Index

Page numbers in *italics* refer to illustrations; those in **bold** refer to tables

1000 Genomes Project 10

a posteriori granddaughter design (APGD) 146
  concordance determination 148–149, 149

ABCG2 gene 63, 151

animal breeding programs 11
  “client” of 17–18
  crossbreeding programs 11
    advantages and disadvantages 14–15
    cumulative discounted returns 19
    traditional dairy cattle selection schemes 12–14, 13
    within-breed selection programs 11–12
  see also genetic gain

backcross (BC) design 24–25, 25
  with flanking markers 81–83, 82

Bayesian methods 27–28, 154
  genomic evaluation 116–117
  granddaughter design analysis 49
  QTL effects 76–79, 116–117
    Bayes A models 77–79, 116
    Bayes B models 77–79, 116, 117
    Bayes C model 116
    simulation results 79
    theory 76–77
    whole genome scans 76–79

BeadChips 9, 106, 139

BEAGLE algorithm 141
  accuracy and speed comparisons 142–143

best linear unbiased predictors (BLUP) 35, 128
  marker-assisted (MA-BLUP) 5

bias 126

biconjugate gradient stabilized (Bi-CGSTAB) algorithm 122

blood group markers 7

Bonferroni correction 69

bootstrap methods 86, 87

BovineHD BeadChip 9, 106
  breeding objective 17
  breeding programs see animal breeding programs
  breeding values (BV) 104
  broiler chickens see poultry

cattle see dairy cattle

chickens see poultry

chromosomal theory of inheritance 1

comparison-wise error rate (CWER) 69, 71–74, 73

competition among breeding companies 18

complete genome sequencing 9–10

confidence intervals (CI) 26–27
  computation of 84–85
  empirical estimation methods 86–87
  simulation studies 85–86

continuous variation 2

copy number variation (CNV) 9, 107

cow genomic evaluations 154
  see also dairy cattle

CRISPR/Cas9 technology 63

crossbreeding programs 11
  advantages and disadvantages 14–15
  selection index theory and 96
  see also animal breeding programs

dairy cattle 12
  GEBV method validation 127–129
    single-step methodology based on actual data 128–129
    two-step methodology based on actual data 127–128
    two-step methodology based on simulated data 127
  genome scans by granddaughter design 65–66
  genome-wide association studies 66, 146
  genomic evaluations 154, 156
    cow and female calves 154
    old versus young bulls 156
dairy cattle (cont’d)
marker assisted selection (MAS) breeding programs 5
parentage validation 135–136
QTL detection 2
QTN determination 63–64
traditional selection schemes 12–14
half-sib breeding program 12, 13, 14, 14
progeny test breeding program 12–14, 13, 14
daughter design 43, 44
haplotype determination 90, 92
interval mapping for 83–84
maximum likelihood estimation of QTL effects 43–45
daughter yield deviation (DYD) 39–40, 46, 83, 112–113
as the dependent variable 40, 113
all markers included as random effects 114–116
genomic analysis 112
daughter yields 112–113
see also daughter yield deviation (DYD)
DGAT1 gene 63, 150
DNA microsatellites 3, 8
parentage validation 135
economic contribution 18
economic evaluation
genetic gain 17–20
national economy versus competition among
breeders 17–18
profit horizon 18–19
electrophoresis 3
EMMAX algorithm 108, 146
epigenetic changes 96
estimated breeding value (EBV) 83, 113–114, 116
evaluation criteria 125–126
see also genomic estimated breeding values (GEBV)
estimated daughter contributions (EDC) 113–114
estimated genetic values see genomic estimated breeding values (GEBV)
expectation–maximization (EM) 26
false discovery rate (FDR) 71–74, 73
false positives, proportion of (PFP) 74–75
family-wise error rate (FWER) 69, 71–75, 73
fastPHASE algorithm 141
accuracy and speed comparisons 142–143
Findhap program 141–142
accuracy and speed comparisons 142–143
fixed variables 21
see also parameter estimation
full-sib design 4
Gauss–Seidel iteration 35–36, 39
genetic evaluation 11–12, 111
bias 126
deregressed genetic evaluation computation 113–114
parent average (PA) 125, 128, 129
reliability 12, 108, 112, 126
see also genomic evaluation; marker-assisted
genetic evaluation
genetic gain
dairy cattle breeding programs 14
economic evaluation 17–20
potential contribution of marker-assisted selection 96–97
predicted gains with genomic estimated breeding values 101–102
genetic group effect 39
genetic manipulation 156–157
genome scans
by granddaughter design 65–66
whole genome scans 69–71, 76–79
genome-wide association studies (GWAS) 145
dairy cattle 66, 146
economic traits 146
recessive lethal gene determination 150
genomic estimated breeding values (GEBV) 101
computation 115–116, 128, 129
based on subsets of SNPs 107–108
evaluation of 104
criteria 125–126
parent average comparisons 125
predicted genetic gains 101–102
reliabilities 116
single-step methodologies 122–123
selection effect 156
validation of methods 126–131
dairy cattle 127–129
multistep methodology based on actual data 127–128, 129–130
multistep methodology based on simulated data 127
plants 130
poultry 129–130
single-step methodology based on actual data 128–130
swine 130
see also estimated breeding value (EBV)
genomic evaluation 103–104, 111
Bayesian methodology 116–117
bias sources 104–105
cow and female calves 154
future directions 153
method improvements 154
imputation effect 143–144
long-term considerations 155–156
multistep methods 112–117
evaluation based on actual data 127–128, 129–130
evaluation based on simulated data 127
old versus young bulls 156
single-step methods 119–124
basic strategy 119–120
evaluation based on actual data 128–130
reliability estimation 122–123
INDEX

versus multistep models 111–112
with unequally weighted marker effects 123–124
validation 125, 126
see also genetic evaluation; genomic estimated
breeding values (GBEV); marker-assisted genetic
evaluation

genomic relationship matrix 108–109, 119
criteria for valid matrices 120
inverting 122
modified matrix when only a fraction of animals
are genotyped 120
solution 121

genotype
development of genotype frequency from expectations 107
incorrect scoring 107
number of genotyped animals 153–154
unequal viability 107
genotype building 156

genotyping
costs 8, 153
high-throughput 8–9
germ-line manipulation 157
Gibbs sampler 28–29
mixed model variance component estimation 55–58
government research institutions 18
granddaughter design 45–46, 45
a posteriori granddaughter design (APGD) 146
Bayesian estimation for segregating QTL 49
genome scans by 65–66
haplotype determination 90, 92
interval mapping for 83–84
prior QTL parameter distribution determination
46–48
Haldane mapping function 1–2, 83
half-sib design 4

dairy cattle breeding program 12, 13, 14
expected annual genetic gains 14
haplotypes 106
determination of 90, 92, 139–140
for imputation 139–140, 140

trends 155
heterosis 14–15, 96
high-density BeadChips 9, 106, 139
high-throughput genotyping 8–9
human genome sequence 9–10

identical by descent (IBD) 91
IBD probabilities 91, 92
imputation 139
accuracy and speed comparisons 142–143
algorithms 141–142
effect on genomic genetic evaluations 143–144
haplotype determination for 139–140, 140
humans versus farm animals 140

IMPUTE program 141
accuracy and speed comparisons 142
in vitro fertilization 157
incorrect parentage identification effects 133
individual animal model (IAM) 38–39
Infinium HD assay 9
Interbull 126
interval mapping see linkage mapping
jackknife method 87
joint linkage 92–93

Kronecker product 37
least absolute shrinkage and selection operator
(LASSO) 117
least squares estimation (LSE) 21–22
lethal recessive determination 150
likelihood ratio test 27, 85
linkage disequilibrium (LD) 4, 89
changes over time 155
estimation in animal populations 89–90
linkage disequilibrium (LD) mapping 81, 89
joint linkage and 92–93
multitrait and multiple QTL LD mapping 93
principles 90–92
linkage mapping 64–65, 81
backcross (BC) design 81–83, 82
daughter and granddaughter designs 83–84
see also linkage disequilibrium (LD) mapping
LOD drop-off method 87
loss function 27–28
low-density BeadChips 139

marker-assisted genetic evaluation 5
Bayesian weighting of marker effects 116–117
development from expected genotype frequencies 107
fixed versus random marker effects 105
individual markers versus haplotypes 106
marker redundancy 106
total versus select markers 106, 107–108
see also genetic evaluation; genomic evaluation
marker-assisted selection (MAS) 5, 59, 95
marker information 99–102
“animal model” genetic evaluations 100–101
relatives’ marker and phenotypic information
inclusion 99
segregating populations 100
maximum selection efficiency 99–100
reduction due to sampling variance 99–100
potential contribution of 96–97
reliability 5
sex-limited traits 98–99
simulation study results 101–102
versus phenotypic selection 97–98
matrix algebra 21
maximum likelihood estimation (MLE) 22–27
  confidence intervals (CI) 26–27
  computation of 84–85
  hypothesis testing and 26–27
  simulation studies 85–86
  likelihood function maximization 26
  mixed model equation solution 51–52
  variance components 52–54
  multiparameter 24–26
  QTL effects 43–46
  daughter design 43–45
  granddaughter design 45–46
  single parameter 22–24
  see also restricted maximum likelihood estimation (REML)
Mendelian theory of genetics 1–2
microsatellites see DNA microsatellites
mid-density BeadChips 9, 106, 139
missing genotypes 139
  see also imputation
missing heritability 61–62, 145
mixed linear model 34
mixed model equations 34–38
  important properties of solutions 36–37
  maximum likelihood solution 51–52
  multivariate analysis 37–38
  solving 35–36, 121–122
  variance component (VC) estimation 52
    Gibbs sampler (GS) 55–58
    Henderson’s Method III 52
    maximum likelihood estimation 52–54
    restricted maximum likelihood estimation 54–55
Morgans 1–2
multiple comparison problem 69
  multiple markers and whole genome scans 69–71
  QTL detection based on false discovery rate 71–75
  QTL detection by permutation tests 71
  multiple ovulation and embryo transplant (MOET) 18
  multiple QTL analysis 93
  multiltrait mapping 93
  multivariate mixed model analysis 37–38
national economy, contribution to 18
next-generation sequencing 9
nonparametric bootstrap method 86, 87
numerator relationship matrix 34, 38, 56, 92, 108
  pseudo relationship matrix 62
oocyte in vitro development and fertilization 157
parameter estimation 21
  Bayesian estimation 27–28
  Gibbs sampling 28–29
  least squares estimation (LSE) 21–22
  see also maximum likelihood estimation (MLE)
parametric bootstrap method 86
parent average (PA) 125, 128, 129
  genomic estimated breeding value comparisons 125
parentage identification and verification 134–135
  incorrect parentage identification effects 133
paternity validation 135–136
  prior to high-density SNP chips 135
  with SNP chips 135–136
  see also parentage identification and verification
degree reconstruction 137
permutation tests 71
phantom parents 39
phenotypic information on relatives 99
phenotypic selection 97–98
pigs, GEBV method validation 130
plants, GEBV method evaluation 130–131
polygenic variance 59–61
polymerase chain reaction (PCR) 3, 7
polymorphism
  information content (PIC) 134
  lack of 106
poly(TG) repeat sequences 3, 8
poultry
  broiler chicken breeding programs 14
  GEBV method validation 129–130
  preconditioned conjugate gradient 36
  predicted transmitting ability (PTA_mate) 40
  prediction error variance (PEV) 26–27, 36, 39, 112
  progeny test breeding programs, dairy cattle 12–14, 13
  expected annual genetic gains 14
  progress–surplus–bankruptcy cycle 18
  proportion of false positives (PFP) 74–75
  proportion of fully informative matings (PFIM) 134–135
  pseudo relationship matrix 62
quantitative trait loci (QTL) 2
  causative mutations 62–63
  see also quantitative trait nucleotides (QTN)
detection 2–4
  false discovery rate 71–74, 73
  permutation tests 71
  effective number of 61
  genotype concordance 148–149
  granddaughter design 45–46
  Bayesian estimation for QTL parameters 49
  prior distribution of QTL parameters 46–48
  linkage mapping of see linkage mapping
  maximum likelihood estimation of QTL effects 43–46
  daughter design 43–45
  missing heritability 61–62
  multiple QTL analysis 93
  biases with estimation 75–76
  new QTL appearance 155
  polygenic variance modeling 59–61
  segregation 2
  number of segregating QTLs, estimation 64–65
  see also Bayesian methods
quantitative trait nucleotides (QTN) 3, 62–63, 145
  conclusive evidence for 147
  concordance 148–149
determination of 146–147
arguments for and against 146–147
dairy cattle 63–64
phase determination for heterozygous sires 149
verification by statistical and biological methods 150–151
see also quantitative trait loci (QTL)

random variables 21
realized genomic reliabilities (RGR) 128
recessive lethal gene determination 150
relationship matrix see genomic relationship matrix;
numerator relationship matrix
relationship validation 136–137
parentage 134–136
relative selection efficiency (RSE) 98
maximum 99
reduction due to sampling variance 99–100
reliability

genetic evaluation 12, 108, 112, 126
genomic estimated breeding values (GEBV) 116
single-step methodologies 122–123
marker-assisted selection (MAS) 5
realized genomic reliabilities (RGR) 128
repetitive DNA 3
restricted maximum likelihood estimation (REML) 24
mixed model variance components 54–55
restriction fragment length polymorphisms (RFLPs) 3, 7
selection index 31
coefficients 32
principles 31–33
situations when not efficient 95–96
variance 33
sex-linked traits 98–99, 107
sib-pair design 4
simple sequence repeats (SSR) 3, 8
single nucleotide polymorphisms (SNPs) 4, 8–9
genome-wide association studies 66
genomic estimated breeding value (GEBV)
computation 107–108
random versus fixed marker effects 105
sire–dam heterosis 14
stutter bands 3, 8
support intervals 85
swine, GEBV method validation 130

variables 21
variance components see mixed model equations
velogenetics 157

whole genome scans 69–71
Bayesian estimation of QTL see Bayesian methods
within-breed selection programs 11–12

yield deviation (YD) 39–40
see also daughter yield deviation (DYD)