

# Index

## A

Alfalfa (*Medicago sativa*), 138  
Asexual propagation, 185  
Autohexaploid sweet potato, 186

## B

Bacterial artificial chromosome (BAC), 191  
Barley, 100, 101, 109, 112, 119, 125  
Bayesian estimation, 39–52  
  Bayes' theorem, 38, 39  
  determination, parameters, 38  
  Gibbs sampling (*see* Gibbs sampling)  
  linear mixed models (*see* Linear models)  
  parameters, normal distributions  
    condition, 39  
    conjugate priors, 43  
    mean with known variance, 40, 41  
    plausible values, 40  
    probability density functions, 39  
    variance with known mean, 41–43  
Bayesian genomic-enabled prediction models,  
  59–62, 65–67, 70, 71, 78–84, 86  
  accuracy, 56  
  BayesB method, 57  
  BayesC threshold model, 57  
  binary trait, 58  
  BLOR model (*see* Bayesian logistic ordinal  
    regression (BLOR) model)  
  BNBR (*see* Bayesian negative binomial  
    regression (BNBR))  
  data sets  
    Fusarium head blight data, 61  
    GLS and Septoria, 60  
  EBV accuracy, 57  
  full conditional distributions, 69, 70  
  genome prediction, 86  
  GBLUP, 58  
  GEBV accuracy, 57  
  Gibbs sampler (*see* Gibbs sampler)  
  GLMMs, 56  
  heritability and training population size, 57  
  MCMC algorithm, 59  
  methodological developments, 86  
  non-Gaussian phenotypes, 55  
  observed and ordinal categorical  
    phenotypes, 56  
  outcomes  
    FHB data set, 83, 84  
    GLS data set, 79–82  
    real data sets, 78–79  
    simulated data sets, 78  
  parameters, 55  
  SNP marker panels, 55  
  statistical, 61  
    applications, 59  
    BLOR (*see* Bayesian logistic ordinal  
      regression (BLOR) model)  
    BNBR, 62  
    Gibbs sampler, 67  
    linear predictor, 61  
    Pois, 62  
    polygenic effects, 66  
    regression coefficients, 65  
    threshold effects ( $\gamma_c$ ), 66  
    variance of polygenic effects, 66  
  TGBLUP, 58  
  threshold methods, 58  
  transformation, 56  
Bayesian LASSO methods, 194

- Bayesian logistic ordinal regression (BLOR)  
 model, 78–80, 88–92  
 CGGLUP, 59  
 cumulative logit model, 64  
 data augmentation approach, 59  
 distributions, 61, 64  
 exclusive and exhaustive categories, 63  
 fit the real data set, 79  
 full conditional distribution  
 liabilities, 88, 89  
 polygenic effects ( $b_h$ ), 91  
 regression coefficients ( $\beta$ ), 90  
 variance, location effects, 92  
 variance, polygenic effects, 91  
 GBLUP, 59  
 Gibbs sampler, 67  
 GLS data set, 79  
 interactions, 60  
 liabilities, 63, 65, 71  
 linear predictor, 61  
 mean and standard deviation values, 73  
 order statistics, 64  
 ordinal data analysis, 59  
 outcomes  
 simulated data sets, 78  
 GLS data set, 79–80  
 parameter vector, 64  
 parameters, 71  
 Pólya-Gamma distribution, 87  
 restriction, 87  
 thresholds, 63  
 validation samples, 80
- Bayesian methods, 194
- Bayesian negative binomial regression  
 (BNBR), 75, 76, 78, 80–84  
 assessing prediction accuracy, 76, 77  
 average values (mean) and standard  
 deviation (SD), 74  
 FHB data set, 85  
 fit the real data set, 79  
 full conditional distribution, 92–94  
 GLS data set, 80, 83  
 hyper-parameters, 68  
 implementation, 59, 72–75  
 interaction, 60  
 linear predictor, 68  
 measures, predictive performance  
 proper scoring rules, 75, 76  
 scoring rules, 75  
 outcomes  
 FHB data set, 83, 84  
 GLS data set, 80–82  
 simulated data sets, 78
- Pólya-Gamma distribution, 68  
 probability distribution, 68  
 scalar parameters, 81  
 standard deviations (SD), 72
- Bayesian LASSO, 188
- Best linear unbiased prediction  
 (BLUP), 164, 188
- Breeding cycle, 187, 192, 194
- Breeding strategies, clonal crops, 186–187
- Breeding strategy, cereals development, 117  
 genomics and phenomics, increased  
 precision, 116, 117  
 germplasm development, introgression  
 of alleles, 115, 116  
 implementation, GS, 113  
 MAS, 113  
 program  
 breeders, 117  
 design, 17  
 data and training populations, 117  
 introgress exotic alleles, 117  
 selection, alleles, 117  
 timing of GS application, 113–115, 117
- C**
- Chickpea (*Cicer arietinum*), 139, 140
- Clonal crops  
 asexual propagation, 185  
 breeding programs, 194–195  
 breeding strategies, 186–188  
 fruit and forest trees, 185  
 generations, 192–193  
 genetic analysis, 186  
 genetic architecture, 191–192  
 genetic effects, 189–190  
 genetic studies, polyploid crops, 186  
 GS, 187–189  
 heritability, 189–190  
 high natural heterozygosity, 185  
 LD, 190  
 polyploid, 185  
 potato breeders, 186  
 proof-of-concept, 193  
 single mother plant, 185
- Combining abilities, hybrid breeding, 152–153
- Conditions, 87
- Crop development, 2–6  
 agriculture land, 1  
 challenges, 2  
 complex traits, 6  
 conventional breeding, 2  
 food demand, 1

- food prices, 1
- food production systems, 1
- GS
  - breeding, 4, 5
  - clonal crops and tree species, 5–6
  - GEVVs, 2
  - methodologies and models, 3, 4
  - QTLs/genes, 2
  - quantitative traits, 2
  - training population, 2
- mapping populations, 2
- NGS technologies, 2
- poverty eradication, 1
- QTL, 2
- Cross-validated prediction accuracy of GS,
  - hybrid breeding
  - biparental populations, 156–158
  - concepts, 154
  - estimation
    - marker effects, 158
    - and validation, 155
  - factorial crosses, 158–160
  - limitations, 157
  - plant material, 155
  - predictabilities, 157
  - program, 155–158
  - same genotypes, 155
  - segregating inbred populations, 155
  - test and target environments, 160–161
- Cross validation, legumes, 134, 136, 138
  
- D**
- Disease-free stocks, 187
- Dominance, hybrid breeding, 150–151
- Dominance within GS model
  - accommodating epistasis, 167–168
  - Bayesian approaches, 165–167
  - benefits, 166
  - genotype  $\times$  environment interactions, 170
  - genotypes and relationship matrices, 164–164
  - marker effects
    - BayesA, 162–163
    - BayesB, 163
    - BayesC $\pi$ , 163–164
    - components, 161
    - linear unbiased prediction, 162
  - metabolomic prediction, 169–170
  - multiple-trait, 168–169
  - W-BLUP, 161
- Dry matter content (DMC), 193
  
- E**
- Eucalyptus, 212
  - breeding, 229
  - genotype, 218
  - GS, 219, 222
  - hybrid breeding, 227
  - populations, 215
  - RAD sequencing, 218
  - size, 214
  - SNP arrays, 217, 218
  - tropical, 237
  
- F**
- Forest tree breeding, 201–203
  - applications, 238
  - breeding cycle, 238
  - eucalyptus breeding, 238
  - genomic selection (*see* Genome-wide selection, forest tree breeding)
  - standard phenotypic selection, 237
  - top ranked seedlings, GEV, 238
  - training dataset, 238
  - trait dissection to genomic prediction, 199–201 (*see also* Genomic prediction in forest trees)
- Fresh root weight (FRW), 193
  
- G**
- Genetic architecture
  - traits and size of training population, 191–192
- Genetic diversity
  - inbreeding and maintenance, 235, 236
- Genetic effects
  - clonal crops, 189–190
- Genetic studies in polyploid crops, 186
- Genome-based selection, 187
- Genome-wide selection, forest tree breeding, 203, 204, 206, 215–221, 228–230, 233–235
  - changes, environment and model retraining, 241
  - complex traits, 203
  - cost-benefit analysis, 240, 241
  - cultivated plant species, 203
  - data analysis, 240
    - marker variables, 228
    - modeling nonadditive effects, 229
    - prediction of genome-estimated breeding, 228

- Genome-wide selection (*cont.*)
- ranking problem, 229, 230
  - risk, 228
  - RR-BLUP method, 228
  - dense marker data and QTL effects, 202
  - EBV, 201
  - effective population size, 213–215
  - environments and ages, 230–233, 239
  - experimental outcomes, 206–212
  - features and outcomes, 207–211
  - G-BLUP, 202
  - GEBV, 202
  - genetic relationship, training population
    - and selection candidates, 239
  - genetic values, 201
  - genomic prediction, 203
  - genotyping, 239
  - genotyping density and SNP
    - DNA marker, 215
    - fixed content, 217, 218
    - GbS approaches, 218–221
    - GEBV accuracy, 216
    - GEBV prediction accuracy, 216
    - LD pattern, 217
    - lower-density marker panels, 217
    - microsatellite markers, 216
    - molecular genetic techniques, 215
    - population size, 216
    - whole-genome sequence data, 221
  - inbreeding and maintenance of genetic
    - diversity, 235, 236
  - logistic issues, 240
  - marker effect, 202
  - MAS, 202
  - performance
    - breeding program, 235
    - G0, G1, and G2, 234
    - long-term gain, 233
    - Pinus pinaster*, 234
    - training and validation sets, 233
  - perspectives
    - accuracy, 204
    - BLUP, 206
    - EBV and true breeding value, 203
    - eucalypts and poplars, 204
    - GEBV/GEGV, 204
    - genetic improvement, 203
    - implementation, 204
    - measures, tree life, 203
    - phenotypes, 203
    - training population, 204
    - tree breeding program, 203, 204
    - prediction, generations, 241
    - stages, development and
      - implementation, 205
    - training GS models, 239
    - trait heritability and genetic architecture,
      - 227, 228
    - training population, 202
    - validation\population, 202
  - Genome-wide SNPs, 191
  - Genomic best linear unbiased prediction
    - (G-BLUP), 14, 15, 188, 194
  - Genomic estimated breeding values (GEBVs),
    - 2, 187, 190, 192
  - Genomic prediction, legumes, 134
  - Genomic prediction in forest trees, 230–233
    - alleles, traits, 199
    - challenges, tree breeders, 199
    - dissecting complex traits, 200
    - environments and ages
      - accuracy, 231, 232
      - crops, 233
      - GEBV, 231
      - interaction, 231
      - management zones, 231
      - marker effect, 233
      - outcomes, 232
      - periodical retraining, 233
      - ranking of individual trees, 230
      - same clones or families, 231
      - selection, same age, 232
      - tree breeding programs, 230
      - variation, 231
    - genomic selection, 201
    - growth and wood quality, 200
    - MAS, 200
    - modeling nonadditive effects, 229
    - QTL and association mapping, 200
    - ranking problem, 229, 230
    - recurrent cycles of selection, mating and
      - testing, 199
    - tree breeding programs, 199
  - Genomic prediction models
    - across environment vs. within environment
      - replication, 18, 19
    - breeding programs, 16
    - MAS, 16
    - preliminary and advanced breeding tests,
      - 17, 18
    - QTL mapping, 16
    - replication and plot allocation, 17
    - replication, phenotypic variance, 16
  - Genomic selection (GS), 2, 185

candidate population, 2  
 clonally propagated crops (*see* Clonal crops)  
 crop (*see* Crop development)  
 GEBVs, 2, 3, 187  
 genome-wide marker data, 2  
 methodologies and models, 3, 4  
 Genomics-assisted breeding (GAB), 132, 133, 136  
 Genotype-by-environment (G x E) interactions, 186, 192  
 Gibbs sampler  
   BNBR  
     parameters, 70  
     simulation, 71  
   BLOR, 67  
 Gibbs sampling  
   Box-Muller method, 50  
   conditional probability distributions, 49  
   parameters estimation, 50–52  
   procedure, 49  
   values, 49  
 Gray leaf spot (GLS), 60  
 Groundnut (*Arachis hypogaea*), 140–141

**H**

Hardy–Weinberg equilibrium, 190, 192  
 Heritability, clonal crops, 189–190  
 Heterosis, hybrid breeding, 149, 150, 152–154  
 High-density genome-wide molecular markers, 187  
 Hybrid breeding  
   combining abilities, 152–153  
   concept of dominance, 151–152  
   GS  
     challenges, 178  
     cross-validation (*see* Cross-validated prediction accuracy, GS)  
     dominance (*see* Dominance within GS model)  
     implementation (*see* Implementation of GS in hybrid breeding)  
     skills and knowledge, plant breeders, 178  
   heterosis, 149–150  
   heterotic group and patterns, 153–154  
   line breeding, 151

**I**

Implementation of GS in hybrid breeding  
   biparental population, 173, 174, 176, 178  
   budget, 171, 173, 174  
   decision variables, 174–175

objective function, 171–173  
 parameters, 175–176  
 plant breeding programs, 171, 173, 177  
 predictability, 171, 177  
 recalibration, 176–178  
 International Institute of Tropical Agriculture (IITA), 191  
 International Potato Center (CIP), 191

**L**

Legume crops  
   Alfalfa, 138  
   breeding, 132  
   chickpea, 139, 140  
   cultivation, 131  
   demand, 132  
   GAB approaches, 133  
   GBS approach, 133  
   generation, genomic resources, 141  
   genomic-assisted breeding, 141  
   genomic prediction, 134  
   genomic resources, 132  
   genomic selection studies, 137  
   groundnut, 140  
   GS, 134–136, 141  
   introgression lines, 133  
   large scale genome-wide genetic markers, 132  
   MABC, 133, 134  
   MARS, 133, 134  
   NGS technologies, 132  
   pea, 139  
   production, 131  
   productivity trends, 132  
   QTLs, 134  
   single nucleotide polymorphisms (SNPs), 132  
   soybean (*Glycine max*), 136–138  
   UN World Food Program, 131  
 Linear models, 44, 46–49  
   Bayesian estimation  
     conditional probability distribution  $\beta$ , 46, 47  
     conditional probability distribution, 49  
     conditional probability distribution,  $u_j$ , 47, 48  
     parameters, 44  
     prior probability distributions, 44  
     quantitative traits, 25–27  
 Linkage disequilibrium (LD), 189  
   markers and quantitative trait loci, 190  
   QTL mapping, 10  
 Linear models, 43, 45, 46  
   Bayesian estimation

- Linear models (*cont.*)  
 joint posterior probability distributions,  
 45–46  
 MCMC methods, 43  
 prior probability distributions, 45
- M**  
 Marker-assisted backcrossing (MABC), 133  
 Marker-assisted selection (MAS), 187  
 crop development, 4  
 oats, 109  
 wheat, 101  
 MARS. *See* Marker-assisted recurrent selection  
 (MARS)
- N**  
 Next-generation sequencing (NGS)  
 technologies, 2
- O**  
 Oats, 100, 109
- P**  
 Pea (*Pisum sativum*), 139  
 Phenotypic data, 87  
 Phenotypic selection (PS), 188, 193  
*Picea*, 214, 217, 218  
 Pinus  
 generations, 212  
 genotyping, 218  
 SNP arrays, 217  
 Population size, legumes, 135, 136, 138–140  
 Potato breeders, 186  
 Prediction accuracy, 5, 6, 188, 189  
 Public sector, 193
- Q**  
 Qualitative and quantitative trait loci  
 (QTL), 187  
 Quantitative Trait Loci  
 LD, 190–191  
 Quantitative traits, 23, 24, 28, 29, 31–52  
 Bayesian estimation (*see* Bayesian  
 estimation)  
 breeding values and heritability  
 additive genotypic value, 32  
 broad-sense, 32  
 deviations, 31  
 genetic factors, 31  
 genetic markers, 32, 33  
 GS, 32  
 genotypic value, 32  
 LD, 33  
 narrow-sense, 32  
 traditional genomic selection, 32  
 values, 31  
 decomposition of variance and genetic  
 markers, 30, 31  
 genetic factors, 23  
 genetic markers, 52  
 GS  
 dataset, 24  
 genetic markers, 24  
 process, 23  
 strategies, 23  
 LD, 53  
 least-squares estimation  
 genetic markers, 34  
 number of parameters, 37, 38  
 one-marker model, 34, 35  
 parameters, 33  
 two-marker model without  
 interactions, 35–37  
 linear models, 25–27  
 one-marker model, 27, 28  
 prediction, 52  
 two-marker model  
 with interactions, 29  
 without interactions, 28–29
- R**  
 Random regression-best linear unbiased  
 prediction model (RR-BLUP), 193,  
 194  
 Reproducing kernel Hilbert space (RKHS),  
 188, 194  
 Ridge-regression best linear unbiased  
 prediction (RR-BLUP), 14, 17, 188  
 Rye, 100, 101
- S**  
 Shoot weight (SW), 193  
 Single mother plant, 185  
 Small grains, 101, 109–125  
 breeding methods (*see* Breeding strategy,  
 cereals development)  
 breeding programs, 125

- GS
- cost benefit analysis, 124, 125
  - pure line breeding program, 114
  - studies, 102–105, 107
- GS prediction
- accuracies, 110
  - effect, number of markers and individuals, 112
  - model, 111
  - research, 110
  - training and validation population, 111, 112
- GS research
- DArT, 109
  - GBS, 109
  - oat breeding program, 109
  - oats, 109
  - wheat, 101
- GxE
- applications, GS models, 118–119
  - environmental covariates and crop models, 123, 124
  - modeling marker replication and interaction, 119–121
  - phenotypic variation, 118
  - TPE, 118
  - treating environments, multiple traits, 121, 122
  - marker effects, 120
  - optimization, training population, 121
- Small grains
- biparental cross of maize, 100
  - breeding cycle, 100
  - breeding program, 100
  - cereals, 100
  - estimated breeding values, 100
  - genotype-by-environment interaction (GxE), 101
  - pseudo-cereals, 100
  - RR-BLUP, 101
- Soybean (*Glycine max*), 136–138
- Starch yield (SY), 193
- T**
- Target population of environments (TPE), 118
- Time consuming and resource consuming, 186
- Training population, 8–15
- breeding programs, 7
  - design
    - composition, 8, 9
    - co-segregation of alleles, 15
    - diverse germplasm collection, 12–14
    - family-specific, 9
  - G-BLUP perspective, 14
  - genetic relationships, 15
  - genomic prediction, 9, 10, 14
  - genomic selection, 8, 9
  - genotyping, 10
  - LD source of information, 15
  - maize breeding, 9
  - marker-based optimization, 10
  - marker-QTL LD, 15
  - optimization, 9
  - phenotypic data, 8
  - QTL mapping, 10
  - researchers, 14
  - RR-BLUP perspective, 14
  - sizes and marker numbers, 10
  - unrelated genotypes, 11, 12
- development, 8
- genetic value, 7
- genotyping costs, 8
- hypothetical and empirical implementation, GS, 8
- literature, 19
- molecular markers, 7
- optimization, 121
- phenotypic testing procedures, 7
- phenotyping, 7, 19
- resource allocation, phenotyping, 16–19
- tree breeding
- clonal replication, 223
  - eucalyptus, 222
  - GS in hybrid breeding, 226, 227
  - NIRS, 222
  - phenotyping, 222
  - QTL distribution, 222
  - selection accuracy, 222
  - and selection candidates, 223–226
  - SNP effects, 221
- Trait heritability and genetic architecture, 227, 228
- Triticale, 100, 116
- V**
- Vegetative propagules, 187
- W**
- Wheat
- bread, 101
  - breeding programs, 101, 112

Wheat (*cont.*)

buckwheat, 100

durum wheat, 101, 120

environmental covariates, 122

GBS datasets, 109

GS, 100, 109

investigation, 111

lines, 123

low-intensity winter, 124

MET, 122

phenotypes, 120

quality, 125

SiriusQuality, 123