stress that care should be taken in the interpretation of the inferred value of *K*. To begin with, due to the very high dimensionality of the parameter space, we found it difficult to obtain reliable estimates of $\Pr(X \mid K)$ and have chosen to use a fairly *ad hoc* procedure that we have found gives sensible results in practice. Second, it has been observed that in Bayesian model-based clustering, the posterior distribution of *K* tends to be quite dependent on the priors and modeling assumptions, even though estimates of the other parameters (*e.g.*, *P* and *Q* here) may be reasonably robust (see Richardson and Green 1997; Stephens 2000a, for example).

There are also biological reasons to be careful interpreting *K*. The population model that we have adopted here is obviously an idealization. We anticipate that it will be flexible enough to permit appropriate clustering for a wide range of population structures. However, as we pointed out in our discussion of data set 3 (*Choice of K for simulated data*), clusters may not necessarily correspond to "real" populations. As another example, imagine a species that lives on a continuous plane, but has low dispersal rates, so that allele frequencies vary continuously across the plane. If we sample at *K* distinct locations, we might infer the presence of *K* clusters, but the inferred number *K* is not *biologically* interesting, as it was determined purely by the sampling scheme. All that can usefully be said in such a situation is that the migration rates between the sampling locations are not high enough to make the population act as a single unstructured population.

In summary, we find that the method described here can produce highly accurate clustering and sensible choices of *K*, both for simulated data and for real data from humans and from the Taita thrush. In the latter example, we find it particularly encouraging that using a relatively small number of loci (seven) we can detect a very strong signal of population structure and assign individuals appropriately.

The algorithms described in this article have been implemented in a computer software package *structure*, which is available at http://www.stats.ox.ac.uk/~pritch/home.html.

We thank Peter Galbusera and Lynn Jorde for allowing us to use their data, Augie Kong for a helpful discussion, Daniel Falush for suggesting comparison with neighbor-joining trees, Steve Brooks and Trevor Sweeting for helpful discussions on inferring *K*, and Eric Anderson for his extensive comments on an earlier version of the manuscript. This work was supported by National Institutes of Health grant GM19634 and by a Hitchings-Elion fellowship from Burroughs-Wellcome Fund to J.K.P., by a grant from the University of Oxford and a Wellcome Trust Fellowship (057416) to M.S., and by grants GR/M14197 and 43/MMI09788, from the Engineering and Physical Sciences Research Council and Biotechnology and Biological Sciences Research Council, respectively, to P.D. The work was initiated while the authors were resident at the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK.

LITERATURE CITED

Balding, D. J., and R. A. Nichols, 1994 DNA profile match probability calculations: how to allow for population stratification, relatedness, database selection and single bands. Forensic Sci. Int. **64:** 125–140.

Balding, D. J., and R. A. Nichols, 1995 A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. Genetica **96:** 3–12.

Bowcock, A. M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. Kidd *et al.*, 1994 High resolution of human evolutionary trees with polymorphic microsatellites. Nature **368:** 455–457.

Cavalli-Sforza, L. L., P. Menozzi and A. Piazza, 1994 *The History and Geography of Human Genes*. Princeton University Press, Princeton, NJ.

Chib, S., 1995 Marginal likelihood from the Gibbs output. J. Am. Stat. Assoc. **90:** 1313–1321.

Chib, S., and E. Greenberg, 1995 Understanding the Metropolis-Hastings algorithm. Am. Stat. **49:** 327–335.

Davies, N., F. X. Villablanca and G. K. Roderick, 1999 Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. TREE **14:** 17–21.

DiCiccio, T., R. Kass, A. Raftery and L. Wasserman, 1997 Computing Bayes factors by posterior simulation and asymptotic approximations. J. Am. Stat. Assoc. **92:** 903–915.

Ewens, W. J., and R. S. Spielman, 1995 The transmission/disequilibrium test: history, subdivision, and admixture. Am. J. Hum. Genet. **57:** 455–464.

Felsenstein, J., 1993 PHYLIP (phylogeny inference package) version 3.5c. Technical report, Department of Genetics, University of Washington, Seattle.

Foreman, L., A. Smith and I. Evett, 1997 Bayesian analysis of DNA profiling data in forensic identification applications. J. R. Stat. Soc. A **160:** 429–469.

Galbusera, P., L. Lens, E. Waiyaki, T. Schenck and E. Mattysen, 2000 Effective population size and gene flow in the globally,

critically endangered Taita thrush, *Turdus helleri*. Conserv. Genet. (in press).
Gilks, W. R., S. Richardson and D. J. Spiegelhalter, 1996a Introducing Markov chain Monte Carlo, pp. 1–19 in *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson and D. J. Spiegelhalter. Chapman & Hall, London.
Gilks, W. R., S. Richardson and D. J. Spiegelhalter (Editors), 1996b *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.
Goldstein, D. B., and D. Pollock, 1997 Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. J. Hered. **88:** 335–342.
Green, P. J., 1995 Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika **82:** 711–732.
Hudson, R. R., 1990 Gene genealogies and the coalescent process, pp. 1–44 in *Oxford Surveys in Evolutionary Biology*, Vol. 7, edited by D. Futuyma and J. Antonovics. Oxford University Press, Oxford.
Jorde, L. B., M. J. Bamshad, W. S. Watkins, R. Zenger, A. E. Fraley *et al.*, 1995 Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am. J. Hum. Genet. **57:** 523–538.
Mountain, J. L., and L. L. Cavalli-Sforza, 1997 Multilocus genotypes, a tree of individuals, and human evolutionary history. Am. J. Hum. Genet. **61:** 705–718.
Paetkau, D., W. Calvert, I. Stirling and C. Strobeck, 1995 Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. **4:** 347–354.
Parra, E. J., A. Marcini, J. Akey, J. Martinson, M. A. Batzer *et al.*, 1998 Estimating African American admixture proportions by use of population-specific alleles. Am. J. Hum. Genet. **63:** 1839–1851.
Pritchard, J. K., and N. A. Rosenberg, 1999 Use of unlinked genetic markers to detect population stratification in association studies. Am. J. Hum. Genet. **65:** 220–228.
Raftery, A. E., 1996 Hypothesis testing and model selection, pp. 163–188 in *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson and D. J. Spiegelhalter. Chapman & Hall, London.
Rannala, B., and J. L. Mountain, 1997 Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA **94:** 9197–9201.
Richardson, S., and P. J. Green, 1997 On Bayesian analysis of mixtures with an unknown number of components. J. R. Stat. Soc. Ser. B **59:** 731–792.
Roeder, K., M. Escobar, J. B. Kadane and I. Balazs, 1998 Measuring heterogeneity in forensic databases using hierarchical Bayes models. Biometrika **85:** 269–287.
Smouse, P. E., R. S. Waples and J. A. Tworek, 1990 A genetic mixture analysis for use with incomplete source population-data. Can. J. Fish. Aquat. Sci. **47:** 620–634.
Spiegelhalter, D. J., N. G. Best and B. P. Carlin, 1999 Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models. Available from http://www.mrc-bsu.cam.ac.uk/publications/preslid.shtml.
Stephens, M., 2000a Bayesian analysis of mixtures with an unknown number of components—an alternative to reversible jump methods. Ann. Stat. (in press).
Stephens, M., 2000b Dealing with label-switching in mixture models. J. R. Stat. Soc. Ser. B (in press).

Communicating editor: M. K. Uyenoyama

APPENDIX

MCMC methods and Gibbs sampling:

MCMC methods are extremely useful for obtaining (approximate) samples from a probability distribution, $\pi(\theta)$, say, which cannot be simulated from directly [in our case $\theta = (Z, P, Q)$ and $\pi(\theta) = \Pr(Z, P, Q|X)$]. The idea is to construct a Markov chain $\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \ldots$ with stationary distribution $\pi(\theta)$. This is often surprisingly straightforward using standard methods devised for this purpose, such as the Metropolis-Hastings algorithm (*e.g.*, Chib and Greenberg 1995) and Gibbs sampling (*e.g.*, Gilks *et al.* 1996a), which we describe in more detail below. Intuitively, if the Markov chain $\theta^{(0)}, \theta^{(1)}, \theta^{(2)}, \ldots$ has stationary distribution $\pi(\theta)$, then $\theta^{(m)}$ will be approximately distributed as $\pi(\theta)$ provided m is sufficiently large. This can be formalized and shown to be true provided the Markov chain satisfies certain technical conditions (*ergodicity*) that hold for the Markov chains considered in this article. Furthermore, for sufficiently large c, $\theta^{(m)}, \theta^{(m+c)}, \theta^{(m+2c)}, \ldots$ will be reasonably independent samples from $\pi(\theta)$. The value of m used is often referred to as the *burn-in* period of the chain; c is often referred to as the *thinning* interval.

In general it is very difficult to know how large m and c should be. The values required to obtain reliable results depend heavily on the amount of correlation between successive states of the Markov chain. If successive states are relatively uncorrelated (that is, if the chain moves quickly between reasonably different values of θ), then the chain is said to *mix* well, and relatively small values of m and c will suffice. Conversely, if the chain mixes badly (sometimes known as being *sticky*, as the chain will tend to get stuck moving among very similar values of θ), then very large values of m and c will be required, possibly rendering the method impracticable. One strategy for investigating whether m and c are sufficiently large, and the strategy we adopt here, is to simulate several realizations of the Markov chain, each starting from a different value of $\theta^{(0)}$. If m and c are sufficiently large, then the results obtained should be independent of $\theta^{(0)}$ and should therefore be similar for the different runs. Substantial differences among the results obtained for the different runs indicate that m and c are too small. It is then necessary either to increase m and c or (if this makes the method computationally infeasible) to construct a Markov chain with better mixing properties. In the examples presented in this article we have chosen $c = 1$.

Gibbs sampling is a method of constructing a Markov chain with stationary distribution $\pi(\theta)$, which has proved particularly useful for clustering problems. Suppose that θ may be partitioned into $\theta = (\theta_1, \ldots, \theta_r)$, and that although it is not possible to simulate from $\pi(\theta)$ directly, it *is* possible to simulate a random value of θ_i directly from the full conditional distribution $\pi(\theta_i \mid \theta_1, \theta_2, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_r)$ for $i = 1, 2, \ldots, r$. Then the following algorithm may be used to simulate a Markov chain with stationary distribution $\pi(\theta)$:

Algorithm A1: *Starting with initial values* $\theta^{(0)} = (\theta_1^{(0)}, \ldots, \theta_r^{(0)})$, *iterate the following steps for* $m = 1, 2, \ldots$.

Step 1. *Sample* $\theta_1^{(m)}$ *from* $\pi(\theta_1|\theta_2^{(m-1)}, \theta_3^{(m-1)}, \ldots, \theta_r^{(m-1)})$.
Step 2. *Sample* $\theta_2^{(m)}$ *from* $\pi(\theta_2|\theta_1^{(m)}, \theta_3^{(m-1)}, \ldots, \theta_r^{(m-1)})$.

Step r. Sample $\theta_r^{(m)}$ *from* $\pi(\theta_r|\theta_1^{(m)}, \theta_2^{(m)}, \ldots, \theta_{r-1}^{(m)})$.

It is easy to show that if $\theta^{(m-1)} \sim \pi(\theta)$, then $\theta^{(m)} \sim \pi(\theta)$, and so $\pi(\theta)$ is the stationary distribution of this Markov chain.

Inference on *K*, the number of populations

We now provide further details regarding our approach to choosing K (see *Inference for the number of populations*).

The simplest way of estimating $\Pr(X|K)$ is the so-called harmonic mean estimator

$$\frac{1}{\Pr(X|K)} = \int \frac{\Pr(Z, P, Q|X, K)}{\Pr(X|Z, P, Q, K)} dZdPdQ \approx \frac{1}{M}\sum_{m=1}^{M} \frac{1}{\Pr(X|Z^{(m)}, P^{(m)}, Q^{(m)}, K)}. \quad \text{(A1)}$$

This estimator is notoriously unstable, often having infinite variance, and is thus of little use in practice. One theoretically attractive alternative involves estimating $\Pr(P, Q|X)$ for some P, Q (Chib 1995; Raftery 1996). However, our own implementation of versions of this approach has turned out to be computationally infeasible, due to the very high-dimensional parameter space of our problem. While alternative approaches to estimating $\Pr(X|K)$, such as variable-dimension MCMC methods (Green 1995; Stephens 2000a) or importance sampling (DiCiccio *et al.* 1997), may lead to computationally feasible algorithms, the high-dimensional parameter space makes designing efficient versions of these schemes rather challenging. For this reason we take a more *ad hoc* approach, which begins by considering the *Bayesian deviance*

$$D(Z, P, Q) = -2 \log \Pr(X|Z, P, Q). \quad \text{(A2)}$$

The conditional mean and variance of D given X are easily estimated using

$$E(D(Z, P, Q)|X) \approx \frac{1}{M}\sum_{m=1}^{M} -2 \log \Pr(X|Z^{(m)}, P^{(m)}, Q^{(m)}) = \hat{\mu}, \text{ say}, \quad \text{(A3)}$$

and

$$\mathrm{Var}(D(Z, P, Q)|X) \approx \frac{1}{M}\sum_{m=1}^{M} (-2 \log \Pr(X|Z^{(m)}, P^{(m)}, Q^{(m)}) - \hat{\mu})^2 = \hat{\sigma}^2, \text{ say}. \quad \text{(A4)}$$

If we make the (admittedly dubious) assumption that the conditional distribution of D given X is normal, then it follows from (A1) that

$$-2 \log(\Pr(X|K)) \approx \hat{\mu} + \hat{\sigma}^2/4. \quad \text{(A5)}$$

(Replacing the assumption of normality with the assumption of being gamma-distributed may be more asymptotically justifiable and gives similar results.) We then use this to estimate the posterior distribution of K from (11). An alternative interpretation of this method is that model selection is based on penalizing the mean of the Bayesian deviance by a quarter of its variance (*cf.* Spiegelhalter *et al.* 1999, who suggested investigating model fit using a different penalization of the mean of the Bayesian deviance).

Details of the MCMC algorithms

Algorithm A2: Step 1 may be performed by simulating $p_{kl\cdot}$ independently for each (k, l), from

$$p_{kl\cdot}|X, Z \sim \mathcal{D}(\lambda_1 + n_{kl1}, \ldots, \lambda_{l_l} + n_{kll_l}), \quad \text{(A6)}$$

where

$$n_{klj} = \#\{(i,a) : x_l^{(i,a)} = j \text{ and } z^{(i)} = k\} \quad \text{(A7)}$$

is the number of copies of allele j at locus l observed in individuals assigned (by Z) to population k.

Step 2 may be performed by simulating $z^{(i)}$, independently for each i, from

$$\Pr(z^{(i)} = k|X, P) = \frac{\Pr(x^{(i)}|P, z^{(i)} = k)}{\sum_{k'=1}^{K}\Pr(x^{(i)}|P, z^{(i)} = k')}, \quad \text{(A8)}$$

where $\Pr(x^{(i)}|P, z^{(i)} = k) = \prod_{l=1}^{L} p_{klx(i,1)}p_{klx(i,2)}$.

Note that Equation A8 makes an implicit assumption that an equal fraction of the sample is drawn from each population. Alternatively, it might be natural to introduce an additional parameter for the fraction of the sample drawn from each population.

Algorithm A3: Step 1 may be performed by updating P and Q independently. Updating P is achieved as before, using Equation A6 but where the definition (A7) of n_{klj} is modified in the obvious way to

$$n_{klj} = \#\{(i, a) : x_l^{(i,a)} = j \text{ and } z_l^{(i,a)} = k\}. \quad \text{(A9)}$$

Updating Q involves simulating from

$$q^{(i)}|X, Z \sim \mathcal{D}(\alpha + m_1^{(i)}, \ldots, \alpha + m_K^{(i)}), \quad \text{(A10)}$$

where $m_k^{(i)}$ is the number of allele copies in individual i that originated (according to Z) in population k:

$$m_k^{(i)} = \#\{(l, a) : z_l^{(i,a)} = k\}. \quad \text{(A11)}$$

Step 2 may be performed by simulating $z_l^{(i,a)}$, independently for each i, a, l, from

$$\Pr(z_l^{(i,a)} = k|X, P) = \frac{q_k^{(i)}\Pr(x_l^{(i,a)}|P, z_l^{(i,a)} = k)}{\sum_{k'=1}^{K} q_{k'}^{(i)}\Pr(x_l^{(i,a)}|P, z_l^{(i,a)} = k')}, \quad \text{(A12)}$$

where $\Pr(x_l^{(i,a)}|P, z_l^{(i,a)} = k) = p_{klx_l^{(i,a)}}$.

Step 3 may be performed by simulating a proposal α', from a normal distribution with mean α, and some variance σ_α^2. The proposal is automatically rejected if $\alpha' \leq 0$, and otherwise it is accepted with the appropriate Metropolis-Hastings probability.

Model with correlated allele frequencies

For very closely related populations it is natural to assume that allele frequencies are correlated across populations. For completeness, we describe a model that is implemented in the program *structure*, allowing allele-frequency correlations.

Recall that we model allele frequencies by $p_{kl\cdot} \sim \mathcal{D}(\lambda_1, \lambda_2, \ldots, \lambda_{J_l})$. For all the results presented in this article, we took $\lambda_1 = \lambda_2 = \cdots = \lambda_{Jl} = 1.0$, which gives a uniform distribution on allele frequencies, where J_l is the number of alleles at lows l. To model closely related populations, we consider an alternative model, where

$$p_{kl\cdot} \sim \mathcal{D}(f^{(l)}J_l\mu_1^{(l)}, f^{(l)}J_l\mu_2^{(l)}, \ldots, f^{(l)}J_l\mu_{J_l}^{(l)}). \quad \text{(A13)}$$

Here, $\mu_i^{(l)}$ is the mean sample frequency of allele i at locus l, and $f^{(l)} > 0$ determines the strength of the correlations across populations at locus l. When $f^{(l)}$ is large, the allele frequencies in all populations tend to be similar to the mean allele frequencies in the sample. In our implementation of this model, we placed a gamma prior on each $f^{(l)}$ and used a Metropolis-Hastings update step. The proposal $f^{(l)'}$ was chosen from a normal with mean $f^{(l)}$ and some variance σ_f^2. It was automatically rejected if $f^{(l)'} \leq 0$.

There are several possible alternative models to considering a factor f for each locus. One would be to consider a factor for each population, and another would be to give each type of locus (*e.g.*, SNPs and dinucleotide and trinucleotide repeats) a shared value of f.

THE INTERPRETATION OF POPULATION STRUCTURE BY F-STATISTICS WITH SPECIAL REGARD TO SYSTEMS OF MATING

SEWALL WRIGHT
Department of Genetics, University of Wisconsin, Madison, Wisconsin[1, 2]

Accepted May 15, 1965

Kimura and Crow (1963b) have recently made an interesting comparison between two classes of systems of mating within populations of constant size: ones in which there is maximum avoidance of consanguine mating and ones in which all matings are between close relatives around an unbroken circle. These are illustrated in Figs. 1 and 2 in populations of eight. The rate of decrease of heterozygosis in the former class had, as they note, been found long before to approach $1/(4N)$ asymptotically with increasing size of population, N (Wright, 1921, 1933a). Two cases with patterns of mating similar to those of Kimura and Crow's second class, except that the matings were between neighbors along infinitely extended lines instead of around a circle, had also been considered in these papers. These systems consisted of exclusive mating of half-sibs or of first cousins, otherwise with a minimum of relationship. It was found that there is no equilibrium in either case short of complete fixation locally, in spite of the linear increase in number of different ancestors with increasing number of ancestral generations. This was in contrast to systems (half first cousin or second cousin) in which this increase is more than linear and a steady state is rapidly attained with respect to heterozygosis.

Kimura and Crow were surprised to find that the limiting rates of decrease of heterozygosis in their circular systems are much less than under maximum avoidance approaching $[\pi/(2N + 4)]^2$ in the case of half-sib matings and $[\pi/(N + 12)]^2$ under first-cousin matings with large N. Maximum avoidance delays the onset of the decrease in heterozygosis but after a great many generations the lower rates under the circular systems cause these to fall below the systems of maximum avoidance of corresponding population number, in progress toward fixation.

The authors' surprise was occasioned by the fact that maximum avoidance had been reported in my early papers as approximately halving the rate of decrease in heterozygosis found under random mating, in populations of the same size. A later demonstration (Wright, 1938c, 1939) that the effective size of populations is approximately doubled by using exactly two offspring per parent as parents in the next generation in maintaining population size (assumed in the regular systems of mating) instead of by drawing offspring at random, was unfortunately not applied to this case until its implication was noted by Kimura and Crow. Actually, the limiting rate of decrease of heterozygosis is very slightly *greater* under maximum avoidance than under random mating if exactly two offspring are used from each parent in both cases.

I fully agree with the interpretation of the paradox given by Kimura and Crow. It is, however, I think, instructive to apply to these systems the set of F-coefficients devised for describing population structure in breeds of livestock and in natural populations (Wright, 1943a, 1946, 1951). The following discussion is divided into three parts: first, the justification for using the theoretical correlation coefficients between gametes as the basis for description of population structure; second, a review of previous applications of a set of such correlations; and third, the application to the

[1] Paper No. 988 from the Department of Genetics, University of Wisconsin.

[2] This study was conducted under grant No. GB 1317 from the National Science Foundation.

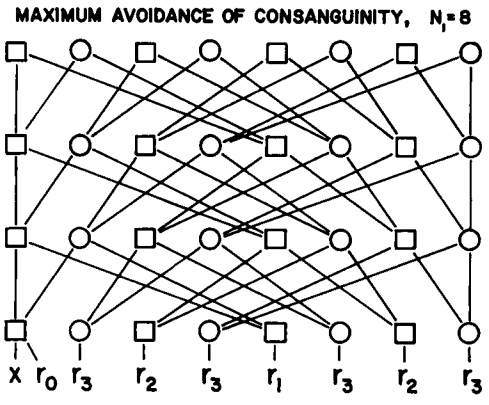


FIG. 1.

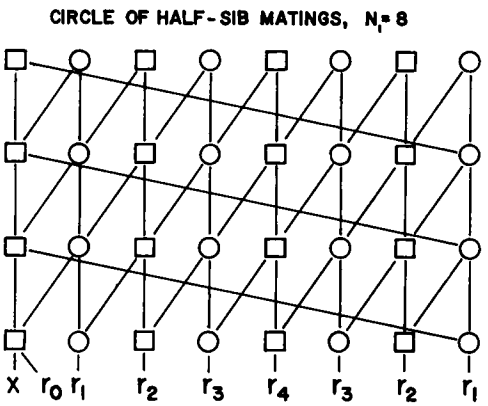


FIG. 2.

systems of maximum avoidance and of circular mating.

PART I. THE FIXATION INDEX F AS THE CORRELATION BETWEEN UNITING GAMETES

The basic coefficient in the set of statistics referred to above is a coefficient F (f in early papers) defined in the 1921 paper as the correlation between homologous genes of uniting gametes under a given mating pattern, relative to the total array of these in random derivatives of the foundation stock. While in most of the cases considered, the basis for the correlation was relationship, formulae were also obtained where the basis was phenotypic similarity in assortative mating. In a paper in the next year (1922a) this coefficient was designated the coefficient of inbreeding in cases in which it is based on relationship of the parents. It should be emphasized that this is a narrower concept than that of the 1921 paper. Later (Wright, 1951), F in the broader context was designated the fixation index. The quantity P ($= 1 - F$), which is often more convenient than F, was designated the panmictic index. It is to be noted that an individual, as representative of a certain mating system, or a population may have more than one fixation index. The most important as a rule, is the inbreeding coefficient pertaining to neutral autosomal disomic loci but there are other inbreeding coefficients pertaining to neutral sex linked (Wright, 1933a, 1951) and polysomic loci (1938b, 1951). There may also be fixation indices relating to loci subject to assortative mating or even to differential selection.

The value of F in terms of that in preceding generations was determined by the form of correlation analysis, designated path analysis. It is desirable to sketch this briefly for comparison with alternative methods.

Path analysis (Wright, 1954, 1963a) is based essentially on the algebraic manipulation of standardized partial regression coefficients (unidirectional path coefficients) in systems of variables, measurable or hypothetical, in which each one that is not treated as an ultimate factor is represented, usually by an arrow diagram, as completely determined by certain others. This requires as a rule an independent residual factor to give the postulated completeness. These determining factors are often represented as determined similarly by more remote ones, and so on, until all lines of determination end in ultimate factors which must be assumed to be correlated with each other. These correlations are indicated in a diagram by bidirectional arrows, if the factors are not known to be independent, in order that the system may be a completely self-contained one.

If all relations are linear, it holds rigorously that the correlation between any two

variables is the sum of contributions from all legitimate paths through the system that tend to contribute to it, the value of the contribution being the product of the path coefficients for the component paths, of which one may be a correlation coefficient (bidirectional) the rest all unidirectional. In the most condensed form, the correlation between two variables, X and Y, may be expressed as $r_{XY} = \Sigma p_{Xi} r_{Yi}$ where p_{Xi} is the unidirectional path coefficient relating X to one its immediate determiners and r_{Yi} is the correlation between this and Y (which reduces to 1 if Y is itself an immediate determiner). This expression may be expanded as far as the diagram permits, by application of this sort of equation to r_{Yi} itself. The self-correlations yield very useful equations of the type, $r_{XX} = \Sigma p_{Xi} r_{Xi} = 1$, in complete systems. In tracing connecting paths through a diagram one must never go forward along an arrow (including an end of a bidirectional one) and then back along another. At least two critics have considered this to be an arbitrary rule but common "descendants" do not tend to contribute to correlation between "ancestors."

A pedigree is a system of the sort described above. Under disomic autosomal heredity there are two sorts of elementary paths. One is for a path from zygote (Z) to a determining gamete (G) and has the value $a\ (= p_{ZG}) = \sqrt{1/[2(1+F)]}$, which follows at once from the equation expressing complete and equal determination by the pertinent genes of the two uniting gametes, $r_{ZZ} = 2a(a + aF) = 1$. The other is for a path from gamete to parental zygote (Z') with the value $b\ (= p_{GZ'}) = \sqrt{\tfrac{1}{2}[1+F']}$, where the prime indicates preceding generation, derivable at once from the fact that the correlation $r_{GZ'}$ which equals b because there is only one connecting path, must equal the correlation $r_{Z'G'} = a'(1+F')$, if there is no intervening selection. Note that there is no implicit assumption in either of these equations with respect to number or frequency of alleles or to values attributed to them. It is, however, assumed that there are no selective differences. In the case of sex-linked heredity (φXX, δXY) the path coefficient relating a male zygote to the determining egg has the value 1 there being no path to the sperm with no X chromosome) and the path coefficient relating an X bearing sperm to the male zygote that produced it is also 1 because of complete determination. In polysomic systems, the analysis must be in terms of pairs of genes that are brought together, not to the sets of alleles of gametes.

In practice, most use is made of two compound path coefficients: a unidirectional one relating a gamete of one generation to one of the two back of it in the preceding generation, which under disomic autosomal heredity always has the value $ba' = \tfrac{1}{2}$, and the bidirectional correlation between two random gametes from the same individual with value $b^2 = \tfrac{1}{2}(1+F')$. Thus the correlation between two gametes (whether uniting ones or not) has the value $\Sigma[(\tfrac{1}{2})^n(1+F_A)]$ where n is the number of individuals along a path that contributes to it, there being $n-1$ unidirectional compound components with the value $\tfrac{1}{2}$ each, and one bidirectional one, connecting two gametes of the common ancestor, A, with the value $\tfrac{1}{2}(1+F_A)$ where F_A is the inbreeding coefficient of this ancestor. This formula is the usual one for use in calculating inbreeding coefficients and correlations between gametes other than uniting ones, where there is only sporadic inbreeding. Closely related is the correlation between two zygotes Z_1 and Z_2 as sums of contributions from the uniting gametes, $Z_1 = G_1 + G_2$, $Z_2 = G_3 + G_4$ (Wright, 1922a).

$$r_{Z_1 Z_2} = \sum \frac{(\tfrac{1}{2})^{n-1}(1+F_A)}{\sqrt{(1+F_{Z_1})(1+F_{Z_2})}}$$

If there is systematic inbreeding, it is more convenient to trace the paths back of the gametes under consideration for only one gamete-to-gamete generation in each case before completing them by correla-

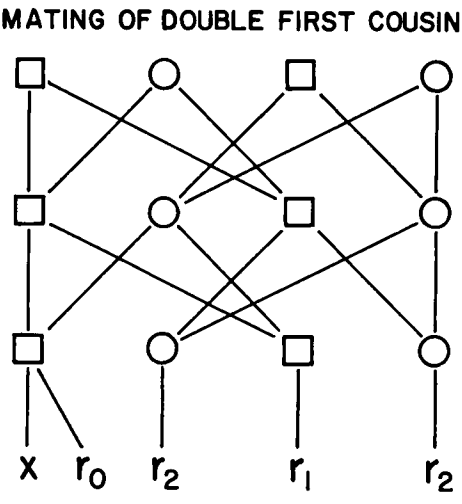


FIG. 3. Gametic correlations with x.

tional paths. Thus if one traces G_1 to G_5 and G_6, and G_2 to G_7 and G_8, we have $r_{12} = \frac{1}{4}[r_{57} + r_{58} + r_{67} + r_{68}]$.

As a simple example, consider the case of mating of double first cousins (Fig. 3), which comes under the heads of both maximum avoidance and of circular first-cousin mating. There are three kinds of correlations among gametes of the same generation: between gametes of the same individual (r_0), between gametes produced by sibs (r_1), and between ones produced by first cousins (r_2). The last is also the inbreeding coefficient, F. These may be expressed in terms of r's of the preceding generation by inspection (Fig. 3).

$$r_0 = \tfrac{1}{2}(1 + r_2')$$
$$r_1 = \tfrac{1}{4}(2r_0' + 2r_2')$$
$$r_2 = \tfrac{1}{4}(2r_1' + 2r_2')$$

The desired expression for F is obtained by substitution in the third. Using numbers of primes to indicate the numbers of preceding generation $F = \frac{1}{2}F' + \frac{1}{4}F'' + \frac{1}{8}F''' + \frac{1}{8}$. There is usually simplification by substitution of $1 - P$ for F.

$$P = \tfrac{1}{2}P' + \tfrac{1}{4}P'' + \tfrac{1}{8}P'''$$

On putting $\lambda = P/P' = P'/P'' = P''/P'''$ to obtain the limiting ratio of P to P', expected here because of the constant size of population.

$$8\lambda^3 - 4\lambda^2 - 2\lambda - 1 = 0, \qquad \lambda = 0.91964$$

The most important property of P, assuming that the pedigrees trace to a random breeding foundation stock, is that it gives the amount of heterozygosis (y) relative to that under random mating, $y_0 = 2q(1-q)$, where q is gene frequency. By constructing a 2×2 correlation array of uniting alleles in the two allele case, it comes out at once that $F = 1 - (y/y_0)$ and thus that $P = y/y_0$. From this the zygotic composition of the population with respect to the pair of alleles $[qA + (1-q)a]$ can be written at once in terms of deviations from complete fixation, $[q - Pq(1-q)]AA + [2Pq(1-q)]Aa + [(1-q) - Pq(1-q)]aa$ (Wright, 1922b) or in two algebraically equivalent ways: as deviations from panmixia, $[q^2 + Fq(1-q)]AA + [2q(1-q) - 2Fq(1-q)]\ Aa + [(1-q)^2 + Fq(1-q)]\ aa$ or as the weighted average of panmictic and fixed components, $[Pq^2 + Fq]AA + [2Pq(1-q)]Aa + [P(1-q)^2 + F(1-q)]\ aa$. A coefficient α proposed by Bernstein (1930) in this connection is the same as F.

These formulae can be extended to any number of alleles by the principle that any group of alleles *e.g.* $(q_iA_i + q_jA_j)$, can be treated formally as if one, $(q_i + q_j)A_{ij}$ (Wright, 1931, 1951). Thus the frequency of any homozygote A_iA_i (opposed to all of its alleles collectively) is of the type $[Pq_i^2 + Fq_i]$ in which F is obtained by path analysis, and that of any collective pair, A_{ij}, as opposed to all of the rest collectively, is of the type $[P(q_i + q_j)^2 + F(q_i + q_j)]$ by the same analysis, whence by subtraction, that of any heterozygote, A_iA_j, is of the type $2Pq_iq_j$. The ratio of the total amount of heterozygosis, $2\Sigma Pq_iq_j$, $(i < j)$, to that in the randombred foundation stock, $2\Sigma q_iq_j$, is thus $P(= 1 - F)$ irrespective of number, frequencies, or values of alleles. F, as calculated by path analysis, is thus a rather unusual sort of correlation coefficient, but the principle may easily be verified by actually setting up

the correlation array with arbitrary frequencies and values for each allele in a multiple system and calculating the correlation coefficient by the usual formula (Wright, 1963a).

If the average inbreeding coefficient is calculated from pedigrees that trace to a foundation stock that is itself inbred, it gives the correlation between gametes with histories of the sorts indicated, relative to the gene frequencies of this foundation stock, not to those of any more remote foundation stock to which the latter traces.

Thus the inbreeding coefficient (or more generally the fixation index), F, and the panmictic index, P, must always be specified as relative to the particular foundation stock to have any meaning.

It is not necessary here to discuss the alternative method of calculating the decline in heterozygosis under systems of mating, based on types of matings instead of gametic unions. This method, as developed by Bartlett and Haldane (1934, 1935) gives a more complete account of the consequences of systems of mating but is much more cumbersome than path analysis in arriving at the most significant parameter and does not permit generalization from small to large populations as readily. Thus the analysis of double first-cousin matings which requires a 12×12 matrix and thus solution of a characteristic equation of the 12th degree by the method of Bartlett and Haldane (Fisher, 1949) required only solution of a cubic equation by path analysis (as indicated above) and could easily be generalized to any population of size 2^m with maximum avoidance of consanguine mating (Wright, 1921, 1933a). The equations of path analysis are also expressible in matrix form but this has been needed only in exceptionally complicated cases (Wright, 1963a).

Another approach to such problems was that of Malécot (1948). He introduced the interpretation of the inbreeding coefficient F as the probability of identity of representatives of a locus, due to descent from a common origin. Since the probability that an autosomal gene is identical with one of the two a generation back of it is ½ (in the absence of selection) and thus the same as the compound path coefficient ba', and since the probability that two random gametes of the same individual have the same gene by origin is $\frac{1}{2}(1 + F')$, and thus equal to the compound bidirectional coefficient b^2, and since probabilities, like path coefficients, compound by multiplication along paths, the resulting formula $F = \Sigma[(\frac{1}{2})^n(1 + F_A)]$ is identical with the basic formula of path analysis in this case. There is similar agreement in the cases of sex linkage and polysomy. This mathematical identity of the methods also applies if probabilities are traced only to the preceding gamete generation before combining with the probabilities of identity by origin of the two gametes to which these lead.

This method is therefore identical as far as calculation is concerned with path analysis as applied to gametes. The only difference is in interpretation. The same relation can be interpreted either as the correlation between two gametes, relative to the foundation stock, or as the probability of identity by origin from this foundation stock.

The concepts are thus very closely related. The principle of equal probabilities at segregation is indeed directly involved in assigning the value of the elementary path coefficient b. The principles of probability were also explicitly involved in the application of path analysis to assortative mating (Wright, 1921) and to all applications to random breeding populations (*e.g.*, ones with N_m males and N_f females, Wright, 1931).

An especially important example of the use of principles of probability in calculating correlations has been in determining the phenotypic correlation between relatives with respect to loci involving dominance (Malécot, 1948) and factor interaction (Kempthorne, 1954; Cockerham, 1954). Path analysis was developed as a method of dealing rigorously with correla-

tions in systems of multiple variables connected exclusively by linear relations and as such did not seem competent to deal with the non-linear relations usually involved in phenotypic correlations (except in cases in which there could be no correlation between the dominance or interaction deviations of the relatives in question).

It is possible, however, to extend path analysis to include contributions to correlation from parallel all-or-none joint contributions of two or more variables to the variances of the two non-inbred relatives. The value of the compound joint path coefficient is the product of all of the elementary coefficients in all of the parallel paths and is thus zero if any one of these is zero. With this addition, the formulae for the effects of dominance and the various types of factor interaction agree with those of the above authors (Wright, 1963b).

Returning to relations between gametes, the significance of the inbreeding coefficient and of the similar parameter for pairs of gametes in general, is undoubtedly much enhanced by the fact that they can always be interpreted either as correlation coefficients or as probabilities of identity of origin, in both cases relative to a specified foundation stock.

There is, however, an important difference between those interpretations in applicability in the broader context of a system of parameters useful in concise discrimination among population structures. Correlation coefficients vary between -1 and $+1$ while probabilities can only vary between 0 and $+1$. A correlation between gametes as calculated from a pedigree cannot be negative (as may be seen from the general formula $F = \Sigma[(\frac{1}{2})^n(1 + F_A)]$ and thus can always be identified with a probability. It may, however, be useful to find the correlation between such gametes relative to the array of gametes of their own generation. Such a correlation (not capable of direct calculation from the pedigree) may be negative. Thus it is negative by definition, if there is maximum avoidance of consanguinity, which is one of the classes of mating systems with which this paper is especially concerned.

There may also be a possibility of negative correlation between uniting gametes (and thus more heterozygosis than under random mating) in the broader context of F as a fixation index in cases in which there are other reasons for correlation than consanguinity. The effects of assortative mating, based on a phenotypic correlation, r, between mates, or genotypic correlation $m = h^2r$ where h^2 is the heritability of the character, affected by a system of multiple (n) pairs of alleles with equal frequencies, no dominance and equivalent effects, with respect to which the assortative mating occurs, were dealt with by path analysis at the same time as the effects of systems of inbreeding (Wright, 1921). The phenotypic correlation was taken as constant. Less restrictive postulates may be made but this case suffices for the present purpose.

The genotypic correlation, m, may be expected in general to reach a steady state, $\hat{m}$, at less than 1. The correlation between uniting gametes (with respect to the array of loci involved) was represented by f; that between uniting genes of a single locus by f_u. The latter, however, may be considered a case of the fixation index since the amount of heterozygosis at each locus relative to that is the random breeding foundation stock is given by $1 - \hat{f}_u$.

$$F = \hat{f}_u = \hat{m}/(2n - 2n\hat{m} + \hat{m})$$

$$\hat{f} = \hat{m}/(2 - \hat{m})$$

These are positive if m is positive but negative if m is negative. In the extreme case of perfect disassortative mating ($r = -1$) and complete heritability ($h^2 = 1$, $m = -1$), $F = -1/(4n - 1)$ and $\hat{f} = -\frac{1}{3}$. The amount of heterozygosis relative to that in the foundation stock ($\frac{1}{2}$ under the assumptions) is given by $P = 1 - F = 4n/(4n - 1)$ and thus is greater than 1 in this case. In absolute terms, $y = 2n/(4n - 1)$ under the assumption that gene frequency is $\frac{1}{2}$.

The case of a locus in which there is a steady state because of selection against both homozygotes in favor of the heterozygotes similarly gives a negative fixation index.

	Frequency	w	
AA	q^2	$1-s$	$\bar{w}=1-sq^2-t(1-q)^2$
Aa	$2q(1-q)$	1	$\Delta q=-(s+t)q(1-q)\times[q-\hat{q}]$, $\hat{q}=t/(s+t)$
aa	$(1-q)^2$	$1-t$	$\hat{w}=1-[st/(s+t)]$

Calculation of the correlation between uniting gametes gives the negative fixation index $F=-st/[s+t-st]$, and $P=(s+t)/(s+t-st)$ is in excess of 1.

Part II. The F-Statistics

In studying the history of the British Shorthorn cattle from the herdbook records (Wright, 1923a, 1923b; McPhee and Wright, 1925, 1926) it became obvious that mere specification of the average inbreeding coefficient at successive periods was not adequate. It was found desirable to supplement this in ways which will be discussed briefly later.

A system was developed from this start (Wright, 1943a, 1946, 1951) for describing the properties of hierarchically subdivided natural populations. Three parameters were proposed in the 1951 paper in terms of a total population (T), subdivisions (S), and individuals (I). F_{IT} is the correlation between gametes that unite to produce the individuals, relative to the gametes of the total population. F_{IS} is the average over all subdivisions of the correlation between uniting gametes relative to those of their own subdivision. F_{ST} is the correlation between random gametes within subdivisions, relative to gametes of the total population. The list can be extended if there are further subdivisions.

The above three F-statistics are not independent. One of two demonstrations of their interrelation, given in the 1943 paper, is repeated below in the later symbolism. The effects of accidents of sampling are here ignored.

The amount of heterozygosis (y_T) in the total population, whatever its structure, is as follows in terms of total gene frequency q_T, as brought out earlier.

$$y_T=2q_T(1-q_T)(1-F_{IT})$$

Assume first that the total is divided into many (n) random breeding demes (D) with varying gene frequencies, q_D, and amounts of heterozygosis, $y_D=2q_D(1-q_D)$. The total amount of heterozygosis is the average of these.

$$y_T=\frac{2}{n}\sum^{n}q_D(1-q_D)=2\left[q_T-\frac{1}{n}\sum^{n}q_D^2\right]$$

The value of the term $\frac{1}{n}\sum^{n}q_D^2$ may be obtained from the variance of gene frequencies of demes within the total, $\sigma_{q(DT)}^2$:

$$\sigma_{q(DT)}^2=\frac{1}{n}\sum^{n}(q_D-q_T)^2=\frac{1}{n}\sum^{n}q_D^2-q_T^2.$$

Thus $y_T=2[q_T(1-q_T)-\sigma_{q(DT)}^2]$.

This formula was enunciated first by Wahlund (1928). On equating the two expressions for y_T,

$$\sigma_{q(DT)}^2=q_T(1-q_T)F_{IT}.$$

It should be noted that the correlation between random gametes drawn from demes (F_{DT}) is the same as F_{IT} if there is random mating within the deme.

Consider next, division of the total into subdivisions that are themselves inbred. As before, $y_T=2q_T(1-q_T)(1-F_{IT})$ but as an average $y_T=\frac{2}{n}\sum^{n}[q_S(1-q_S)](1-F_{IS})$ $=2(1-\bar{F}_{IS})\left(q_T-\frac{1}{n}\sum^{n}q_S^2\right)$ assuming that F_{IS} and q_S are independent. Again,

$$\sigma_{q(ST)}^2=\frac{1}{n}\sum^{n}(q_S-q_T)^2=\frac{1}{n}\sum^{n}q_S^2-q_T^2,$$

$$y_T=2[(1-\bar{F}_{IS})q_T(1-q_T)-\sigma_{q(ST)}^2].$$

Thus $\sigma_{q(ST)}^2=q_T(1-q_T)[F_{IT}-F_{IS}]/[1-F_{IS}]$, dropping the bar over F_{IS}.

If now, completely random mating were instituted in the subdivisions there would be no change in their gene frequen-

cies and hence none in $\sigma_{q(ST)}^2$, but now $\sigma_{q(ST)}^2 = q_T(1-q_T)F_{ST}$ where F_{ST} is defined as the correlation between random gametes of the subdivision.

Thus $F_{ST} = (F_{IT} - F_{IS})/(1 - F_{IS})$.

This is simplified if expressed in terms of P's.

$$P_{IT} = P_{IS}P_{ST}$$

If there are secondary subdivisions into local races (R) which may themselves be inbred ($F_{IR} \neq 0$), $P_{IS} = P_{IR}P_{RS}$ and $P_{IT} = P_{IR}P_{RS}P_{ST}$.

Such analysis may be continued as far as there is hierarchic subdivision.

These equations add another interpretation of the F-statistics to the three already treated (as correlations, as functions of the relative amount of heterozygosis, and, in some cases, as probabilities of identity by origin). F_{IT} gives the ratio of the variance $\sigma_{q(DT)}^2$ of gene frequencies of random breeding subdivisions (D) (if these occur) to its maximum possible value $q_T(1-q_T)$, expected if the subdivisions are completely isolated and each completely fixed, thus forming the array $q_T AA + (1 - q_T)aa$. $F_{IS} = \frac{1}{m}\sum^{m}\sigma_{q(DS)}^2/q_S(1-q_S)$ gives the average of such ratios among subdivisions. Most importantly, however, F_{ST} is the ratio of the actual variance of gene frequencies of subdivisions to its limiting value, irrespective of their own structures. F_{ST} is thus necessarily positive. F_{IS}, while usually positive, is negative if there is systematic avoidance of consanguine mating within the subdivisions. F_{IT} is positive if there is systematic subdivision, whether into demes ($F_{IS} = 0$, $F_{IT} = F_{ST}$) or into inbred groups, but can be negative if there is no systematic subdivision and there is prevailing avoidance of consanguine mating.

If pedigrees can be traced some distance back, a pedigree F can be obtained which differs from F_{IT} above in relating to a still more comprehensive total than currently exists, the total of all hypothetically similar populations, derivable from the foundation stock of the earlier period. This, as noted, is necessarily positive and interpretable as the probability of identity of origin of the uniting gametes in the current total populations.

In addition to the above properties, the F-statistics have important relations to the statistics of quantitatively varying characters. If the effects of genes are completely additive (semidominance and no factor interactions) the mean of the character in the total population or its subdivision depends only on the gene frequencies, unaffected by the F-statistics. The variances are, however, dependent on the latter. The variance of individuals in the total population, σ_{IT}^2, is analyzed below into the variance of subdivision means, σ_{ST}^2, and the mean of the variances within subdivisions, σ_{IS}^2. The general formula does not seem to have been given before. That for randombred subdivisions has long been known (Wright, 1921, 1943b).

The variance that would be found with the same gene frequencies but random mating throughout is represented by $\sigma_{IT(0)}^2$.

	General	Randombred subdivisions ($F_{IS} = 0$, $F_{ST} = F_{IT}$)
σ_{ST}^2	$2F_{ST}\sigma_{IT(0)}^2$	$2F_{IT}\sigma_{IT(0)}^2$
σ_{IS}^2	$(1 + F_{IT} - 2F_{ST})\sigma_{IT(0)}^2$	$(1 - F_{IT})\sigma_{IT(0)}^2$
σ_{IT}^2	$(1 + F_{IT})\sigma_{IT(0)}^2$	$(1 + F_{IT})\sigma_{IT(0)}^2$

If there is other than semidominance, the mean is affected in a way that is responsible for the well-known effects of inbreeding. At a given gene frequency, the mean with a given value of F_{IT}, represented by $M(F_{IT})$, is related to the means under random mating, $M(0)$, and under complete fixation of subdivisions, $M(1)$, by the formula (Wright, 1922b).

$$M(F_{IT}) = M(0) + F_{IT}[M(1) - M(0)]$$

This applies to subgroups on replacing F_{IT} by F_{IS}.

The total variance under inbreeding, $\sigma_{IT(F)}^2$, is related to that under random mating throughout by the formula (Wright, 1951).

$$\sigma_{IT(F)}^2 = (1 - F_{IT})\sigma_{IT(0)}^2 + F\sigma_{IT(1)}^2 + F_{IT}(1 - F_{IT})[M(1) - M(0)]^2$$

The analysis of this into the components, σ_{ST}^2 and σ_{IS}^2, requires third and fourth moments of the distribution of gene frequencies among subdivisions. The somewhat surprising results for dominant characters under progressive inbreeding of completely isolated subdivisions have been presented by Alan Robertson (1952). The somewhat different ones where there is a steady state because of a balance between local inbreeding and immigration have also been presented (Wright, 1952).

The theory of the distribution of gene frequencies under joint action of systematic and random processes deals with aspects of population structure that are not in general amenable to path analysis because of the non-linear action of selection. A connection can be established, however, in cases in which there is linearity. The general formula for the steady state distribution at a single locus is as follows letting Δq represent the change per generation that systematic factors tend to produce and $\sigma_{\Delta q}^2$, the contribution to variance from random ones (Wright, 1938a).

$$\phi(q) = (C/\sigma_{\Delta q}^2) \exp\left[2\int(\Delta q/\sigma_{\Delta q}^2)\, dq\right]$$

Immigration is responsible for a linear pressure on gene frequency $\Delta q_S = -m(q_S - q_T)$ in which m is the amount of replacement of local genes (frequency q_S) by immigrant ones (frequency q_T). The sampling variance is $\sigma_{\Delta q}^2 = q_S(1-q_S)/2N$. Substitution yields:

$$\phi(q) = \frac{\Gamma(4Nm)}{\Gamma(4Nmq_T)\Gamma[4Nm(1-q_T)]} \times q_S^{4Nmq_T-1}(1-q_S)^{4Nm(1-q_T)-1}.$$

This formula was obtained in a different way earlier (Wright, 1931).

$$\bar{q}_S = \int_0^1 q_S \phi(q_S)\, dq_S = q_T$$

$$\sigma_{q(ST)}^2 = \int_0^1 (q_S - q_T)^2 \phi(q_S)\, dq_S = q_T(1-q_T)/[4Nm+1]$$

Thus $F_{ST} = \dfrac{1}{4Nm+1}$ in this case (Wright, 1931).

By path analysis $F_{ST} = (1-m)^2\{1/(2N) + [1-1/(2N)]F_{ST'}\}$.

At equilibrium $F_{ST'} = F_{ST}$, giving (Wright, 1943a, 1951) $F_{ST} = (1-m)^2/[2N-(2N-1)(1-m)^2] \approx 1/(4Nm+1)$.

The two approximate determinations of F_{ST} thus agree. More generally F_{ST} in the broad sense can always be obtained, at least empirically, for the variance of distribution of gene frequencies even in cases involving selection, from the formula $F_{ST} = \sigma_{q(ST)}^2/q_T(1-q_T)$. The results, of course, apply only to the particular loci in question.

A more direct determination of F_{ST} is possible in the important case in which there is a balance between local selection pressure and immigration, that varies among the subdivisions. In the simplest case (additive gene effects, s always less than m in absolute value).

$$\Delta q_S = sq_S(1-q_S) - m(q_S - q_T)$$

$$\hat{q}_S = q_T + \frac{s}{m} q_T(1-q_T) \quad \text{approximately}$$

(Wright, 1931)

$$\sigma_{\hat{q}_S}^2 = \sigma_{q(ST)}^2 = \sigma_{(s/m)}^2 q_T^2(1-q_T)^2$$

$$F_{ST} = \sigma_{(s/m)}^2 q_T(1-q_T)$$

The situation, if s varies from positive values greater than m to negative ones also greater in absolute value, is more complicated but it would be possible to calculate F_{ST} for the locus in question, again as a fixation index, not an inbreeding coefficient. The point here is that many different aspects of population structure can be brought under a common viewpoint by means of the F-statistics.

There was nothing in the derivation of the three basic F-statistics that implies anything about the degree of isolation of the subdivisions or their arrangement in space. They may be completely isolated, at one extreme, or merely arbitrarily bounded portions of a continuum, at the other. Their gene frequencies may be distributed at random in the total population or in an orderly cline. The full account of an actual population obviously requires a map, and detailed accounts of the various regions within it and their relations. There are, however, additional general aspects of

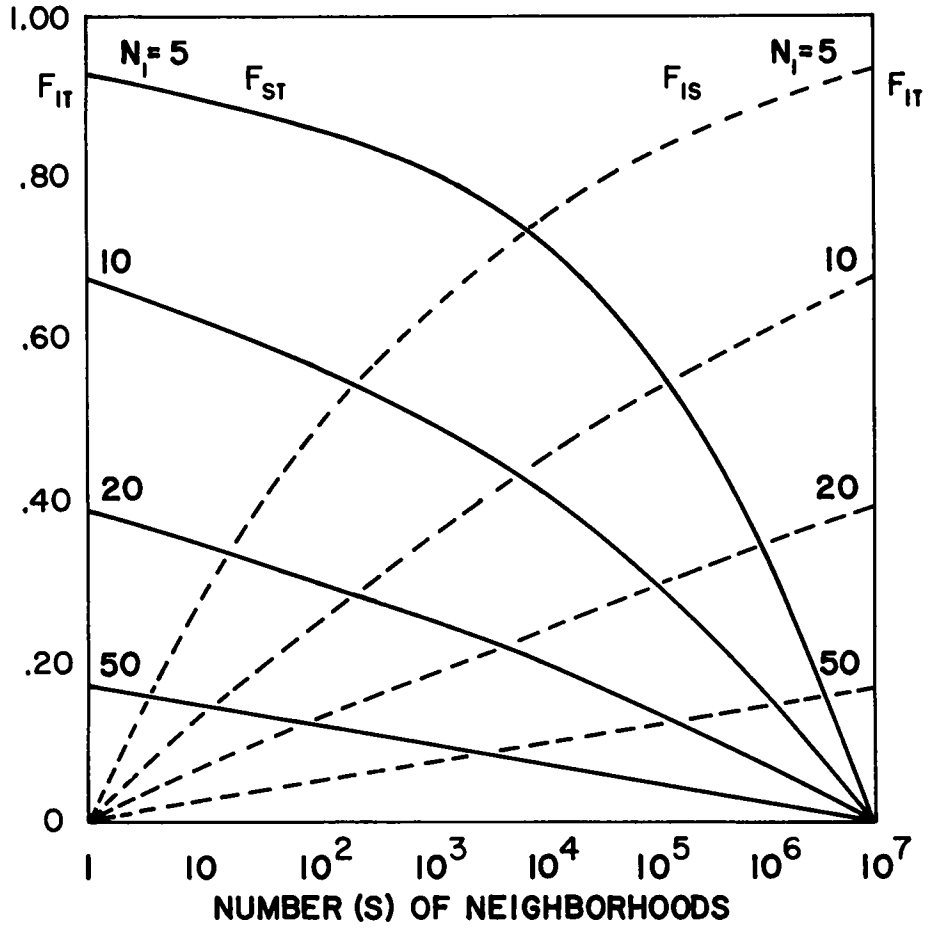


FIG. 4.

the structure than can be brought out by extension of the hierarchic pattern. This can be illustrated by considering the application of F-statistics to the properties of a continuum in which there is some degree of local isolation because of restricted dispersion (Wright, 1940, 1943a, 1946, 1951).

The most important unit is the "neighborhood," the size of population (N_1) from which the parents may be treated as if drawn at random. The properties of the system depend largely on N_1. If the distribution of birthplaces of parents are distributed normally relative to those of offspring in the case of a linear continuum, such as a shoreline, effective N_1 is equivalent to the number of reproducing individuals along a strip $2\sqrt{\pi\sigma}$ long. If the distribution is normal in all directions in an area continuum, effective N_1 is equivalent to the number in a circle of radius 2σ (Wright, 1946).

Arbitrary subdivisions (S) may be of any size from that of the neighborhood up

to the total (T) under consideration. Various systems of mating have been considered (Wright, 1946). It will suffice here to review only the simplest case, that of random union of gametes in a monoecious population since other cases do not differ very much. It will be convenient to use x for a varying number of ancestral generations.

In the case of an area continuum, ancestors of generation x are drawn from an effective population of size xN_1. It was shown that for a subdivision (S generations):

$$F_{IS} = \sum_1^{S-1} t_x \Big/ \left[2 - \sum_1^{S-1} t_x \right], \quad [t_1 = 1/N_1]$$

$$t_x = \frac{(x-1) - (1/N_1)}{x} t_{x-1}.$$

Σt_x can be calculated as the sum of such terms, or approximated by an integration formula or, as pointed out by D. J. Hooton, by the formula $\sum_1^{S-1} t_x = 1 - St_S$.

F_{IT} is analogous for the total (T generations).

$$F_{ST} = (F_{IT} - F_{IS})/(1 - F_{IS})$$

In the case of a linear continuum, ancestors of generation x are drawn from an effective population of size $\sqrt{x}\, N_1$. If F_{IS} refers to a strip containing $\sqrt{S}$ neighborhoods:

$$F_{IS} = \sum_1^{S-1} t_x / [2 - \sum_1^{S-1} t_x], \qquad [t_1 = 1/N_1]$$

$$t_x = \frac{[\sqrt{x-1} - (1/N_1)]}{\sqrt{x}} t_{x-1}$$

$$\sum_1^{S-1} t_x = 1 - \sqrt{S}\, t_S.$$

The population structure may best be indicated (as in Fig. 4) by plotting F_{IS} ($= \sigma_{q(IS)}^2 / q_S(1-q_S)$) and F_{ST} ($= \sigma_{q(ST)}^2 / q_T(1-q_T)$) against log S over the range 0 to log T. F_{IT} is the last value of the former and first of the latter. The interpretation of F_{IS} as a measure of the variability among neighborhoods within areas of increasing size suffers from the disadvantage that its denominator increases with increasing S, though less rapidly than the numerator. F_{ST}, however, is without this drawback and shows how differentiation of neighborhoods builds up decreasing amounts of differentiation in areas of larger size.

The comparison of the curves for different sizes of neighborhood shows, however, that N_1 must be very small (in the case of an area continuum) for an appreciable differentiation even of neighborhoods. This is not the case with a linear continuum.

It should be added that reversible mutation (u, v) or a small amount of universal dispersion (m) imposes a limit beyond which F_{IS} cannot increase with increasing S. This can be estimated by an integration formula but more easily for an area continuum by a formula suggested by Alan Robertson $\Sigma t = 1 - [1 - (1-m)^2]^{1/N_1}$. Fig. 5 shows how F_{ST}, ($N_1 = 20$) is affected by increasing m. The rates of mutation to (v) and from (u) the gene may be incorporated into m. Balanced universal selection has somewhat similar effects.

If the distributions of parental distances are not normal, the forms of these curves are somewhat modified. The distributions of ancestral populations approach normality. With leptokurtic parental distributions, which seem to be most usual (Bateman, 1950), there is little modification in the case of areas but considerably more damping in the case of linear ranges.

This type of theory has been extended to the important model of population structure in which there are clusters of high density distributed uniformly over the range of the species. The size of population of such clusters may be large but the rise in F_{IS} depends on the small increments to the size of the ancestral populations due to small amounts of dispersion per generation in the population as a whole. If this increment is treated as the significant "N_1," the rise of F_{IS} from zero for the cluster is substantially the same irrespective of the cluster size (Wright, 1951).

Malécot (1948) has attacked the problem of isolation by distance in a very dif-

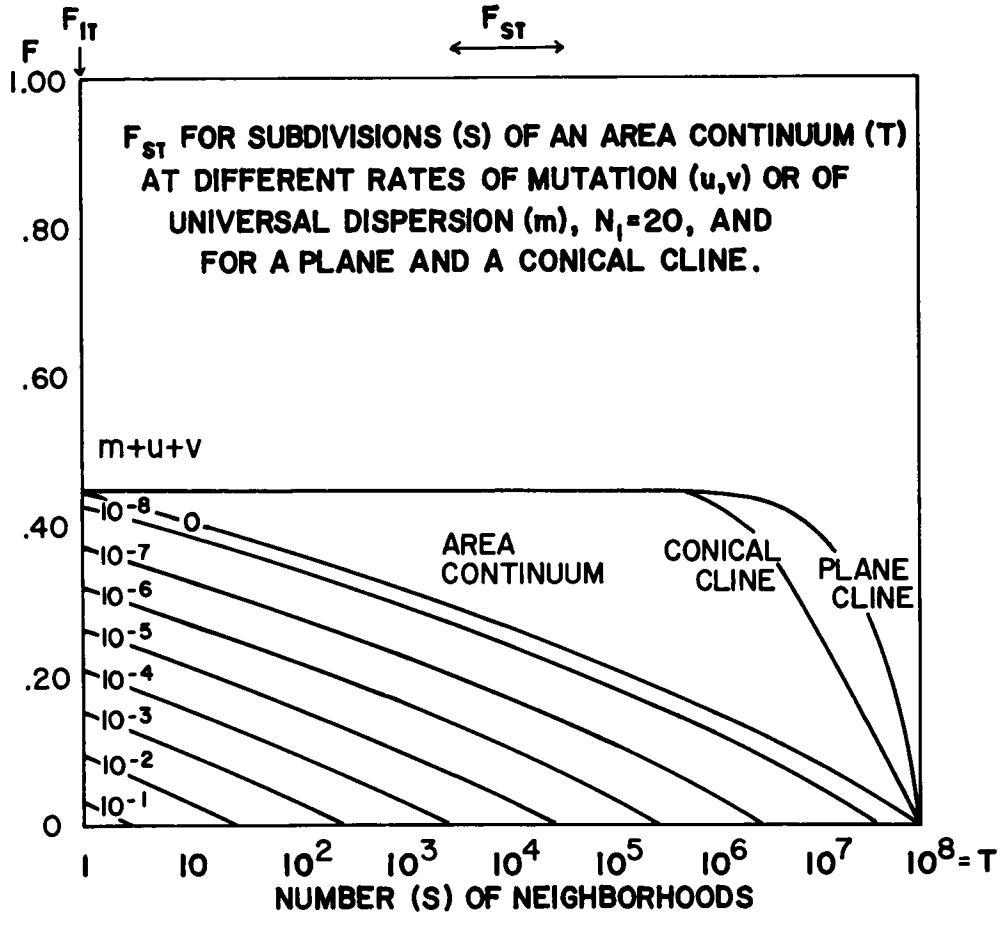


FIG. 5.

ferent way. His primary objective was the correlation between regions as a function of their distance apart. Kimura (1953) studied what he called the stepping stone model which was similar to a cluster model in one dimension. Kimura and Weiss (1964) have extended this to two and three dimensions. They also dealt primarily with the correlation between regions as functions of distance but also considered variances. Points of contact have been established between these models and those of the *F*-statistics but no full comparison has yet been made because of differences in mathematical form.

If, as assumed above, local differentiation is due only to the balancing of local inbreeding and dispersion, F_{ST} for an area continuum falls off much more rapidly with increasing S than it does in the case of a cline, based on orderly regional changes in the conditions of selection. Consider the case of a uniform gradient in one direction over a square total area (a plane cline). If this area is divided into n squares, the variance is only $1/n$ less than for very small subdivisions, corresponding to neighborhoods in the preceding case. The curve for F_{ST} is virtually level up to large fractions of the total (Fig. 5).

In the case of a conical cline, with uniform change of average in all directions from a high (or low) center, there is more rapid falling off of amount of differentia-

tion with increasing size of the areas under consideration than in the case of a plane cline, but still much less than in the chaotic pattern of differentiation due to balancing between inbreeding and dispersion. With seven equal areas (a central one surrounded by six others), the variance of area averages is 51.6 per cent of that of very small subdivisions. With 19 such areas in two concentric zones about a central one, the corresponding figure is 78.9 per cent. With 37 such areas in three concentric zones about a central one, it is 88.7 per cent and in the case of 61, it is 93 per cent. The curve for F_{ST} is still virtually level over most of the range of values of S in a large total population (Fig. 5).

We return here to the very different sort of application of the theory that has been made for the study of breed history. In this case, pedigree F, treated as F_{IT} measures inbreeding relative to the foundation stock and thus to a hypothetical contemporary total (T). The contemporary breed of any period under study is treated as a subdivision, S, of this total. F_{ST} is the correlation between random gametes from the parental generation relative to the same total. This can be calculated from random pedigrees by random association of sires and dams, while F_{IS}, the correlation between mating gametes within the contemporary array of gametes cannot be calculated directly from pedigrees but must be calculated from the formula $F_{IS} = (F_{IT} - F_{ST})/(1 - F_{ST})$. It is negative if F_{ST} is greater than F_{IT}.

The method actually used in studying the history of the British Shorthorns (McPhee and Wright, 1925, 1926) was to select a considerable number of bulls and cows at random from Coates' herdbook of a given year, and construct random samples of the pedigrees of the sire and dam of each (Wright and McPhee, 1925). At first, single random lines, the sequence of sires and dams being determined by coin tossing, were carried back of all four grandparents but later it was recognized that it was preferable to trace only two such lines, one back of each parent. The average value of F_{IT} for the breed at the time is given by $(k/2m)(1 + \bar{F}_A)$ if there are k cases of a common entry among m such pairs of pedigree lines. This approaches 1 if nearly every pair shows a common entry and the average inbreeding coefficient of this animal itself (A) approaches 1. In this form a standard error is easily calculated. The calculation of F_{ST} is exactly the same except that the pedigree samples of sires and dams of different animals are matched at random for common entries.

In the original study, the calculation of F_{IT} was supplemented by calculations of the correlation between random animals of each period and two bulls, Charles Colling's Favourite born in 1793, about which the breed was formed, and Amos Cuickshank's Champion of England, born 1859 about which it was largely reformed nearly a century later. These correlations were compared with the correlation between random animals of the same period. The latter is closely related to F_{ST} which as "the amount of inbreeding due to random mating" was calculated but not wholly correctly. Table 1 and Fig. 6 show F_{IT} separately for males and females as well as the average revised figures for F_{ST}, F_{IS}, and the other correlations referred to.

This table and figure show that a very substantial amount of inbreeding was built up in the course of a century; $F_{ST} = 0.249$ in 1920. This was possible because of relatively small effective numbers, especially in early years.

It was noted (Wright, 1931) that the effective number, N_e, may differ from the number of mature individuals for at least three reasons: an extreme sex ratio, differences in productivity beyond expectation under random sampling, and passage through a bottleneck of small size. The effect of sex ratio was indicated by the formula $N_e = 4N_mN_f/(N_m + N_f)$ which approaches $4N_m$ if there are many more females than males. The effect of differential productivity is given in a monoecious population with random union of gametes

TABLE 1. *F-statistics and coefficients of relationship for the British Shorthorn cattle at successive periods.*

	F_{IT}			F_{ST}	F_{IS}	Relationship		
	♂	♀	avg.			Favourite	C. of E.	Interse
1780	0	0	0	–	–	–	–	–
1810	19.1 ± 1.5	14.3 ± 1.3	16.6	12.8	+4.4	44.3	26.3	22.0
1825	21.9 ± 2.1	16.1 ± 3.7	19.9	16.1	+4.5	51.3	29.9	26.8
1850	18.4 ± 1.9	16.9 ± 3.7	18.0	20.0	−2.5	50.1	26.1	33.9
1875	25.9 ± 2.1	29.0 ± 3.8	27.4	24.1	+4.3	57.6	32.9	37.8
1900	23.2 ± 2.1	22.6 ± 3.8	22.9	24.1	−1.6	52.1	39.2	39.3
1920	25.4 ± 4.0	26.7 ± 4.1	26.0	24.9	+1.5	55.2	45.5	39.5
Dairy	–	27.5 ± 2.1	–	25.3	+2.9	56.1	42.1	39.7

by the formula $N_e = (4N-2)/(2+\sigma_k^2)$ where σ_k^2 is the variance of number of offspring that reach maturity about the mean $\bar{k}=2$ in a population of constant size (Wright, 1939). Kimura and Crow (1963a) arrived at $(4N-4)/(2+\sigma_k^2)$ as the effective number with $N/2$ males and $N/2$ females to which, as in all cases with separate sexes 0.5 must be added for comparison with the case of random union of gametes. Finally the effective number is the harmonic mean of the varying numbers over a number of generations and thus is dominated by the numbers at bottlenecks of population size (Wright, 1938c, 1939).

All of these played roles in the Shorthorn breed. The total numbers were relatively small in the foundation period which was thus something of a bottleneck. The number of sires was much less than the number of dams and the effect of this was exaggerated by wide differences in the productivity of the sires that were used, especially with respect to sons that became sires. The breed was essentially founded by intensive inbreeding to the bull Favourite to which the breed as a whole came to have a relationship somewhat closer than between parent and offspring. This was made possible to a large extent by the building up of strains such as Thomas Bates' Duchesses with average inbreeding coefficient of 0.409 over eight generations (64 cows) (Fig. 6) and average relationship to Favourite of 0.656 (Wright, 1923b).

It may be seen that in the early years, the bulls were of more inbred origin than the cows. In the later years, increasing relationship was built up to the bull Champion of England until this reached 0.455 by 1920 with the diffusion through the breed of the beefy Scotch strain, centering in this bull.

In the earlier years and one later period there was excess contemporary inbreeding ($F_{IS} > +0.043$) indicating the building up of somewhat inbred strains such as the Duchesses. In two periods, however, F_{IS} was actually negative, indicating a slight excess tendency toward crossing differentiated strains. Thus the relative steadiness in the rise of F_{ST} contrasts with wide fluctuations in F_{IT}.

A phenotypically somewhat differentiated group, those with registered milking records (McPhee and Wright, 1926) showed little differentiation in pedigree from the contemporary breed (1920) but showed relatively high F_{IS} suggesting a slight tendency toward segregation, substantiated by a relatively low relationship to Champion of England.

PART III. APPLICATION OF *F*-STATISTICS TO SYSTEMS OF MATING

In applying the set of *F*-statistics to systems of mating, the closed lines are to be considered as subdivisions (S) of an indefinitely large random array of such lines (T), derived similarly from a random breeding foundation stock. The inbreeding coefficients are thus of the type F_{IT}. F_{ST} is the average of the correlations among all gametes of a line relative to this total. It

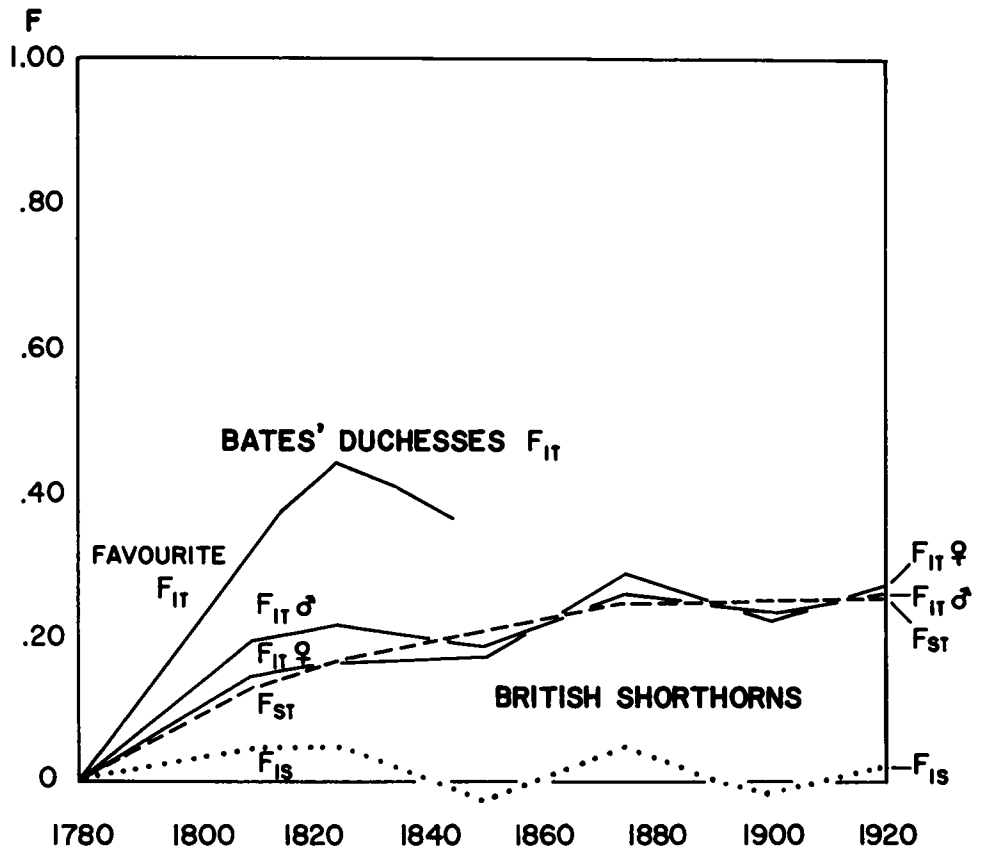


FIG. 6.

might seem that with separate sexes, one should average only the correlations between eggs and sperms, but if a very large random breeding population is to be produced in a given generation, the gametes of the N zygotes should contribute equally. It is the average correlation among gametes of such a random breeding strain from each line, relative to the total of all possible lines that constitutes F_{ST}. F_{IS} is the average correlation between uniting gametes relative to the gene frequencies of their own lines.

It will be convenient to begin with the extreme case of sib-mating. There are two kinds of correlations in a generation: between gametes of the same individual (r_0) and between gametes of different individuals ($r_1 = F_{IT}$).

$$r_0 = \tfrac{1}{2}(1 + r_1')$$
$$F_{IT} = r_1 = \tfrac{1}{2}(r_0' + r_1')$$
$$= \tfrac{1}{4}[2F' + F'' + 1],$$
$$P_{IT} = \tfrac{1}{2}P_{IT}' + \tfrac{1}{4}P_{IT}''$$

Since these types are equally numerous

$$F_{ST} = \tfrac{1}{2}(r_0 + r_1) = \tfrac{1}{4}[2F + F' + 1]$$
$$P_{ST} = \tfrac{1}{2}P_{IT} + \tfrac{1}{4}P_{IT}'$$
$$P_{IS} = P_{IT}/P_{ST} = P_{ST}'/P_{ST}$$

The values of F_{IT}, F_{ST}, and F_{IS} in seven generations of sib mating are given in Table 2.

The correlation (F_{IT}) between uniting gametes, relative to the total, lags a generation behind that (F_{ST}) between random gametes, also relative to the total and is smaller by a ratio that oscillates about λ ($= \tfrac{1}{4}[1 + \sqrt{5}] = 0.80902$). The average

TABLE 2. *Values of* F_{IT}, F_{ST}, *and* $F_{IS} = (F_{IT} - F_{IS})/(1 - F_{IS})$ *by generation.*

	0	1	2	3	4	5	6	7
F_{IT}	0	0.2500	0.3750	0.5000	0.5937	0.6719	0.7344	0.7852
F_{ST}	0.2500	0.3750	0.5000	0.5937	0.6719	0.7344	0.7852	0.8262
F_{IS}	–0.3333	–0.2000	–0.2500	–0.2308	–0.2381	–0.2353	–0.2364	–0.2360

correlation between uniting gametes (F_{IS}) relative to its line, is thus negative and rapidly approaches $(\lambda - 1)/\lambda = -0.23607$ in oscillatory fashion.

The question may arise as to whether F_{IS} has any meaning in such small populations. In populations consisting of one male and one female, the correlation coefficient between egg and sperm is indeterminate (0/0), unless both parents are heterozygous, in which case, it is zero. If the correlation arrays are made diagonally symmetrical by tabulating gamete against gamete, irrespective of kind, the coefficient is indeterminate only if both parents are homozygous in the same allele; but this is a class of matings that approaches 100 per cent in frequency as the inbreeding proceeds.

This difficulty disappears, however, if the various types of mating are weighted by their variances in the direct calculations of F_{IS}. The essentially different kinds of mating, their correlation arrays, their variances, and correlation coefficients are as follows, noting that A and a may be exchanged with no essential change in kind.

$AA \times AA$	$AA \times aa$	$AA \times Aa$	$Aa \times Aa$
0 1	0.50 0	0.25 0.50	0.25 0.25
0 0	0 0.50	0 0.25	0.25 0.25
0 1	0.50 0.50	0.25 0.75	0.50 0.50
$\sigma^2 = 0$	$\sigma^2 = 0.25$	$\sigma^2 = 0.1875$	$\sigma^2 = 0.25$
$r = 0/0$	$r = -1$	$r = -0.3333$	$r = 0$

The frequencies of these types in seven successive generations of sib mating and the weighted averages of the coefficients are given in Table 3.

It may be seen that the actual weighted averages of the correlations between uniting gametes agree exactly with the theoretical values. F_{IS} can thus be given a concrete meaning even in this very extreme case. It cannot, of course, be interpreted as a probability because of its negative value.

In maximum avoidance systems with population number $N = 2^m$ (illustrated in Fig. 1 for $N = 8$), there are $m + 1$ different kinds of correlations between gametes of the same generation. Labeling these r_0, $r_1 \cdots r_m$ in order of remoteness, $F_{IT} = r_m = \frac{1}{2}(r_{m-1}' + r_m')$ and r_0, which as always is $\frac{1}{2}(1 + F_{IT}')$, can be written $\frac{1}{2} \times (1 + r_m')$. All of the others are of the form $r_x = \frac{1}{2}(r_{x-1}' + r_m')$. $F_{IT} = \frac{1}{2}F_{IT}' + \frac{1}{4}F_{IT}'' \cdots \left(\frac{1}{2^{m+1}}\right)F_{IT}^{(m+1)primes}$ or, after starting, $F_{IT}' - [1/(4N)]F_{IT}^{(m+2)primes}$. $P_{IT} = P_{IT}' - [1/(4N)]P^{(m+2)primes}$. Thus k $(= 1 - \lambda)$ approaches 1 $(4N)$ asymptotically (Wright, 1933a). A closer approximation is given by the equation $k = 1/[4N(1-k)]^{m+1}$ or approximately $k \approx [1/(4N)][1 + k]^{m+1} \approx [1/(4N)][1 + (m+1)k]$, $k = 1 - \lambda \approx 1/[4N - (m+1)]$ (communicated by Alan Robertson).

The average correlation between gametes is given by $F_{ST} = (1/N)[r_0 + r_1 + 2r_2 + 4r_3 \cdots 2^{m-1}r_m]$.

All of the r's can readily be expressed in terms of F_{IT} $(= r_m)$ of the same and following generations ($r_{m-1}' = 2r_m - r_m'$ etc.) and similarly with the P's. Thus in the case of 16-fold fourth cousins $P_{ST} = \frac{1}{128}[64P_{IT} + 32P_{IT}' + 15P_{IT}'' + 7P_{IT}''' + 3P_{IT}^{IV} + P_{IT}^{V}]$ and $P_{IS} = P_{IT}/P_{IS}$: limiting value, $128\lambda^5/[64\lambda^5 + 32\lambda^4 + 15\lambda^3 + 7\lambda^2 + 3\lambda + 1] = 1.04207$. $F_{IS} = 1 - P_{IS}$: limiting value, –0.04207. Values of $1 - \lambda$ are given in Table 4 and of F_{IS} in Table 5 for various values of N.

In the case of circular mating of half-sibs, population number $N = 2^m$ (illustrated in Fig. 2 for $N = 8$), there are $[(N/2) + 1]$ different kinds of correlations. Letting $n = (N/2)$ and numbering

TABLE 3. *Frequencies of types of matings by generations.*

	$r\ (=F_{IS})$	σ^2	0	1	2	3	4	5	6	7
$AA \times AA$, $aa \times aa$	0/0	0	0.1250	0.2812	0.4141	0.5254	0.6157	0.6891	0.7484	0.7965
$AA \times aa$	−1	0.2500	0.1250	0.0313	0.0391	0.0254	0.0220	0.0172	0.0141	0.0113
$AA \times Aa$, $aa \times Aa$	−0.3333	0.1875	0.5000	0.3750	0.3437	0.2734	0.2246	0.1812	0.1469	0.1187
$Aa \times Aa$	0	0.2500	0.2500	0.3125	0.2031	0.1758	0.1377	0.1125	0.0906	0.0734
			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$\overline{F}_{IS} = \Sigma r\sigma^2 f / \Sigma \sigma^2 f$			−0.3333	−0.2000	−0.2500	−0.2308	−0.2381	−0.2353	−0.2364	−0.2360

these from r_0 to r_n, $F_{IT} = r_1$ and $r_0 = \frac{1}{2}(1 + r_1')$. The smallest of these correlations is $r_n = \frac{1}{2}(r_{n-1}' + r_n')$. All the rest are of the type $r_x = \frac{1}{4}[r_{x-1}' + 2r_x' + r_{x+1}']$. On making the substitution $r_x = 1 - P_x$, $P_0 = \frac{1}{2}P_1'$, and the others are as above with replacement of r's by P's. The values can be calculated generation after generation. The limiting ratio $\lambda = P/P'$ was obtained by Kimura and Crow by arranging equivalents of these P-equations in a matrix. The characteristic equation could then be expressed as a function of a known determinant (Wolstenholme's) on making the substitution $\lambda = \frac{1}{2}(1 + \cos\theta)$. This yielded the solution $\sin\theta = \cot n\theta$.

It is not, however, necessary to set up the matrix and characteristic equation, given this substitution. The equation $P_n = \frac{1}{2}(P_{n-1}' + P_n')$ can be written $P_{n-1}' = P_n'(2\lambda - 1) = P_n' \cos\theta$. The equation $P_{n-1} = \frac{1}{4}[P_{n-2}' + 2P_{n-1}' + P_n']$ can be written $P_{n-2}' = P_n'[2\cos^2\theta - 1] = P_n' \cos 2\theta$. The equation $P_{n-2} = \frac{1}{4}[P_{n-3}' + 2P_{n-2}' + P_{n-1}']$ can be written $2P_n' \cos 2\theta \cos\theta = P_{n-3}' + P_n' \cos(2\theta - \theta)$ giving $P_{n-3}' = P_n'(\cos 2\theta \cos\theta - \sin 2\theta \sin\theta) = P_n' \cos 3\theta$. In general $P_x = P_n \cos[(n-x)\theta]$, including $P_0 = P_n \cos n\theta$. Also, $P_0 = \frac{1}{2}P_1' = P_n \cos[(n-1)\theta]/(1 + \cos\theta)$.

$\text{Cos } n\theta(1 + \cos\theta) = \cos[(n-1)\theta] = \cos n\theta \cos\theta + \sin n\theta \sin\theta$; $\sin\theta = \cos n\theta / \sin n\theta$ as given by Kimura and Crow.

With large n, θ becomes small, $\sin\theta$ and $\cos n\theta$ approach θ, and $\sin n\theta$ approaches 1. Thus θ approaches $\pi/[2(n+1)]$, $\cos\theta$ approaches $\sqrt{1-\theta^2} \approx 1 - \theta^2/2$ and $1 - \lambda = \frac{1}{2}[1 - \cos\theta]$ approaches $[\pi/4(n+1)]^2$ or $[\pi/(2N+4)]^2$, again as given by Kimura and Crow.

All of the correlations appear twice around the circle except r_0 and r_n since $r_{N-x} = r_x$. Thus $F_{ST} = (1/N)[r_0 + 2\sum^{n-1} r_x + r_n]$. P_{IS} approaches a limiting value that can be obtained as the limiting value of $NP_1/[P_0 + 2\sum^{n-1} P_x + P_0]$ in terms of the above cosine formula, after determining the value of θ. Values of $1 - \lambda$ and F_{IS} are given in Tables 4 and 5 respectively.

In the case of circular mating of first cousins, $N = 2^m$, there are $(N/4) + 2$ different kinds of correlations between gametes. Letting $n = N/4$ in this case, and numbering the correlations according to remoteness, $F_{IT} = r_2$ and thus $r_0 = \frac{1}{2}(1 + r_2')$. The correlation between gametes of sibs is $r_1 = \frac{1}{2}(r_0' + r_2')$. The correlation between the most remote gametes is $r_{n+1} = \frac{1}{2}[r_n' + r_{n+1}']$. All of the others are of the type $r_x = \frac{1}{4}[r_{x-1}' + 2r_x' + r_{x+1}']$. Replacing r_x by $1 - P_x$, $P_0 = \frac{1}{2}P_2'$ and in the others merely replace r's by corresponding P's. Kimura and Crow again solved for the limiting ratio $\lambda = P/P'$ by setting up the P-matrix and making the same substitution, $\lambda = \frac{1}{2}(1 + \cos\theta)$, as in the preceding case.

TABLE 4. *The limiting rates of decrease of heterozygosis $(1-\lambda)$ under six closed systems of mating at different sizes of population (N). In the first two systems, the parents are drawn at random ($\sigma_k^2 = 2$); in the others just two are used from each parent of the preceding generation ($\sigma_k^2 = 0$).*

N	1 ♂ $(N-1)$ ♀'s $\sigma_k^2 = 2$	$N/2$ ♂'s $N/2$ ♀'s $\sigma_k^2 = 2$	Maximum avoidance $(\sigma_k^2 = 0)$	$N/2$ ♂'s $N/2$ ♀'s $(\sigma_k^2 = 0)$	Circle of first cousins $(\sigma_k^2 = 0)$	Circle of half sibs $(\sigma_k^2 = 0)$
2	0.1910	0.1910	0.1910	0.1910	—	—
4	0.1396	0.1096	0.0804	0.0764	0.0804	0.0727
8	0.1228	0.0586	0.0362	0.0344	0.0347	0.0249
16	0.1159	0.0303	0.0170	0.0164	0.0142	0.0076
32	0.1120	0.0154	0.0082	0.0080	0.0053	0.0021
Large N	0.1096	$1/[2N+1]$	$1/[4N-(m+1)]$	$1/[4N-3]$	$[\pi/(N+12)]^2$	$[\pi/(2N+4)]^2$

By the same sort of reasoning as before, $P_n = P_{n+1}\cos\theta$ and in general $P_x = P_{n+1}\cos[(n+1-x)\theta]$ except for P_0 for which there are two equations to be solved.

$$P_0 = P_{n+1}\cos[(n-1)\theta]/(1+\cos\theta)$$
$$P_0 = 2\lambda P_1 - P_2 = P_{n+1}(1+\cos\theta)\cos n\theta - \cos[(n-1)\theta]$$

From these $\sin\theta[2+\cos\theta] = \cos n\theta/\sin n\theta$ as given by Kimura and Crow.

This requires that θ be approximately $\pi/2(n+3)$, leading to the approximate value, $1-\lambda = \theta^2/4 = [\pi/4(n+3)]^2 = [\pi/(N+12)]^2$ as given by Kimura and Crow.

The correlations r_0 and r_1 appear only once around the circle, r_{n+1} appears twice while all of the others appear four times.

$$F_{ST} = \frac{1}{N}[r_0 + r_1 + 4\sum_2^n r_x + 2r_{n+1}]$$

P_{IS} approaches a limiting value that can readily be found by the limiting cosine formulae for the P's. Values of $1-\lambda$ and F_{IS} are again given in Tables 4 and 5 respectively.

TABLE 5. *The limiting correlation between uniting gametes relative to the array of gametes of the same line (F_{IS}) in systems in which there is not random mating.*

N	Maximum avoidance	Circle of first cousins	Circle of half sibs
2	−0.2361	—	—
4	−0.1824	−0.1824	−0.0784
8	−0.1170	−0.0733	+0.2221
16	−0.0711	+0.1011	+0.5160
32	−0.0421	+0.3293	+0.7274

Fig. 7 compares the modes of approach toward fixation (F_{IT}) under various systems of mating in populations of eight individuals, starting in all cases from the first generation that shows an increase and thus not allowing for lags on starting from a random breeding stock. The most rapid of these is that with one male and seven females with replacement by a random male and seven random females. The next is similar except that there are four males and four females. In terms of N_m males, N_f females (Wright, 1931).

$$F = F' + \tfrac{1}{2}N_e(1 - 2F' + F''),$$
$$N_e = \frac{4N_mN_f}{N_m + N_f}$$
$$(1-\lambda) = \tfrac{1}{2}\left[1 + \frac{1}{N_e} - \sqrt{1 + \frac{1}{N_e^2}}\right] \approx \frac{1}{2N_e + 1}$$

If $N_m = 1$, $N_f = 7$, $(1-\lambda) = \tfrac{1}{4}[9 - \sqrt{53}] = 0.1228$

If $N_m = 4$, $N_f = 4$, $(1-\lambda) = \tfrac{1}{16}[9 - \sqrt{65}] = 0.0586$

The maximum avoidance system (quadruple second cousins) comes next $(1-\lambda = 0.03622)$, but differs little from the following two: circular mating of first cousins $(1-\lambda = 0.03475)$ and mating of four males and four females with replacement by selection of just two offspring per parent in all three. In the last case, the effective number in relation to the system of completely random replacement is given by Kimura and Crow (1963a) as $(4N-4)/(2+\sigma_k^2)$ which with $\sigma_k^2 = 0$ give $2N-2$

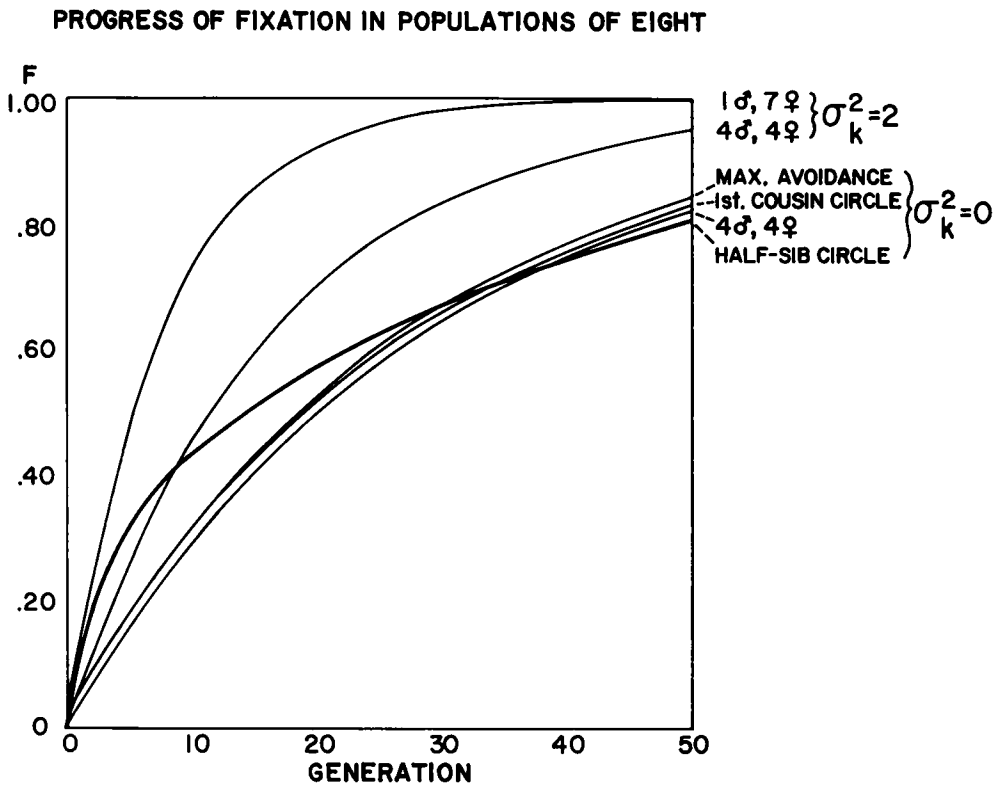


FIG. 7.

and thus 14 if $N = 8$. This replaces N_e in the formula above by 14.

$$1 - \lambda = \tfrac{1}{28}[15 - \sqrt{197}] = 0.0344$$

This differs little from $1/(4N - 3)$ which $1 - \lambda$ approaches with large N.

Progress under circular mating of half-sibs follows a very different course from any of the others. There is almost as rapid an early rise as with one male, seven females followed by crossing of all of the other lines because of much the lowest ultimate rate (0.0249).

Table 4 compares the limiting rates of decrease of heterozygosis, $1 - \lambda$ in various sizes of population. This is always slightly greater under maximum avoidance ($N > 2$) than under random mating with equal numbers of males and females and selection of just two offspring per parent in both. Circular mating of first cousins is the same system as maximum avoidance if $N = 4$ but has a slight lower rate if $N = 8$. It has a very slightly higher rate (if $N = 8$) than under the above random system. With $N = 16$ or more the rate becomes less than either of the preceding. The limiting rate under circular half-sib mating is always the lowest of those considered here and becomes very much so as N is increased. The greater tendency under this system to maintain heterozygosis than under maximum avoidance was the paradox discussed by Kimura and Crow (1963b).

Table 5 gives the correlation, F_{IS}, between uniting gametes relative to their own lines in the three systems in which there is not random mating. F_{IS} is, of course, always negative under maximum avoidance but it rapidly approaches zero as N is increased.

Circular first cousin mating passes from

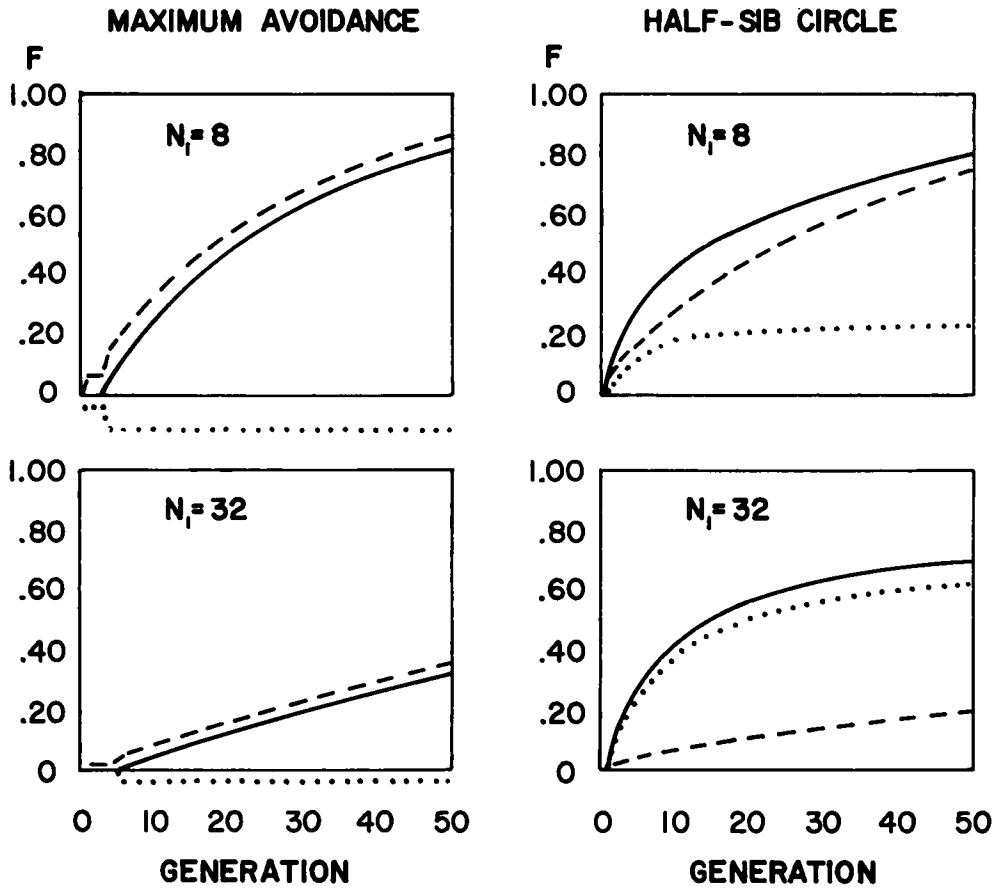


FIG. 8. Solid lines = F_{IT}; broken lines = F_{ST}; dotted lines = F_{IS}.

negative F_{IS} if $N = 4$ (also maximum avoidance) to positive at some population size between 8 and 16. Even with circular half-sib mating F_{IS} is negative if $N = 4$ but becomes positive between this and $N = 8$ and reaches the high value +0.73 if $N = 32$.

There is one property of groups of size N that has been found to be the same whether mating of consanguineous individuals is avoided as much as possible, or the reverse, or is at random. It was shown in a previous paper (Wright, 1933b) that under maximum avoidance an array of lines that started from the mating AB/ab, where A and B show the proportion c of recombination, ultimately arrive at correlation $r_\infty = (1 - 2c)/[1 + 2(2N - 1)\ c]$ and a proportion of recombinant lines of $2X_\infty = \frac{1}{2}(1 - r_\infty) = 2Nc/[1 + 2(2N - 1)\ c]$ when fixation is complete. Thus under double first cousin mating ($N = 4$) $r_\infty = (1 - 2c)/(1 + 14c)$ and $2X_\infty = 8c/(1 + 14c)$, (the latter correcting an error in the paper as published). Kimura (1963), applying the probability concept arrived at exactly the same general formula for the proportion of recombinants in the case of circular half-sib mating as given above for maximum avoidance. I have confirmed this by path analysis in the special case $N = 4$.

In the case of random mating in groups of N_m males and N_f females, I obtained $r_\infty = (1 - 2c)/[1 + 2(N_e + 1)\ c]$ and amount of recombination $(N_e + 2)\ c/[1 + 2(N_e + 1)\ c]$ where $N_e = 4N_mN_f/(N_m + N_f)$. This

has been confirmed by Kimura by his method. The correlation approaches $(1-2c)/2Nc$ in large groups instead of $(1-2c)/4Nc$ as in the two preceding cases, but this is because the offspring generation was assumed to be drawn at random instead of just two from each parent. Using the formula of Kimura and Crow (1963a), $N_e = (4N-4)/(2+\sigma_k^2)$, N_e above must be replaced by $2N_e - 2$ for the case in which $\sigma_k^2 = 0$ giving $r_\infty = (1-2c)/[1+2(2N_e-1)\,c]$ exactly as under the other two systems. It appears that the ultimate result of the race between recombination and fixation is exactly the same in these three cases. The differences in rates of fixation are ultimately exactly compensated for by differences in amounts of recombination.

Fig. 8 compares the ways in which F_{IT}, F_{IS}, and F_{ST} change in the course of 50 generations after starting lines of 8 or 32 from random breeding stock under the two extreme systems. Under maximum avoidance, F_{ST} (measuring the permanent inbreeding effect) is always higher than F_{IT} (measuring the total inbreeding) because of negative F_{IS}. It is shown in Fig. 7 that in populations of eight, F_{IT} is very slightly higher under maximum avoidance than under random mating. Thus F_{ST} is still more in excess under the former since $F_{ST} = F_{IT}$ under random mating. These differences are, however, slight even in populations of eight and become less in larger populations as shown for $N = 32$.

The very different character of the progress of F_{IT} under circular half-sib mating as compared with random mating in populations of eight was brought out in Fig. 7. It is evident from Fig. 8 that this is due to the rapid rise and approach to constancy of positive F_{IS}. Because of positive F_{IS}, F_{ST} is always lower than F_{IT}.

With larger N, illustrated by the case of $N = 32$, there is an almost qualitative difference. F_{IT} is almost wholly dominated at first by the large amount of current inbreeding measured by F_{IS} while the more permanent differentiation of lines as wholes, measured by F_{ST}, builds up very slowly. The excess correlation between adjacent individuals tends to maintain different alleles in different regions around the circle. There is an approach to the situation in a population that is broken up into permanently distinct isogenic lines, which has been recognized as the best way to maintain the potentiality for maximum heterozygosis realizable by crossing, since G. H. Shull and D. F. Jones developed the theoretical basis for the enormously successful hybrid corn program. The basis for the exceedingly low rates of decrease of heterozygosis in large populations under circular systems of mating, demonstrated by Kimura and Crow, is obvious from Fig. 8.

Open systems of half-sib and first-cousin mating in infinite populations were studied as first approaches to the problem of isolation by distance. They were not very satisfactory models, however, and at the time the results were merely presented without discussion (Wright, 1921). When more satisfactory models were studied later (Wright, 1940, 1943a, 1946, 1951) it did not seem worthwhile to go back to these very artificial systems. It is, however, instructive in the present context to do so, first for the open system of half-sib mating and then for the corresponding closed circle of 32 individuals.

Fig. 9 shows the values of F under half-sib mating in an infinite population, carried to generation 50 instead of merely to generation 15 as in 1921. This is compared with the progress of F_{IT} under linear isolation by distance with neighborhoods of various sizes. Parent-offspring distances are assumed to be distributed normally, and hence variances of ancestral population rise linearly with number of generations but sizes of these populations rise only as the square roots of the latter. In the case of the half-sib system the parent-offspring distance is constant and equals 0.50 in terms of the distance between adjacent individuals. The grandparental distances have a 1 : 2 : 1 distribution about zero, the great grandparental a 1 : 3 : 3 : 1 distribu-

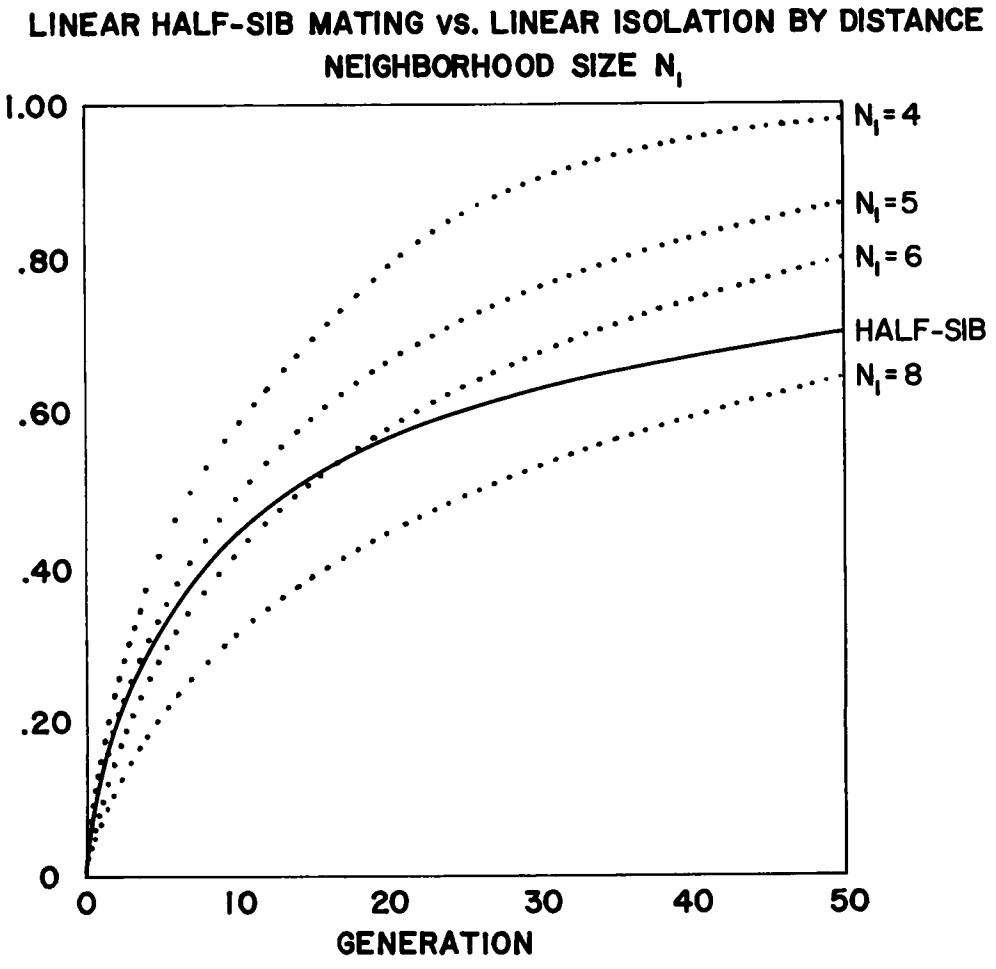


FIG. 9.

tion, the great great grandparental a 1 : 4 : 6 : 4 : 1 distribution. The variances rise linearly (0.25, 0.50, 0.75, 1.00, etc.) as in the model for linear isolation by distance. On the other hand the range also rises linearly (1, 2, 3, 4, etc.) as does the number of different ancestors (2, 3, 4, 5, etc.) in contrast with the sizes of ancestral populations in the model of linear isolation by distance that rise only as the square root of the number of ancestral generations. This difference is due to the fact that in the half-sib model the parental distribution is as platykurtic as possible and the ancestral distributions only gradually approach normality instead of all being normal.

At first sight it would appear that the size of "neighborhood" to be used in making a comparison should be two since each individual is produced by two adjacent individuals. This, however, does not allow for the differences between drawing at random from normally distributed parents and taking exactly two offspring from each parent. The effective number is thus expected to be considerably larger than two.

On calculating the values of F by generations with $N = 4$, by the formula given earlier for the linear model, it turns out that these rise much too rapidly. With $N = 5$ the values start out rather similarly (actually very slightly too small) but soon

are considerably in excess. With $N = 6$, the values are too small for 16 generations but are considerably too large at 50 generations. Even with $N = 8$, the course of F under the linear model is clearly rising toward that under the half-sib model, though still below at 50 generations.

The difference in form of the course of F under the two systems can be accounted for by the differences discussed above. In the half-sib system the effective number of ancestors in ancestral generation X is somewhere between $\sqrt{X}$, expected if parental and all ancestral distributions were normal, and X. There is in consequence greater damping of progress than under the linear model in which this number varies with $\sqrt{X}$. It should be noted that in terms of the linear model the scale of generations of Fig. 9 should be transformed to one of square roots.

The increase in F_{IT} in the circular half-sib system, $N = 32$, does not differ appreciably from that in the infinite population, up to 50 generations. It is of interest to make an analysis in terms of subdivisions of the total circle, similar to that discussed earlier for the models of isolation by distance. Since S has been used for the whole circle in relation to an infinite array of such groups, it will be convenient to use X for fractions of the circle, of length X in the sense used earlier, but giving only half weight to the terminal individuals. The smallest such fraction ($X = 1$) includes two individuals with correlation between gametes of r_0 and r_1 in equal frequencies. The next to be considered ($X = 2$) includes four individuals with gametic correlations of r_0, r_1, and r_2 in frequencies 0.25, 0.50, and 0.25 respectively. The distance $X = 16$, includes all 32 individuals. Again it is the square roots of these "distances" that correspond to the distances on the linear model.

Since the total of all groups (T) is subdivided into groups of size $S = 32$ and these into fractions (X), the panmictic index for individuals relative to the total can be analyzed into three factors, $P_{IT} = P_{IX}P_{XS}P_{ST}$.

TABLE 6. *Analysis of system of circular half-sib mating,* N = 32 *with* S *pertaining to a single group of the 50th generation,* T *to the totality of all such groups and* X *(in* F_{XT}, F_{IX}, *and* F_{XS}*) fractions of groups terminating at "distance"* x *(in* r_x*) from individuals. Terminal distances have half weight.*

Distance X	r_x	F_{XT}	F_{IX}	F_{XS}
0	0.8477	0.8477	—	—
1 = 31	0.6982	0.7730	−0.3293	0.7169
2 = 30	0.5594	0.7009	−0.0090	0.6270
3 = 29	0.4354	0.6331	+0.1776	0.5424
4 = 28	0.3288	0.5703	+0.2977	0.4641
5 = 27	0.2406	0.5132	+0.3801	0.3929
6 = 26	0.1704	0.4619	+0.4392	0.3289
7 = 25	0.1167	0.4164	+0.4829	0.2722
8 = 24	0.0772	0.3765	+0.5160	0.2224
9 = 23	0.0493	0.3417	+0.5416	0.1790
10 = 22	0.0304	0.3115	+0.5617	0.1414
11 = 21	0.0181	0.2854	+0.5777	0.1088
12 = 20	0.0104	0.2628	+0.5906	0.0806
13 = 19	0.0057	0.2432	+0.6012	0.0562
14 = 18	0.0032	0.2261	+0.6100	0.0349
15 = 17	0.0019	0.2113	+0.6174	0.0163
16	0.0015	0.1981	+0.6237	0

Table 6 shows the values of r_X for all values of X from 0 to 31 for the 50th generation based on calculations for all previous generations. F_{IT} is 0.69822 and thus $P_{IT} = 0.30178$, P_{ST} is 0.80187 and P_{IS} ($= P_{IT}/P_{ST}$) is thus 0.37635. Thus F_{ST} ($= 0.19813$) is far below its limiting value while F_{IS} ($= 0.62365$) is not very far below its limiting value 0.72736 (see Table 5). The coefficients involving X were found as follows for each X.

$$P_X = 1 - r_X$$

$$P_{XT} = \frac{1}{X}\left[\tfrac{1}{2}P_0 + \sum_{1}^{x-1} P_i + \tfrac{1}{2}P_X\right], \qquad F_{XT} = 1 - P_{XT}$$

$$P_{IX} = P_{IT}/P_{XT} = 0.30178/P_{XT}, \qquad F_{IX} = 1 - P_{IX}$$

$$P_{XS} = P_{IS}/P_{IX} = 0.37635/P_{IX}, \qquad F_{XS} = 1 - P_{XS}$$

The values of F_{XT} (Table 6) are of little interest in themselves. The correlations F_{IX} (also Table 6) between uniting gametes relative to the array of their own fractional group are, however, of interest. $F_{I1} = -0.329$ is the correlation between

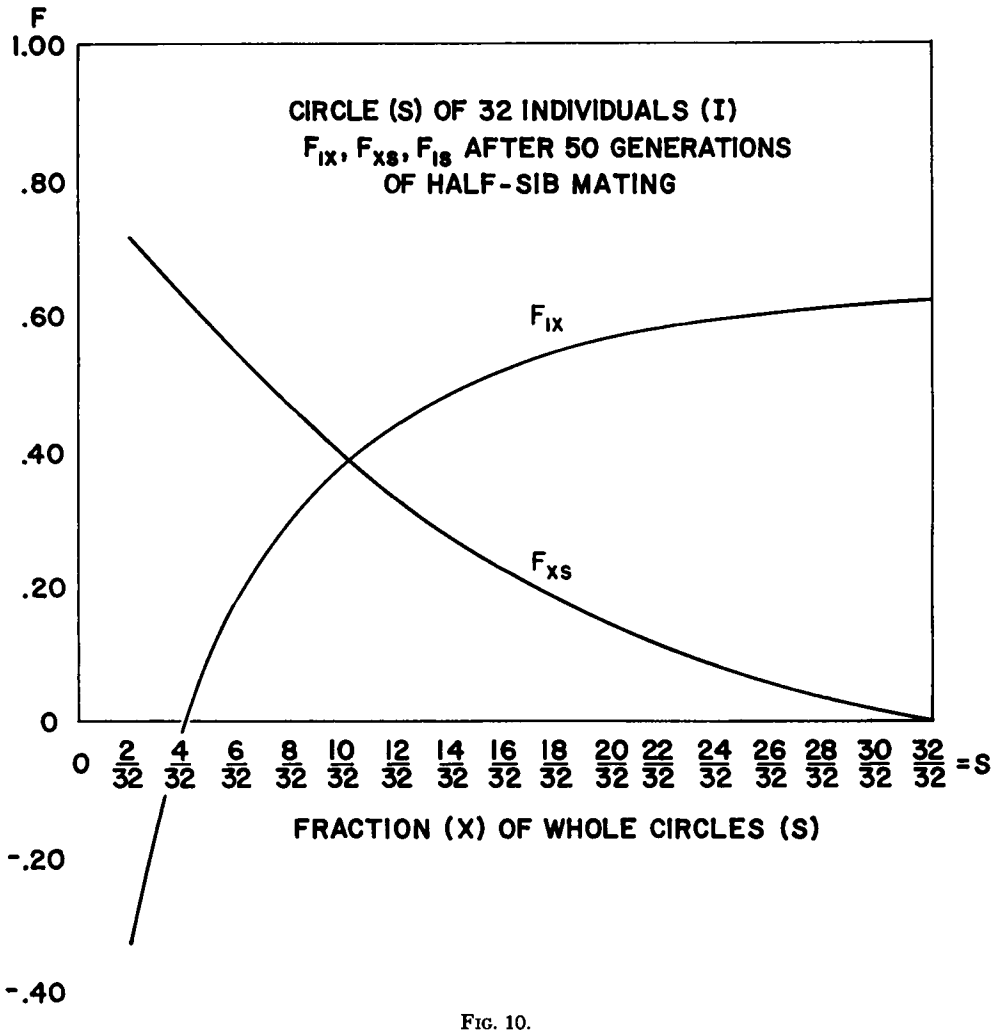


FIG. 10.

uniting gametes relative to the array of gametes from two neighboring individuals. $F_{I2} = -0.009$ is virtually zero, indicating that such groups of four individuals (effective number about six by Kimura and Crow's formula) are roughly equivalent to neighborhoods. Beyond this, F_{IX} rises gradually as X increases, toward the value $F_{IS} = 0.624$. F_{XS} in the last column of Table 6 measures the amount of differentiation of the fractional groups relative to that of complete local fixation $(q_S(1 - q_S))$. It falls from 0.717 for groups of two to 0.222 for groups of 16 (half of the circle) and to 0.016 for groups of 30. That for "neighborhoods" (groups of four) is approximately 0.627. These points are illustrated in Fig. 10, similarly to the case of isolation by distance in an area continuum in Fig. 4. This brings out in another way, the strong local differentiation within the circles that interferes with the progress of fixation of circles as wholes as measured by F_{ST}.

SUMMARY

The general conclusion of part I is that the theoretical correlation between repre-

sentatives of a locus in gametes, uniting or otherwise, relative to one or another array of such representatives (F-statistics), gives a broader basis for comparison of population structures, including progress in fixation, than does the alternative concept: the probability of identity of such representatives by origin. One reason is that correlations vary from -1 to $+1$ while probabilities vary only from 0 to $+1$. The probability concept gives, however, a very useful supplementary interpretation where applicable.

The relation of the basic set of F-statistics, F_{IT}, F_{IS}, F_{ST}, to variances within populations is discussed in part II and applications to diverse patterns of population structure are reviewed (the island model with or without selective differences, isolation by distance in continuous populations under balancing of local inbreeding and dispersion, uniformly distributed clusters under a similar balance, selective clines, breeds of livestock).

In part III, these F-statistics are applied to systems of mating in populations of given small size, in which consanguine mating is either avoided as much as possible, or pursued as much as is possible without any disruption of the group.

The apparently paradoxical result obtained by Kimura and Crow that heterozygosis declines more rapidly under the former than under the latter is discussed from the standpoint of these statistics.

These systems have been found to agree in one respect, the ultimate proportion of recombinant lines in the race between fixation and recombination among lines starting from double heterozygotes.

Literature Cited

Bartlett, M S., and J. B. S. Haldane. 1934. The theory of inbreeding in tetraploids. J. Genet., **29**: 175–180.

——. 1935. The theory of inbreeding with forced heterozygosis. J. Genet., **31**: 327–340.

Bateman, A. J. 1950. Is gene dispersion normal? Heredity, **4**: 353–363.

Bernstein, F. 1930. Fortgesetzte Untersuchungen aus der Theorie der Blutgruppen. Z. indukt. Abstamm.—u Vererb. Lehre, **56**: 233–273.

Cockerham, C. Clark. 1954. An extension of the concept for partitioning hereditary variance for analysis of covariances among relatives, when epistasis is present. Genetics, **39**: 859–882.

Fisher, R. A. 1949. The theory of inbreeding. Oliver and Boyd, Edinburgh, 120 pp.

Kempthorne, O. 1954. The correlations between relatives in a random mating population. Proc. Royal Soc. London, **B143**: 103–113.

Kimura, M. 1953. "Stepping stone" model of population. Ann. Rept. Nat. Inst. Genetics of Japan, **3**: 62–63.

——. 1963. A probability method for treating inbreeding systems, especially with linked genes. Biometrics, **19**: 1–17.

Kimura, M., and J. F. Crow. 1963a. The measurement of effective population number. Evolution, **17**: 279–288.

——. 1963b. On the maximum avoidance of inbreeding. Genet. Res. Cambridge, **4**: 399–415.

Kimura, M., and G. H. Weiss. 1964. The stepping stone model of population structure, and the decrease of genetic correlation with distance. Genetics, **49**: 561–576.

Malécot, G. 1948. Les mathématiques de l'hérédité. Masson et Cie, Paris, 63 pp.

McPhee, H. C., and S. Wright. 1925. Mendelian analysis of the pure breeds of livestock III. The Shorthorns. J. Hered., **16**: 205–215.

——. 1926. Mendelian analysis of the pure breeds of livestock IV. The British Dairy Shorthorns. J. Hered., **17**: 397–401.

Robertson, A. 1952. The effect of inbreeding on the variation due to recessive genes. Genetics, **37**: 189–207.

Wahlund, S. 1928. Zuzammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas, **11**: 65–106.

Wright, S. 1921. Systems of mating. Genetics, **6**: 111–178.

——. 1922a. Coefficients of inbreeding and relationship. Amer. Nat., **56**: 330–338.

——. 1922b. The effects of inbreeding and crossbreeding on guinea pigs. Bull. U. S. Dept. Agric., **1121**: 1–59.

——. 1923a. Mendelian analysis of the pure breeds of livestock I. The measurement of inbreeding and relationship. J. Hered., **14**: 339–348.

——. 1923b. Mendelian analysis of the pure breeds of livestock II. The Duchess family of Shorthorns as bred by Thomas Bates. J. Hered., **14**: 405–422.

——. 1931. Evolution in Mendelian populations. Genetics, **16**: 97–159.

——. 1933a. Inbreeding and homozygosis. Proc. Nat. Acad. Sci., **19**: 411–420.

——. 1933b. Inbreeding and recombination. Proc. Nat. Acad. Sci., **19**: 420–433.

——. 1938a. The distribution of gene frequen-

cies under irreversible mutation. Proc. Nat. Acad. Sci., **24**: 253–259.

——. 1938b. The distribution of gene frequencies of polyploids. Proc. Nat. Acad. Sci., **24**: 372–377.

——. 1938c. Size of population and breeding structure in relation to evolution. Science, **87**: 430–431.

——. 1939. Statistical genetics in relation to evolution. Exposés de biometrie et de statistique biologique. No. 802, 63 pp., Hermann and Cie, Editeurs.

——. 1940. Breeding structure of populations in relation to speciation. Amer. Nat., **74**: 232–248.

——. 1943a. Isolation by distance. Genetics, **28**: 114–138.

——. 1943b. An analysis of local variability of flower color in Linanthus Parryae. Genetics, **28**: 139–156.

——. 1946. Isolation by distance under diverse systems of mating. Genetics, **31**: 39–59.

——. 1951. The genetical structure of populations. Ann. Eugenics, **15**: 323–354.

——. 1952. The theoretical variance within and among subdivisions of a population that is in a steady state. Genetics, **37**: 312–321.

——. 1954. The interpretation of multivariate systems. Chapter 2, pp. 11–33, *in* Statistics and mathematics in biology, edited by O. Kempthorne, W. A. Bancroft, J. W. Gowen, and J. L. Lush.

——. 1963a. Discussion of "Systems of mating used in mammalian genetics" by E. L. Green and D. P. Doolittle. *In* Methodology in mammalian genetics, pp. 41–53, edited by W. J. Burdette. Holden-Day, Inc., San Francisco.

——. 1963b. Discussion of "Estimation of genetic variances" by C. C. Cockerham. *In* Statistical genetics and plant breeding, pp. 93–94, edited by W. D. Hanson and H. T. Robinson. Publ. No. 982, Nat. Acad. Sci. Nat. Res. Council, Washington, D. C.

WRIGHT, S., AND H. C. MCPHEE. 1925. An approximate method of calculating coefficients of inbreeding and relationship. J. Agric. Res., **31**: 377–383.