

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Examining perennial ryegrass (*Lolium perenne*
L.) persistence through identifying genetic
shifts within two cultivars after nine years in
the field.**

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

In

Plant Breeding

At Massey University,

Palmerston North, New Zealand

Jessica Richelle O'Connor

2019

ABSTRACT

Perennial ryegrass (*Lolium perenne* L.) is a commercially important forage species in New Zealand agriculture. Ryegrass persistence is important for farmers as it substantially decreases the costs associated with reseeded pastures. Breeding for ryegrass persistence is difficult because of the complex interaction between genotype and environment; and the short time of field trial assessment compared to the expected longevity of pasture. A nine year old cultivar comparison trial at Poukawa, Hawkes Bay, New Zealand was identified and plants surviving in the cultivar plots of 'Grasslands Samson' and 'Commando' were retrieved. These populations that had survived were termed Persistent. A sample of commercially sourced seed of these cultivars were also grown to represent the 'Original' genetic pool of the cultivars sown in the field. Persistent populations were compared to Original cultivar seed to characterise morphology and underlying genetics associated with persistence. Results were interpreted to determine if a genetic shift had occurred in Persistent populations due to advantageous phenotypes surviving.

Three methodologies were used to compare populations: 1) In a glasshouse, eight morphological traits were measured after 10 weeks growth for Original and Persistent populations of 'Grasslands Samson'; 2) Half-sibling families were generated from Persistent and Original populations for both cultivars and were assessed for additive genetic variation of seven traits as one metre rows in the field over 13 months; 3) Simple sequence repeat (SSR) markers were used to explore the genetic composition of Original and Persistent populations of each cultivar. Analysis and interpretation of data showed genetic shifts were cultivar specific. The greatest differences were identified between populations of 'Grasslands Samson'. Compared to the Samson Original population, Samson Persistent plants had significantly greater means for four traits in the glasshouse

and half-sibling families showed evidence of shifting population means of traits associated with animal grazing avoidance. SSR marker results were confounded by late detection of contamination in samples. Analysis of a reduced sample size showed no significant differences between any of the four populations using F statistics and genetic structure analysis.

These results suggest future studies could reduce risk of contamination by collecting single tillers from the field of Persistent populations. Further investigation of the genetics of persistence should focus on the role of lamina sheath lengths in tiller production, and using the half-sibling families identified in this study for germplasm development and quantification of genotype-by-environment interactions.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisors Dr. Zulfi Jahufer, Dr. Vaughan Symonds, and Dr Jim Crush. With their patient guidance and expertise I have been able to undertake this thesis and see it through.

I am grateful for all the help and technical skills I learnt in the lab from Prashant Joshi and Won Hong. I would also like to thank Lily Ouyang for all her help with setting up and maintaining the glasshouse trial in Hamilton and Trevor Holloway for doing the tractor work and organising grazing for the field trial when required. I am particularly grateful for all of Catherine McKenzie and Tom Lyons help and advice. Tom and Zulfi have both inspired and stimulated my passion for plant breeding.

I would like to acknowledge AgResearch for funding of this project through the SSIF Forage Value programme, (Contract No. A19054) and for the use of the facilities at Grasslands campus to complete this work.

Finally, I would like to thank my friends and family who have supported me through my studies. Especially Rowan who experienced every emotional high and low of this process with me.

ABBREVIATIONS

%amp	Success rate of marker amplification
%P	Percentage of polymorphic loci
μ	Population mean
AFLP	Amplified Fragment Length Polymorphism
AMH	Aftermath heading score
AMOVA	Analysis of molecular variance
AR1	Commercially available endophyte strain in perennial ryegrass cultivars
BLUE	Best linear unbiased estimator
BLUP	Best linear unbiased predictor
bp	Base pairs
CASS	Cheaply amplified size standard
CV_A	Co-efficient of additive variation
df	Degrees of freedom
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DW	Dry weight (per plant)
DWT	Dry weight (per one metre row)
E	Environment
EST-SSRs	Expressed sequence tag simple sequence repeats
Est. Var.	Estimated variance components
F_{IS}	Component of Wright's (1921) fixation index, used to define within population structure by calculating the average observed heterozygosity of an individual relative to the expected heterozygosity of individuals in the population it belongs to.
F_{ST}	Component of Wright's (1921) fixation index, used to define between population structure by comparing the expected heterozygosity of individuals within a subpopulation to the total expected heterozygosity of individuals across all populations
G	Genotype
G_C	Genetic gain per cycle
h^2_n	Narrow-sense heritability
H_E	Expected heterozygosity

HGs	Herbage growth score	
H _o	Observed heterozygosity	
kg ha ⁻¹	Kilograms per hectare	
LL	Leaf length	
LnP(D)	Mean posterior probability	
LT	Leaf thickness	
LW	Leaf width	
LWs	Leaf width score	
MAF QT	Ministry of Agriculture and Forestry quick test, a soil testing metric	
mL ha ⁻¹	Millilitres per hectare	
MS	Mean of squares	
N	Number of samples	
Na	Number of alleles	
Na _v	Total number of allele variants (within a cultivar)	
N(0,σ ² ε)	Normally and independently distributed	
N-P-K-S	Units of nitrogen, phosphorus, potassium, and sulphur. Used to express fertiliser contents.	
n _r	Number of replications	
n _s	Number of seasons	
P	Phenotype	
PC	Principal components	
PCA	Principal components analysis	
PCR	Polymerase chain reaction	
PHs	Plant habit score	
RAPD	Random amplified polymorphic DNA	
REML	Residual maximum likelihood	
RP	Percentage of reproductive tillers	
RT	Number of reproductive tillers.	
RTD	Reproductive tiller development score	
Ru	Rust score	
S	Selection differential (average superiority of the selected parents), used to predict genetic gain per cycle	S

SA_P	Proportion of shared alleles
SL	Leaf sheath length
SS	Sum of squares
SSRs	Simple sequence repeats, a type of microsatellite marker.
TN	Tiller number
Var. %	Percentage of variation
σ^2_ϵ	Experimental error variance
σ^2_a	Additive genetic variance
σ^2_{as}	Additive additive-by-season interaction variance
σ^2_D	Dominance variance
σ^2_g	Genetic variance
σ^2_P	Phenotypic variance
\bar{x}	Population mean

TABLE OF CONTENTS

Contents

Abstract	iii
Acknowledgements	v
Abbreviations	vi
Table of Contents	ix
List of Figures	xv
List of Tables.....	xix
List of Equations	xxii
1.0 Introduction	1
1.1 Perennial ryegrass	1
1.2 Perennial ryegrass in New Zealand.....	2
1.3 Breeding in New Zealand.....	5
1.4 Pasture persistence	8
1.4.1 Plant morphology	9
1.4.2 Heading dates	12
1.4.3 Endophyte compatibility	13
1.4.4 Long term data sets	15
1.5 Genetic Markers	16
1.5.1 Microsatellites to assess genetic variation	18
1.6 Focus of this research.....	20

2.0 Comparing plant morphology of Persistent Ryegrass plants to plants of the Original cultivar under glasshouse conditions.	21
2.1 Abstract	21
2.2 Introduction	22
2.2.1 Objectives.....	23
2.3 Materials and methods	24
2.3.1 Plant material	24
2.3.2 Experimental design.....	25
2.3.3 Measurements	25
2.3.4 Data analysis	27
2.4 Results	29
2.4.1 Within population variation	29
2.4.2 Multivariate trait analysis.....	30
2.4.2.1 Original population	30
2.4.2.2 Persistent population	32
2.4.3 Between population variation	34
2.5 Discussion	36
2.5.1 Persistent population means were significantly different ($P<0.05$) for four traits when compared to the Original population.....	36
2.5.2 Persistent plants had significantly ($P=0.05$) greater total tiller number compared to the Original population.	37
2.5.3 Tiller number influenced reproductive tiller number.....	39

2.5.4 Sheath length.....	39
2.5.5 Greater tiller numbers contributed to greater dry weights.	40
2.5.6 The Persistent population did not have less within population phenotypic variation.....	41
2.6 Conclusions.....	42
3.0 Comparative genetic analyses of the Persistent and Original Populations.	44
3.1 Abstract.....	44
3.2 Introduction.....	45
3.2.1 Objectives.....	48
3.3 Methods.....	48
3.3.1 Long term trial at Poukawa research station.....	48
3.3.2 Trial site.....	50
3.3.3 Experimental design.....	51
3.3.4 Trial management.....	52
3.3.6 Data collection.....	53
3.3.6.1 Herbage growth score.....	53
3.3.6.2 Leaf width score.....	53
3.3.6.3 Plant habit score.....	54
3.3.6.4 Reproductive tiller development score.....	54
3.3.6.5 Aftermath heading score.....	54
3.3.6.6 Rust score.....	56
3.3.6.7 Dry weight.....	56

3.3.7 Data analysis	57
3.3.7.1 Residual maximum likelihood analysis	57
3.3.7.2 Narrow-sense heritability	58
3.3.7.3 Co-efficient of additive variation	59
3.3.7.4 Cluster analysis	59
3.3.7.5 Principal components analysis	59
3.4 Results	60
3.4.1 Additive variation	60
3.4.2 Trait associations and cluster analysis	64
3.4.2.1 Commando Original.....	64
3.4.2.2 Commando Persistent	66
3.4.2.3 Samson Original.....	68
3.4.2.4 Samson Persistent	70
3.5 Discussion	72
3.5.1 Genetic variation	72
3.5.2 Plant habit and leaf widths	73
3.5.3 Reproductive traits	75
3.5.4 Herbage growth scores.....	77
3.5.5 Trait associations and cluster analyses.....	78
3.6 Conclusions.....	81
4.0 Comparing Original and Persistent population genetics using SSR markers.	82
4.1 Abstract	82

4.2 Introduction	83
4.2.1 Objectives.....	87
4.3 Methods.....	87
4.3.1 Plant material and DNA collection	87
4.3.2 DNA extraction and genotyping	88
4.3.3 Markers	90
4.3.3.1 Scoring marker peaks	91
4.3.4 Identifying contaminated samples	92
4.3.4 Assessing genetic variation	93
4.3.5 Resolving genetic structure and differentiation	94
4.4 Results	94
4.4.1 Marker variation.....	94
4.4.1.1 Samson Original.....	95
4.4.1.2 Samson Persistent	96
4.4.1.3 Commando Original.....	97
4.4.1.4 Commando Persistent	98
4.4.2 Population variation	99
4.4.3 Genetic structure	100
4.5 Discussion	105
4.5.1 Contamination reduced sample size from 185 to 58 genotypes in total	105
4.5.2 SSR marker screens	107
4.5.3 Population variation	108

4.5.4 F_{ST} and population structures	110
4.6 Conclusions.....	112
5.0 Conclusions.....	114
5.1 Introduction.....	114
5.2 Findings.....	116
5.3 Limitations	118
5.4 Future Directions and Recommendations	120
6.0 References.....	123
Appendix 1; Layout of half-sibling family field experiment.....	132
Appendix 2: Sample field trial growth score and leaf width calibration data.....	133

LIST OF FIGURES

1.1; Perennial ryegrass (<i>L. perenne</i>) grows by tillering. Tillers consist of A) Roots; B) Terminal apex; C) Folded lamina sheaths; D) Lamina (leaf blade); E) Daughter tiller	1
2.1; Lamina morphology of a perennial ryegrass tiller. Measurements were taken from three mature tillers of each plant. Data collected included a) lamina length; b) lamina sheath length, c) lamina width; and d) lamina thickness	26
2.2; Biplot generated using standardised Best Linear Unbiased Estimate values for five traits measured from the 30 perennial ryegrass genotypes of the Original population. PC1 accounted for 38.7% and PC2 32.7% of the variation present. Traits are indicated by the directional vectors: Total number of tillers (TN), lamina length (LL), lamina sheath length (SL), lamina width (LW), and lamina thickness (LT). Colours indicate cluster groups: Red = Group 1; Green = Group 2; Blue = Group 3.....	31
2.3; Biplot generated using standardised Best Linear Unbiased Estimate values for five traits measured from the 30 perennial ryegrass genotypes of the Persistent population. PC1 accounted for 42.7% and PC2 27.4% of the variation present. Traits are indicated by the directional vectors: Total number of tillers (TN), lamina length (LL), lamina sheath length (SL), lamina width (LW), and lamina thickness (LT). Colours indicate cluster groups: Red = Group 1; Green = Group 2; Blue = Group 3.....	33
2.4; Box plots of each population for each trait. The traits presented are a) tiller number; b) reproductive tiller number; c) percentage of reproductive tillers; d) lamina length; e) lamina sheath length; f) lamina width; g) lamina thickness; and h) dry weight	35

3.1; Four pollen exclusion tents were used to generate half-sibling families. Each population had 60 plants placed into a pollen exclusion tent in the summer of 2014/15.....50

3.2; Monthly rainfall (mm) and maximum and mean monthly minimum air temperature at AgResearch Palmerston North. During the trial period May 2016-July 2017.....51

3.3; After grazing, plants were mowed to 4cm to homogenise the height of the rows52

3.4; One metre long half-sibling rows were visually scored for plant habit, leaf width, and herbage growth before each defoliation53

3.5; Reproductive tiller development (RTD) was scored in mid-November 2016. Reproductive tillers were assessed on development stage. a) Score of 1- some tillers elongated; b) Score of 2- most tillers elongated, tip of spikelet visible; c) Score of 3 - spikelets clearly visible from elongated tillers; d) Score of 4 - spikelets fully emerged; and e) Score of 5 - spikelets fully emerged and flowering (anthers visible)55

3.6; Summer dry weight harvested in February 2017. Each row was harvested with electric shears. All harvested herbage from each row were placed in separate perforated bags prior to drying56

3.7; Biplots generated using standardised Best Linear Unbiased Predictor values for three traits from the 30 half-sibling families of the Commando Original population. PC1 accounted for 49% and PC2 29% of the variation present. The different symbols indicate groups 1 to 4 generated from cluster analysis. Traits are indicated by the directional vectors: herbage growth (HGs), leaf width score (LWs), and plant habit score (PHs)65

3.8; Biplots generated using standardised Best Linear Unbiased Predictor values for three traits from the 30 half-sibling families of the Commando Persistent population. PC1 accounted for 53% and PC2 29% of the variation present. The different symbols indicate groups 1 to 4 generated from cluster analysis. Traits are indicated by the directional vectors: herbage growth (HGs), leaf width score (LWs), and plant habit score (PHs)	67
3.9; Biplots generated using standardised Best Linear Unbiased Predictor values for three traits from the 30 half-sibling families of the Samson Original population. PC1 accounted for 51% and PC2 33% of the variation present. The different symbols indicate groups 1 to 3 generated from cluster analysis. Traits are indicated by the directional vectors: herbage growth (HGs), leaf width score (LWs), and plant habit score (PHs)	69
3.10; Biplots generated using standardised Best Linear Unbiased Predictor values for five traits from the 30 half-sibling families of the Samson Persistent population. PC1 accounted for 50% and PC2 34% of the variation present. The different symbols indicate groups 1 to 5 generated from cluster analysis. Traits are indicated by the directional vectors: herbage growth (HGs), plant habit score (PHs), and aftermath heading (AMH).....	71
4.1; Plot of mean posterior probability (LnP(D)) values per cluster (K), based on 15 iterations per K from STRUCTURE analyses (Pritchard et al., 2000).....	102
4.2; STRUCTURE cluster assignment of 58 individuals from the populations Samson Original (1), Commando Original (2), Samson Persistent (3), and Commando Persistent (4). Box graphs show population structure for putative ancestral populations of a) K=2; b) K=3; and c) K=4	103

4.3; Principal co-ordinates analysis of all 11 markers across a total of 58 individuals.
Principal co-ordinates I and II accounted for 8.20% and 8.17% of the variation
respectively. Symbols indicate identity of individuals from each of the four populations
.....104

LIST OF TABLES

2.1; Analysis of variance (ANOVA) for each trait for the Original population. Mean, range, $LSD_{0.05}$ and F_{pr} are presented	29
2.2; Analysis of variance (ANOVA) for each trait for the Persistent population. Mean, range, $LSD_{0.05}$ and F_{pr} are presented.....	29
2.3; Cluster analysis produced three groups in the Original population. Number of genotypes and means for tiller number (TN), lamina length (LL), lamina width (LW), lamina thickness (LT), and sheath length (SL) are presented.....	31
2.4; Cluster analysis identified three groups in the Persistent population. Number of genotypes and means for tiller number (TN), lamina length (LL), lamina width (LW), lamina thickness (LT), and lamina sheath length (SL) are presented.....	33
2.5; Mean values of the Persistent Original populations for of each of the eight traits measured. Least significant difference ($L.S.D_{0.05}$) between means and significance level (F_{pr}) are presented.....	34
3.1; Means, ranges, estimated additive (σ^2_a), additive-by-season interaction (σ^2_{as}) and experimental error (σ^2_ϵ) variance components and their associated standard errors (\pm), coefficient of additive variation (CV_A), and half-sibling family mean narrow sense heritability (h^2_n) estimated for the traits herbage growth, leaf width, plant habit and dry weight, among the half-sibling families within each of the four populations.....	62
3.2; Means, ranges, estimated additive (σ^2_a) and experimental error (σ^2_ϵ) variance components and their associated standard errors (\pm), co-efficient of additive variation (CV_A), and half-sibling family mean narrow sense heritability (h^2_n) estimated for the traits reproductive tiller development, aftermath heading, and rust among the half-sibling families within each of the four populations.....	63

3.3; Cluster analysis produced four groups in the Commando Original population. Number of half-sibling families (HS) means for leaf width score (LWs), plant habit score (PHs) and herbage growth score (HGs) for each group are presented	65
3.4; Cluster analysis produced four groups in the Commando Persistent population. Number of half-sibling (HS) families for each group and means for leaf width score (LWs), plant habit score (PHs), and herbage growth score (HGs) for each group are presented	67
3.5; Cluster analysis produced three groups in the Samson Original population. Number of half-sibling (HS) families for each group and means for leaf width score (LWs), plant habit score (PHs), and herbage growth score (HGs) for each group are presented.....	69
3.6; Cluster analysis produced five groups in the Samson Persistent population. Number of half-sibling (HS) families for each group and means for plant habit score (PHs), and aftermath heading (AMH), and herbage growth score (HGs) for each group are presented	71
4.1; Characteristics of 11 <i>Lolium perenne</i> SSR markers selected for screening. Locus name, expected range, forward (F) and reverse (R) primer sequences, and fluorescent dye used for each marker is presented	91
4.2; Characteristics of 11 microsatellite loci for 20 samples of the Samson Original population. Base pair size range (BP range), percentage amplification (% amp.), number of alleles (Na), observed heterozygosity (H _O), and expected heterozygosity (H _E)	95
4.3; Characteristics of 11 microsatellite loci for six samples of the Samson Persistent population. Base pair size range (BP range), percentage amplification (% amp.), number of alleles (Na), observed heterozygosity (H _O), and expected heterozygosity (H _E).....	96

4.4; Characteristics of 11 microsatellite loci used on 20 samples of the Commando Original population. Base pair size range (BP range), percentage amplification (% amp.), number of alleles (Na), observed heterozygosity (H _O), and expected heterozygosity (H _E)	97
4.5; Characteristics of 11 microsatellite loci for 12 samples of the Commando Persistent population. Base pair size range (BP range), percentage amplification (% amp.), number of alleles (Na), observed heterozygosity (H _O), and expected heterozygosity (H _E)	98
4.6; For each of the 11 microsatellite loci, allele variants were compared between Original and Persistent populations within Samson and Commando. Presented are the total number of allele variants (Na _v) and shared proportion of alleles (SA _P) for each cultivar.....	99
4.7; Average characteristics of 11 microsatellite loci for each of the four populations. Number of samples (N), percentage amplification (% amp.), average number of alleles (Na), average observed heterozygosity (H _O), average expected heterozygosity (H _E), and estimates of Wright's fixation index within each population (F _{IS}) are presented.....	100
4.8; Pairwise F _{ST} values of the four populations of perennial ryegrass derived from 11 microsatellite markers.....	101
4.9; AMOVA results for the partitioning of microsatellite variation. Presented are the degrees of freedom (df), sum of squares (SS), mean of squares (MS), Estimated variance components (Est. Var.) and percentage of variation (Var. %)	104

LIST OF EQUATIONS

- 1.1; Phenotype (P) = genotype (G) + environment (E), and genotype (G) × environment (E): $P = G + E + (G \times E)$ 5
- 1.2; Genetic variation (σ_g^2) = additive genetic variation (σ_a^2) + dominance variation (σ_D^2)
 $\sigma_g^2 = \sigma_a^2 + \sigma_D^2$:.....7
- 1.3; Narrow-sense heritability (h_n^2) = additive genetic variation (σ_a^2)/ phenotypic variation (σ_P^2): $h_n^2 = \sigma_a^2 / \sigma_P^2$:7
- 1.4; Genetic gain per cycle (G_c) = narrow-sense heritability (h_n^2) x selection differential (S): $G_c = h_n^2 \times S$7
- 1.5; F_{IS} = expected heterozygosity (H_E) - Observed frequencies of heterozygotes (H_O) / expected heterozygosity (H_E) of individuals within the population assessed: $F_{IS} = \frac{H_E - H_O}{H_E}$ 18
- 1.6; F_{ST} = expected heterozygosity (H_{ET}) – observed frequencies of heterozygotes (H_{ES}) / expected heterozygosity (H_{ET}) of individuals among populations assessed: $F_{ST} = \frac{H_{ET} - H_{ES}}{H_{ET}}$ 19
- 3.1; (Y_{ijkl}) value of an attribute measured from half-sibling family i in row k and column l of replicate j and i . half-sibling families (f), replicates (b), rows (r) and columns (c); M is the overall mean; f_i is the random effect of HS family i , $N(0, \sigma_f^2)$; b_j is the random effect of replicate j $N(0, \sigma_b^2)$; r_{jk} is the random effect of row k within replicate j , $N(0, \sigma_r^2)$; c_{jl} is the random effect of column l within replicate j , $N(0, \sigma_c^2)$; ε_{ijkl} is the residual effect of HS family i in row k and column l of replicate j : $Y_{ijkl} = M + f_i + b_j + r_{jk} + c_{jl} + \varepsilon_{ijkl}$,57
- 3.2; Y_{ijklm} is the value of an attribute measured from half-sibling family i in row l and column m of replicate k nested in season j and i . Half-sibling families (f), replicates (b),

rows (r) and columns (c); M is the overall mean; f_i is the random effect of HS family i , $N(0, \sigma_f^2)$; s_j is the fixed effect of season j ; $(fs)_{ij}$ is the effect of the interaction between HS family i and season j , $N(0, \sigma_{fs}^2)$; b_{jk} is the random effect of replicate k within season j , $N(0, \sigma_b^2)$; r_{jkl} is the random effect of row l within replicate k within season j , $N(0, \sigma_r^2)$; c_{jklm} is the random effect of column m within replicate k within season j , $N(0, \sigma_c^2)$; ε_{ijklm} is the residual effect of HS family i in row l and column m of replicate k in season j : $Y_{ijklm} = M + f_i + s_j + (fs)_{ij} + b_{jk} + r_{jkl} + c_{jklm} + \varepsilon_{ijklm}, \dots\dots\dots 58$

3.3; Narrow-sense heritability (h_n^2) = additive genetic variation (σ_a^2) / additive genetic variation (σ_a^2) + additive by season interaction variance (σ_{as}^2) / number of seasons (n_s) + experimental error (σ_ε^2) / number of replications (n_r) x number of seasons (n_s): $h_n^2 = \frac{\sigma_a^2}{\sigma_a^2 + \frac{\sigma_{as}^2}{n_s} + \frac{\sigma_\varepsilon^2}{n_r n_s}} \dots\dots\dots 59$

3.4; The co-efficient of additive variation ($CV_A(\%)$) = (square root of phenotypic (σ_a^2) variance component / trait population mean (\bar{x})) x 100: $CV_A(\%) = \frac{\sqrt{\sigma_a^2}}{\bar{x}} \times 100 \dots\dots\dots 59$

