References

[1] O. R. N. Laboratory, "Biological and environmental research information system." http://genomicscience.energy.gov. [Online; accessed 12-December-2014].

[2] B. Alberts, J. Alexander, L. Julian, R. Martin, R. Keith, and W. Peter, *Molecular Biology of the Cell*, ch. Manipulating Proteins, DNA, and RNA. New York: Garland Science, fourth ed., 2002.

[3] Wikipedia, "Amino acid — wikipedia, the free encyclopedia," 2014. [Online; accessed 25-December-2014].

[4] "Access excellence @ the national health museum." http://www.accessexcellence.org/RC/VL/GG/central.php, 1994-2009. [Online; accessed 20-December-2014].

[5] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," *AI magazine*, vol. 17, no. 3, p. 37, 1996.

[6] W. M. Jr and M. J. Zaki, *Data Mining and Analysis*. Cambridge University Press, 2014.

[7] E. Mardis, "Next-generation dna sequencing methods," *Annual Review of Genomics and Human Genetics*, vol. 9, no. 11, pp. 387–402, 2008.

[8] M. Smith, "Nucleotide sequence of bacteriophage phi x174 dna," *Nature*, vol. 265, no. 5596, pp. 687–95, 1977.

[9] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping multidimensional data*, pp. 25–71, Springer, 2006.

[10] J. P. Demuth, T. De Bie, J. E. Stajich, N. Cristianini, and M. W. Hahn, "The evolution of mammalian gene families," *PloS one*, vol. 1, no. 1, p. e85, 2006.

[11] B. Zhao, V. Duan, and S. S.-T. Yau, "A novel clustering method via nucleotide-based fourier power spectrum analysis," *Journal of theoretical biology*, vol. 279, no. 1, pp. 83–89, 2011.

[12] M. Deng, C. Yu, Q. Liang, R. L. He, and S. S.-T. Yau, "A novel method of characterizing genetic sequences: genome space with biological distance and applications," *PloS one*, vol. 6, no. 3, p. e17293, 2011.

[13] H. Pearson, "Genetics: what is a gene?," *Nature*, vol. 441, no. 7092, pp. 398–401, 2006.

[14] J. Watson and F. Crick, "A structure for deoxyribose nucleic acid," *Nature*, vol. 421, no. 6921, pp. 397–398, 1953.

[15] F. Crick *et al.*, "Central dogma of molecular biology," *Nature*, vol. 227, no. 5258, pp. 561–563, 1970.

[16] I. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

[17] C. Bishop, *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., 1995.

[18] T. Dietterich, "Machine learning," in *Nature Encyclopedia of Cognitive Science*, Macmillan. London, 2003.

[19] T. HASTIE, R. TIBSHIRANI, and J. FRIEDMAN, *Elements of statistical learning: Data mining, interference and prediction*. Springer-Verlag, 2001.

[20] T. Graepel, "Statistical physics of clustering algorithms," Master's thesis, Berlin, Germany, 1998.

[21] B. Everitt, S. Landau., and M. Leese, *Cluster Analysis*. Wiley; 5 edition, 2011.

[22] G. Gan, C. Ma, and J. Wu, *Data clustering: theory, algorithms, and applications*, vol. 20. Siam, 2007.

[23] J. Han and M. Kamber, *Data Mining, Southeast Asia Edition: Concepts and Techniques*. Morgan kaufmann, 2006.

[24] J. A. Hartigan, *Clustering algorithms*. John Wiley and Sons, New York, NY, 1975.

[25] G. Erban and G. Moldovan, "A comparison of clustering techniques in aspect mining.," *Informatica*, vol. L 1, no. 1, pp. 69–78, 2006.

[26] P. H. Sneath, "The application of computers to taxonomy," *Journal of general microbiology*, vol. 17, no. 1, pp. 201–226, 1957.

[27] S. T., "A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyzes of the vegetation on danish commons," *Biologiske Skrifter*, vol. 5, p. 1–34, 1948.

[28] H.-P. Kriegel, P. Kröger, J. Sander, and A. Zimek, "Density-based clustering," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 231–240, 2011.

[29] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *Kdd*, vol. 96, pp. 226–231, 1996.

[30] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "Optics: Ordering points to identify the clustering structure," in *ACM Sigmod Record*, vol. 28, pp. 49–60, ACM, 1999.

[31] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[32] B. Everitt, S. Landau, M. Leese, and D. Stahl, *cluster analysis*. wiley publication, fifth ed., 2011.

[33] K.-L. Du, "Clustering: A neural network approach," *Neural Networks*, vol. 23, no. 1, pp. 89–107, 2010.

[34] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[35] T. Kohonen, "Self-organization and associative memory," *Self-Organization and Associative Memory. 100 figs. XV, 312 pages. Springer-Verlag Berlin Heidelberg New York. Also Springer Series in Information Sciences, volume 8*, vol. 1, 1988.

[36] G. McLachlan and K. Basford, *Mixture models : inference and applications to clustering*, vol. 84 of *STATISTICS: Textbooks and Monographs*. New York, N.Y. : M. Dekker, 1988.

[37] J. Banfield and A. Raftery, "Model-based gaussian and non-gaussian clustering," *Biometrics*, pp. 803–821, 1993.

[38] G. McLachlan and D. Peel, *Finite mixture models*. John Wiley & Sons, 2004.

[39] A. Fraley, C .and Raftery, "Model-based clustering, discriminant analysis, and density estimation," *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 611–631, 2002.

[40] G. Schwarz *et al.*, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[41] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.

[42] M. Medvedovic and S. Sivaganesan, "Bayesian infinite mixture model based clustering of gene expression profiles," *Bioinformatics*, vol. 18, no. 9, pp. 1194–1206, 2002.

[43] S. Altschul, W. Gish, W. Miller, E. Myers, and D. Lipman, "Basic local alignment search tool," *Journal of molecular biology*, vol. 215, no. 3, pp. 403–410, 1990.

[44] W. Pearson and D. Lipman, "Improved tools for biological sequence comparison," *Proceedings of the National Academy of Sciences*, vol. 85, no. 8, pp. 2444–2448, 1988.

[45] R. Xu and D. Wunsch, "Survey of clustering algorithms," *Neural Networks, IEEE Transactions on*, vol. 16, no. 3, pp. 645–678, 2005.

[46] I. Dondoshansky and Y. Wolf, "Blastclust (ncbi software development toolkit)," *NCBI, Bethesda, Md*, 2002.

[47] A. Enright and C. Ouzounis, "Generage: a robust algorithm for sequence clustering and domain detection," *Bioinformatics*, vol. 16, no. 5, pp. 451–457, 2000.

[48] I. Uchiyama, "Hierarchical clustering algorithm for comprehensive orthologous-domain classification in multiple genomes," *Nucleic acids research*, vol. 34, no. 2, pp. 647–658, 2006.

[49] W. Li and A. Godzik, "Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences," *Bioinformatics*, vol. 22, no. 13, pp. 1658–1659, 2006.

[50] V. Guralnik and G. Karypis, "A scalable algorithm for clustering sequential data," in *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pp. 179–186, IEEE, 2001.

[51] X. Wan, S. Bridges, J. Boyle, and A. Boyle, "Interactive clustering for exploration of genomic data," *SmartEng Design*, vol. 12, pp. 753–758, 2002.

[52] D. Wei, Q. Jiang, Y. Wei, and S. Wang, "A novel hierarchical clustering algorithm for gene sequences," *BMC bioinformatics*, vol. 13, no. 1, p. 174, 2012.

[53] K. Amano, H. Nakamura, and H. Ichikawa, "Self-organizing clustering: a novel non-hierarchical method for clustering large amount of dna sequences," *GENOME INFORMATICS SERIES*, pp. 575–576, 2003.

[54] G. Elhadi, R. Parouk, and A. Issa, "Protein sequence for clustering dna based on artificial neural networks," *International Journal of Computer Science Issues*, vol. 9, no. 1, pp. 161–167, 2012.

[55] B. Umamageswari, B. Karthikeyan, and T. Nalini, "A comparative analysis of feature selection methods for clustering dna sequences," *International Journal of Computer Science and Security (IJCSS)*, vol. 6, no. 2, p. 120, 2012.

[56] S. Sen, S. Narasimhan, and A. Konar, "Biological data mining for genomic clustering using unsupervised neural learning.," *Engineering Letters*, vol. 14, no. 2, pp. 61–71, 2007.

[57] A. Krogh, M. Brown, I. Mian, K. Sjolander, and D. Haussler, "Hidden markov models in computational biology: Applications to protein modeling," *Journal of molecular biology*, vol. 235, no. 5, pp. 1501–1531, 1994.

[58] Z. Qin, "Clustering microarray gene expression data using weighted chinese restaurant process," *Bioinformatics*, vol. 22, no. 16, pp. 1988–1997, 2006.

[59] S. Jensen, L. Shen, and J. Liu, "Combining phylogenetic motif discovery and motif clustering to predict co-regulated genes," *Bioinformatics*, vol. 21, no. 20, pp. 3832–3839, 2005.

[60] Z. Qin, L. McCue, W. Thompson, L. Mayerhofer, C. Lawrence, and J. Liu, "Identification of co-regulated genes through bayesian clustering of predicted regulatory binding sites," *Nature biotechnology*, vol. 21, no. 4, pp. 435–439, 2003.

[61] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. CRC press, 2013.

[62] A. Y. Lo, "Weighted chinese restaurant processes," *Cosmos*, vol. 1, no. 01, pp. 107–111, 2005.

[63] J. S. Liu, *Monte Carlo strategies in scientific computing*. springer, 2008.

[64] J. S. Liu, W. H. Wong, and A. Kong, "Covariance structure of the gibbs sampler with applications to the comparisons of estimators and augmentation schemes," *Biometrika*, vol. 81, no. 1, pp. 27–40, 1994.

[65] R. Chen and J. S. Liu, "Predictive updating methods with application to bayesian classification," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 397–415, 1996.

[66] D. Wei, Q. Jiang, Y. Wei, and S. Wang, "A novel hierarchical clustering algorithm for gene sequences," *BMC bioinformatics*, vol. 13, no. 1, p. 174, 2012.

[67] S. Vinga and J. Almeida, "Alignment-free sequence comparison—a review," *Bioinformatics*, vol. 19, no. 4, pp. 513–523, 2003.

[68] B. Haubold, F. A. Reed, and P. Pfaffelhuber, "Alignment-free estimation of nucleotide diversity," *Bioinformatics*, vol. 27, no. 4, pp. 449–455, 2011.

[69] Z. Liu, J. Meng, and X. Sun, "A novel feature-based method for whole genome phylogenetic analysis without alignment: application to hev genotyping and subtyping," *Biochemical and biophysical research communications*, vol. 368, no. 2, pp. 223–230, 2008.

[70] M. Domazet-Lošo and B. Haubold, "Efficient estimation of pairwise distances between genomes," *Bioinformatics*, vol. 25, no. 24, pp. 3221–3227, 2009.

[71] M. Domazet-Lošo and B. Haubold, "Alignment-free detection of local similarity among viral and bacterial genomes," *Bioinformatics*, vol. 27, no. 11, pp. 1466–1472, 2011.

[72] A. Kelil, S. Wang, R. Brzezinski, and A. Fleury, "Cluss: clustering of protein sequences based on a new similarity measure," *BMC bioinformatics*, vol. 8, no. 1, p. 286, 2007.

[73] G. Reinert, D. Chew, F. Sun, and M. S. Waterman, "Alignment-free sequence comparison (i): statistics and power," *Journal of Computational Biology*, vol. 16, no. 12, pp. 1615–1634, 2009.

[74] Q. Dai, X. Liu, Y. Yao, and F. Zhao, "Numerical characteristics of word frequencies and their application to dissimilarity measure for sequence comparison," *Journal of theoretical biology*, vol. 276, no. 1, pp. 174–180, 2011.

[75] G. Lu, S. Zhang, and X. Fang, "An improved string composition method for sequence comparison," *BMC bioinformatics*, vol. 9, no. Suppl 6, p. S15, 2008.

[76] T. Aita, Y. Husimi, and K. Nishigaki, "A mathematical consideration of the word-composition vector method in comparison of biological sequences," *BioSystems*, vol. 106, no. 2, pp. 67–75, 2011.

[77] L. Liu, Y.-k. Ho, and S. Yau, "Clustering dna sequences by feature vectors," *Molecular phylogenetics and evolution*, vol. 41, no. 1, pp. 64–69, 2006.

[78] P. H. Sneath, R. R. Sokal, *et al.*, *Numerical taxonomy. The principles and practice of numerical classification.* 1973.

[79] R. Ikeda and D. A. Vaughan, "The distribution of resistance genes to the brown planthopper in rice germplasm," *RGN 8*, pp. 1–3, 1991.

[80] G. Rajkumar, J. Weerasena, K. Fernando, A. Liyanage, and R. Silva, "Genetic differentiation among sri lankan traditional rice (oryza sativa) varieties and wild rice species by aflp markers," *Nordic Journal of Botany*, vol. 29, no. 2, pp. 238–243, 2011.

[81] Y. Kawahara, M. de la Bastide, J. P. Hamilton, H. Kanamori, W. R. McCombie, S. Ouyang, D. C. Schwartz, T. Tanaka, J. Wu, S. Zhou, *et al.*, "Improvement of the oryza sativa nipponbare reference genome using next generation sequence and optical map data," *Rice*, vol. 6, no. 1, p. 4, 2013.

[82] M. Stephens, "Dealing with label switching in mixture models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 62, no. 4, pp. 795–809, 2000.

Appendix A: Input Data sets

Experiment 1

Our original data used in the first experiment contains binary data relating to 784 ALFP profiles for 53 Sri Lankan rice varieties. Each row represents a variety. First column gives the name of the variety and the rest of the columns represents the binary value obtained for individual profile. A sample input line is shown below for a single variety. The pairwise distances calculated based on these binary data is shown in table 6.3.

Siviruwee 1000001110101001110010101101 1110101100011000100010000100000001000
1000000010010000001001000000100000000 0000000000000000000000000000000010000
0000000000000010000100011000000000000 0010000011101101100101001110111101000
1110111010110100010110000011101101110 0000010000000000000000000000001001000
0000000000000000000000000000000000000 0000000000000000000000000000000000000
0000000000001011011101101101001111000 0001000000000000000000000000110100000
0001000001000100000000001000100000000 0000000000000001000010001001000000000
1010000000000000000000000000000000000 0000000000000000000000000100000000000
0100000000100001000000000000000000000 0000000000000010000001010000100010001
1101000000001000000000100000001000000 0010001000001000100001001011100111000
0110100010100000100100000000001101000 10010110000010001

Table 1: Jaccard's pairwise distances for the 53 rice varieties

pairDist

	TN1	Mol-samba	yaka-wee	Hat-wee	Kiriyal	Heenati	Siviru-wee	Dahanala	Pacha-peru	Sulai	Podiwi	vellai	Vella-Illa	Odda-valan	Hathiel	Kurulu	Murun-137	Dikwee	Kahata-wee	Hatapan	Niyan-wee	Peilliane	Dewar-add	Murun-gaka	Kotta-mall	Red leaf	Weda-hee	Mudu-kirie	Sudu-heen	O. rhizom	O. rufipo	Godaal-wee	Gona-baru	Kalu-heen	Balama-wee	Sudu-heena	O. granata	O. nivara	kombili	0.eii	Kalu-kura	Kohu-mawi	Suwa-dasa	Madael	Goda-hee	Honda-rawa	Heen-dikwe	Rathuh	chem-bava	Murun-ga	Chemba	hygroryza	Pokkali
TN1	1	0.172	0.279	0.191	0.208	0.231	0.167	0.162	0.211	0.184	0.108	0.171	0.178	0.216	0.192	0.18	0.153	0.165	0.157	0.121	0.218	0.196	0.169	0.197	0.143	0.154	0.155	0.18	0.197	0.209	0.195	0.146	0.126	0.137	0.164	0.168	0.148	0.131	0.128	0.15	0.122	0.15	0.148	0.149	0.135	0.152	0.139	0.11	0.176	0.147	0.162	0.101	0.203
Mol-samba	0.172	1	0.343	0.277	0.295	0.333	0.332	0.217	0.19	0.198	0.18	0.16	0.241	0.223	0.186	0.245	0.242	0.194	0.237	0.198	0.183	0.233	0.189	0.231	0.208	0.207	0.177	0.196	0.203	0.176	0.156	0.185	0.157	0.162	0.168	0.14	0.194	0.147	0.169	0.182	0.164	0.164	0.178	0.113	0.162	0.142	0.114	0.14	0.161	0.162	0.126	0.079	0.2
yaka-wee	0.279	0.343	1	0.287	0.269	0.342	0.296	0.232	0.236	0.197	0.193	0.217	0.226	0.227	0.185	0.204	0.234	0.242	0.191	0.192	0.237	0.237	0.183	0.241	0.245	0.263	0.161	0.195	0.186	0.197	0.178	0.161	0.144	0.147	0.175	0.127	0.147	0.17	0.161	0.162	0.127	0.127	0.192	0.099	0.138	0.135	0.116	0.141	0.137	0.162	0.159	0.083	0.237
Hat-wee	0.191	0.277	0.287	1	0.282	0.288	0.253	0.237	0.212	0.146	0.179	0.257	0.25	0.266	0.212	0.209	0.232	0.167	0.165	0.191	0.17	0.249	0.211	0.216	0.171	0.167	0.145	0.216	0.17	0.189	0.177	0.146	0.15	0.153	0.19	0.126	0.184	0.191	0.183	0.213	0.148	0.163	0.169	0.083	0.114	0.118	0.084	0.155	0.16	0.153	0.127	0.069	0.198
Kiriyal	0.208	0.295	0.269	0.282	1	0.245	0.26	0.218	0.196	0.176	0.186	0.199	0.236	0.19	0.156	0.179	0.204	0.161	0.191	0.186	0.169	0.169	0.188	0.22	0.189	0.195	0.158	0.196	0.186	0.197	0.21	0.184	0.211	0.168	0.182	0.161	0.176	0.184	0.175	0.225	0.17	0.17	0.219	0.133	0.168	0.175	0.144	0.171	0.157	0.205	0.183	0.079	0.208
Heenati	0.231	0.333	0.342	0.288	0.245	1	0.273	0.214	0.211	0.19	0.174	0.15	0.218	0.281	0.223	0.251	0.231	0.261	0.26	0.234	0.228	0.246	0.172	0.244	0.223	0.237	0.163	0.188	0.238	0.183	0.171	0.142	0.159	0.164	0.135	0.13	0.164	0.124	0.156	0.157	0.159	0.18	0.171	0.091	0.141	0.131	0.092	0.116	0.126	0.149	0.153	0.089	0.226
Siviruwee	0.167	0.332	0.296	0.253	0.26	0.273	1	0.253	0.215	0.237	0.191	0.152	0.205	0.184	0.137	0.202	0.24	0.151	0.167	0.167	0.172	0.215	0.162	0.247	0.216	0.234	0.157	0.207	0.179	0.207	0.135	0.151	0.186	0.172	0.125	0.179	0.171	0.171	0.157	0.187	0.153	0.16	0.184	0.084	0.124	0.114	0.078	0.133	0.143	0.138	0.135	0.059	0.197
Dahanala	0.162	0.217	0.232	0.237	0.218	0.214	0.253	1	0.345	0.289	0.251	0.238	0.274	0.254	0.18	0.266	0.206	0.16	0.157	0.176	0.172	0.249	0.162	0.23	0.194	0.155	0.185	0.213	0.185	0.233	0.178	0.164	0.192	0.186	0.163	0.146	0.206	0.225	0.184	0.249	0.173	0.153	0.205	0.11	0.158	0.141	0.104	0.153	0.136	0.158	0.169	0.077	0.223
Pachaperu	0.211	0.19	0.236	0.212	0.196	0.211	0.215	0.345	1	0.264	0.173	0.21	0.25	0.272	0.224	0.274	0.22	0.172	0.145	0.124	0.187	0.217	0.171	0.199	0.143	0.147	0.184	0.224	0.177	0.156	0.184	0.138	0.178	0.131	0.13	0.133	0.176	0.23	0.168	0.22	0.156	0.171	0.184	0.106	0.114	0.118	0.107	0.171	0.121	0.138	0.158	0.096	0.213
Sulai	0.184	0.198	0.197	0.146	0.176	0.19	0.237	0.289	0.264	1	0.157	0.177	0.19	0.183	0.119	0.242	0.191	0.187	0.162	0.211	0.174	0.208	0.175	0.169	0.149	0.159	0.193	0.162	0.203	0.181	0.175	0.137	0.203	0.195	0.145	0.225	0.147	0.202	0.146	0.143	0.159	0.177	0.183	0.17	0.165	0.194	0.15	0.198	0.154	0.216	0.161	0.112	0.256
Podiwi	0.108	0.18	0.193	0.179	0.186	0.174	0.191	0.251	0.173	0.157	1	0.194	0.219	0.215	0.193	0.282	0.161	0.134	0.142	0.141	0.145	0.164	0.101	0.131	0.171	0.15	0.15	0.243	0.162	0.192	0.154	0.15	0.163	0.198	0.213	0.144	0.178	0.252	0.195	0.229	0.181	0.181	0.223	0.123	0.131	0.118	0.114	0.144	0.13	0.151	0.119	0.067	0.194
vellai	0.171	0.16	0.217	0.257	0.199	0.15	0.152	0.238	0.21	0.177	0.194	1	0.26	0.305	0.257	0.221	0.219	0.155	0.18	0.163	0.196	0.216	0.189	0.159	0.139	0.144	0.167	0.206	0.161	0.153	0.186	0.149	0.187	0.177	0.183	0.16	0.177	0.185	0.185	0.189	0.198	0.162	0.203	0.142	0.178	0.186	0.133	0.171	0.167	0.177	0.164	0.116	0.201
VellaIlla	0.178	0.241	0.226	0.25	0.236	0.218	0.205	0.274	0.25	0.19	0.219	0.26	1	0.34	0.244	0.314	0.228	0.193	0.213	0.192	0.188	0.2	0.182	0.194	0.148	0.173	0.176	0.21	0.201	0.169	0.155	0.176	0.201	0.177	0.199	0.186	0.185	0.224	0.192	0.237	0.213	0.204	0.224	0.104	0.152	0.149	0.138	0.155	0.135	0.177	0.174	0.072	0.206
Oddavalan	0.216	0.223	0.227	0.266	0.19	0.281	0.184	0.254	0.272	0.183	0.215	0.305	0.34	1	0.327	0.32	0.19	0.197	0.22	0.181	0.226	0.215	0.142	0.172	0.161	0.159	0.142	0.2	0.178	0.129	0.171	0.136	0.141	0.15	0.175	0.125	0.171	0.184	0.149	0.193	0.152	0.187	0.205	0.094	0.114	0.118	0.088	0.124	0.107	0.116	0.126	0.067	0.177
Hathiel	0.192	0.186	0.185	0.212	0.156	0.223	0.137	0.18	0.224	0.119	0.193	0.257	0.244	0.327	1	0.235	0.175	0.149	0.229	0.127	0.177	0.183	0.125	0.152	0.099	0.125	0.119	0.181	0.137	0.133	0.179	0.134	0.113	0.171	0.147	0.089	0.149	0.143	0.163	0.195	0.151	0.151	0.157	0.088	0.103	0.122	0.096	0.135	0.132	0.119	0.108	0.067	0.157
Kurulu	0.18	0.245	0.204	0.209	0.179	0.251	0.202	0.266	0.274	0.242	0.282	0.221	0.314	0.32	0.235	1	0.212	0.178	0.171	0.165	0.196	0.209	0.137	0.167	0.142	0.163	0.187	0.23	0.213	0.163	0.148	0.137	0.172	0.188	0.152	0.163	0.196	0.194	0.161	0.196	0.148	0.164	0.186	0.137	0.127	0.141	0.102	0.164	0.152	0.153	0.149	0.083	0.218
Murun137	0.153	0.242	0.234	0.232	0.204	0.231	0.24	0.206	0.22	0.191	0.161	0.219	0.228	0.19	0.175	0.212	1	0.233	0.255	0.239	0.28	0.266	0.278	0.289	0.149	0.163	0.248	0.303	0.236	0.25	0.164	0.193	0.235	0.162	0.192	0.155	0.211	0.201	0.254	0.202	0.206	0.223	0.218	0.138	0.17	0.177	0.14	0.216	0.205	0.187	0.184	0.124	0.2
Dikwee	0.165	0.194	0.242	0.167	0.161	0.261	0.151	0.16	0.172	0.187	0.134	0.155	0.193	0.197	0.149	0.178	0.233	1	0.295	0.31	0.337	0.224	0.218	0.186	0.155	0.167	0.168	0.147	0.183	0.168	0.126	0.141	0.121	0.115	0.122	0.093	0.123	0.13	0.128	0.129	0.124	0.157	0.149	0.093	0.107	0.117	0.107	0.164	0.106	0.141	0.124	0.123	0.175
Kahatawee	0.157	0.237	0.191	0.165	0.191	0.26	0.167	0.157	0.145	0.162	0.142	0.18	0.213	0.22	0.229	0.171	0.255	0.295	1	0.304	0.316	0.211	0.269	0.211	0.147	0.16	0.151	0.157	0.173	0.178	0.167	0.158	0.142	0.145	0.131	0.14	0.173	0.133	0.158	0.137	0.161	0.168	0.153	0.109	0.137	0.134	0.11	0.139	0.136	0.138	0.127	0.062	0.172
Hatapan	0.121	0.198	0.192	0.191	0.186	0.234	0.167	0.176	0.124	0.211	0.141	0.163	0.192	0.181	0.127	0.165	0.239	0.31	0.304	1	0.314	0.274	0.298	0.235	0.185	0.165	0.143	0.15	0.202	0.204	0.14	0.158	0.161	0.173	0.143	0.139	0.166	0.145	0.165	0.147	0.153	0.153	0.139	0.137	0.158	0.155	0.1	0.146	0.113	0.181	0.147	0.071	0.207
Niyanwee	0.218	0.183	0.237	0.17	0.169	0.228	0.172	0.172	0.187	0.174	0.145	0.196	0.188	0.226	0.177	0.196	0.28	0.337	0.316	0.314	1	0.271	0.249	0.21	0.174	0.176	0.164	0.155	0.19	0.201	0.152	0.149	0.168	0.157	0.14	0.167	0.19	0.15	0.141	0.145	0.136	0.16	0.173	0.179	0.165	0.154	0.116	0.159	0.147	0.157	0.179	0.088	0.166
Peilliane	0.196	0.233	0.237	0.249	0.169	0.246	0.215	0.249	0.217	0.208	0.164	0.216	0.2	0.215	0.183	0.209	0.266	0.224	0.211	0.274	0.271	1	0.341	0.344	0.199	0.16	0.181	0.192	0.259	0.216	0.19	0.141	0.168	0.174	0.163	0.175	0.199	0.173	0.156	0.152	0.152	0.144	0.158	0.124	0.211	0.145	0.125	0.194	0.156	0.2	0.153	0.096	0.253
Dewaradd	0.169	0.189	0.183	0.211	0.188	0.172	0.162	0.162	0.171	0.175	0.101	0.189	0.182	0.142	0.125	0.137	0.278	0.218	0.269	0.298	0.249	0.341	1	0.298	0.13	0.121	0.134	0.15	0.175	0.16	0.146	0.135	0.133	0.113	0.135	0.13	0.115	0.138	0.1	0.128	0.116	0.123	0.123	0.102	0.134	0.147	0.119	0.137	0.125	0.167	0.116	0.092	0.181
Murungaka	0.197	0.231	0.241	0.216	0.22	0.244	0.247	0.23	0.199	0.169	0.131	0.159	0.194	0.172	0.152	0.167	0.289	0.186	0.211	0.235	0.21	0.344	0.298	1	0.187	0.182	0.159	0.2	0.214	0.215	0.169	0.174	0.156	0.152	0.151	0.146	0.207	0.119	0.145	0.173	0.147	0.162	0.19	0.121	0.152	0.165	0.099	0.147	0.128	0.168	0.141	0.055	0.258
Kottamall	0.143	0.208	0.245	0.171	0.189	0.223	0.216	0.194	0.143	0.149	0.171	0.139	0.148	0.161	0.099	0.142	0.149	0.155	0.147	0.185	0.174	0.199	0.13	0.187	1	0.311	0.182	0.171	0.2	0.256	0.191	0.15	0.169	0.158	0.156	0.136	0.175	0.182	0.142	0.146	0.145	0.153	0.167	0.143	0.167	0.173	0.109	0.152	0.148	0.167	0.163	0.089	0.19
Red leaf	0.154	0.207	0.263	0.167	0.195	0.237	0.234	0.155	0.147	0.159	0.15	0.144	0.173	0.159	0.125	0.163	0.163	0.167	0.16	0.165	0.176	0.16	0.121	0.182	0.311	1	0.161	0.16	0.192	0.221	0.144	0.102	0.192	0.148	0.114	0.157	0.177	0.17	0.155	0.134	0.144	0.144	0.143	0.105	0.105	0.115	0.105	0.102	0.098	0.12	0.152	0.076	0.15
Wedahee	0.155	0.177	0.161	0.145	0.158	0.163	0.157	0.185	0.184	0.193	0.15	0.167	0.176	0.142	0.119	0.187	0.248	0.168	0.151	0.143	0.164	0.181	0.134	0.159	0.182	0.161	1	0.488	0.479	0.313	0.278	0.304	0.341	0.357	0.316	0.295	0.378	0.38	0.373	0.266	0.351	0.408	0.329	0.279	0.275	0.277	0.262	0.329	0.328	0.319	0.358	0.163	0.221
Mudukirie	0.18	0.196	0.195	0.216	0.196	0.188	0.207	0.213	0.224	0.162	0.243	0.206	0.21	0.2	0.181	0.23	0.303	0.147	0.157	0.15	0.155	0.192	0.15	0.2	0.171	0.16	0.488	1	0.409	0.342	0.243	0.313	0.292	0.283	0.323	0.24	0.33	0.375	0.37	0.353	0.309	0.367	0.347	0.217	0.233	0.218	0.205	0.289	0.333	0.297	0.267	0.126	0.241
Suduheen	0.197	0.203	0.186	0.17	0.186	0.238	0.179	0.185	0.177	0.203	0.162	0.161	0.201	0.178	0.137	0.213	0.236	0.183	0.173	0.202	0.19	0.259	0.175	0.214	0.2	0.192	0.479	0.409	1	0.279	0.296	0.307	0.326	0.338	0.234	0.242	0.407	0.315	0.307	0.261	0.273	0.333	0.274	0.286	0.333	0.338	0.236	0.313	0.263	0.372	0.304	0.193	0.347
O.rhizom	0.209	0.176	0.197	0.189	0.197	0.183	0.207	0.233	0.156	0.181	0.192	0.153	0.169	0.129	0.133	0.163	0.25	0.168	0.178	0.204	0.201	0.216	0.16	0.215	0.256	0.221	0.313	0.342	0.279	1	0.252	0.245	0.31	0.252	0.232	0.262	0.271	0.265	0.269	0.255	0.207	0.252	0.277	0.231	0.196	0.219	0.168	0.233	0.255	0.194	0.276	0.124	0.21
0.rufipo	0.195	0.156	0.178	0.177	0.21	0.171	0.135	0.178	0.184	0.175	0.154	0.186	0.155	0.171	0.179	0.148	0.164	0.126	0.167	0.14	0.152	0.19	0.146	0.169	0.191	0.144	0.278	0.243	0.296	0.252	1	0.341	0.231	0.326	0.367	0.224	0.301	0.306	0.258	0.287	0.227	0.24	0.243	0.276	0.23	0.275	0.185	0.217	0.256	0.295	0.31	0.204	0.298
Godaalwee	0.146	0.185	0.161	0.146	0.184	0.142	0.151	0.164	0.138	0.137	0.15	0.149	0.176	0.136	0.134	0.137	0.193	0.141	0.158	0.158	0.149	0.141	0.135	0.174	0.15	0.102	0.304	0.313	0.307	0.245	0.341	1	0.274	0.291	0.37	0.213	0.313	0.289	0.294	0.272	0.274	0.259	0.289	0.308	0.25	0.324	0.217	0.282	0.333	0.325	0.324	0.163	0.309
Gonabaru	0.126	0.157	0.144	0.15	0.211	0.159	0.186	0.192	0.178	0.203	0.163	0.187	0.201	0.141	0.113	0.172	0.235	0.121	0.142	0.161	0.168	0.168	0.133	0.156	0.169	0.192	0.341	0.292	0.326	0.31	0.231	0.274	1	0.242	0.245	0.28	0.29	0.351	0.274	0.279	0.269	0.297	0.309	0.247	0.247	0.232	0.188	0.233	0.244	0.215	0.313	0.121	0.218
Kaluheen	0.137	0.162	0.147	0.153	0.168	0.164	0.172	0.186	0.131	0.195	0.198	0.177	0.177	0.15	0.171	0.188	0.162	0.115	0.145	0.173	0.157	0.174	0.113	0.152	0.158	0.148	0.357	0.283	0.338	0.252	0.326	0.291	0.242	1	0.272	0.282	0.423	0.367	0.36	0.305	0.286	0.32	0.227	0.305	0.299	0.367	0.226	0.278	0.275	0.343	0.343	0.136	0.264
Balamawee	0.164	0.168	0.175	0.19	0.182	0.135	0.125	0.163	0.13	0.145	0.213	0.183	0.199	0.175	0.147	0.152	0.192	0.122	0.131	0.143	0.14	0.163	0.135	0.151	0.156	0.114	0.316	0.323	0.234	0.232	0.367	0.37	0.245	0.272	1	0.267	0.279	0.33	0.354	0.352	0.317	0.317	0.286	0.265	0.231	0.213	0.214	0.217	0.311	0.256	0.222	0.148	0.216
Suduheena	0.168	0.14	0.127	0.126	0.161	0.13	0.179	0.146	0.133	0.225	0.144	0.16	0.186	0.125	0.089	0.163	0.155	0.093	0.14	0.139	0.167	0.175	0.13	0.146	0.136	0.157	0.295	0.24	0.242	0.262	0.224	0.213	0.28	0.282	0.267	1	0.259	0.31	0.227	0.233	0.25	0.235	0.267	0.258	0.257	0.239	0.224	0.181	0.221	0.19	0.231	0.101	0.198
o.granata	0.148	0.194	0.147	0.184	0.176	0.164	0.171	0.206	0.176	0.147	0.178	0.177	0.185	0.171	0.149	0.196	0.211	0.123	0.173	0.166	0.19	0.199	0.115	0.207	0.175	0.177	0.378	0.33	0.407	0.271	0.301	0.313	0.29	0.423	0.279	0.259	1	0.369	0.397	0.371	0.359	0.342	0.307	0.292	0.27	0.271	0.221	0.303	0.25	0.312	0.311	0.138	0.245
o.nivara	0.131	0.147	0.17	0.191	0.184	0.124	0.171	0.225	0.23	0.202	0.252	0.185	0.224	0.184	0.143	0.194	0.201	0.13	0.133	0.145	0.15	0.173	0.138	0.119	0.182	0.17	0.38	0.375	0.315	0.265	0.306	0.289	0.351	0.367	0.33	0.31	0.369	1	0.365	0.325	0.329	0.299	0.356	0.23	0.247	0.231	0.25	0.262	0.275	0.286	0.333	0.178	0.252
kombili	0.128	0.169	0.161	0.183	0.175	0.156	0.157	0.184	0.168	0.146	0.195	0.185	0.192	0.149	0.163	0.161	0.254	0.128	0.158	0.165	0.141	0.156	0.1	0.145	0.142	0.155	0.373	0.37	0.307	0.269	0.258	0.294	0.274	0.36	0.354	0.227	0.397	0.365	1	0.295	0.372	0.372	0.318	0.25	0.234	0.216	0.183	0.266	0.28	0.259	0.225	0.124	0.218
0.eii	0.15	0.182	0.162	0.213	0.225	0.157	0.187	0.249	0.22	0.143	0.229	0.189	0.237	0.193	0.195	0.196	0.202	0.129	0.137	0.147	0.145	0.152	0.128	0.173	0.146	0.134	0.266	0.353	0.261	0.255	0.287	0.272	0.279	0.305	0.352	0.233	0.371	0.325	0.295	1	0.327	0.268	0.325	0.188	0.161	0.202	0.133	0.205	0.191	0.234	0.22	0.107	0.254
Kalukura	0.122	0.164	0.127	0.148	0.17	0.159	0.153	0.173	0.156	0.159	0.181	0.198	0.213	0.152	0.151	0.148	0.206	0.124	0.161	0.153	0.136	0.152	0.116	0.147	0.145	0.144	0.351	0.309	0.273	0.207	0.227	0.274	0.269	0.286	0.317	0.25	0.359	0.329	0.372	0.327	1	0.405	0.329	0.206	0.278	0.261	0.266	0.311	0.257	0.269	0.284	0.115	0.188
Kohumawi	0.15	0.164	0.127	0.163	0.17	0.18	0.16	0.153	0.171	0.177	0.181	0.162	0.204	0.187	0.151	0.164	0.223	0.157	0.168	0.153	0.16	0.144	0.123	0.162	0.153	0.144	0.408	0.367	0.333	0.252	0.24	0.259	0.297	0.32	0.317	0.235	0.342	0.299	0.372	0.268	0.405	1	0.314	0.262	0.243	0.243	0.227	0.276	0.257	0.253	0.234	0.141	0.237
Suwadasa	0.148	0.178	0.192	0.169	0.219	0.171	0.184	0.205	0.184	0.183	0.223	0.203	0.224	0.205	0.157	0.186	0.218	0.149	0.153	0.139	0.173	0.158	0.123	0.19	0.167	0.143	0.329	0.347	0.274	0.277	0.243	0.289	0.309	0.227	0.286	0.267	0.307	0.356	0.318	0.325	0.329	0.314	1	0.213	0.232	0.231	0.216	0.293	0.244	0.241	0.284	0.116	0.24
Madael	0.149	0.113	0.099	0.083	0.133	0.091	0.084	0.11	0.106	0.17	0.123	0.142	0.104	0.094	0.088	0.137	0.138	0.093	0.109	0.137	0.179	0.124	0.102	0.121	0.143	0.105	0.279	0.217	0.286	0.231	0.276	0.308	0.247	0.305	0.265	0.258	0.292	0.23	0.25	0.188	0.206	0.262	0.213	1	0.373	0.444	0.341	0.364	0.34	0.375	0.377	0.19	0.273
Godahee	0.135	0.162	0.138	0.114	0.168	0.141	0.124	0.158	0.114	0.165	0.131	0.178	0.152	0.114	0.103	0.127	0.17	0.107	0.137	0.158	0.165	0.211	0.134	0.152	0.167	0.105	0.275	0.233	0.333	0.196	0.23	0.25	0.247	0.299	0.231	0.257	0.27	0.247	0.234	0.161	0.278	0.243	0.232	0.373	1	0.442	0.38	0.371	0.397	0.403	0.361	0.133	0.241
Hondarawa	0.152	0.142	0.135	0.118	0.175	0.131	0.114	0.141	0.118	0.194	0.118	0.186	0.149	0.118	0.122	0.141	0.177	0.117	0.134	0.155	0.154	0.145	0.147	0.165	0.173	0.115	0.277	0.218	0.338	0.219	0.275	0.324	0.232	0.367	0.213	0.239	0.271	0.231	0.216	0.202	0.261	0.243	0.231	0.444	0.442	1	0.391	0.455	0.333	0.439	0.418	0.212	0.272
Heendikwe	0.139	0.114	0.116	0.084	0.144	0.092	0.078	0.104	0.107	0.15	0.114	0.133	0.138	0.088	0.096	0.102	0.14	0.107	0.11	0.1	0.116	0.125	0.119	0.099	0.109	0.105	0.262	0.205	0.236	0.168	0.185	0.217	0.188	0.226	0.214	0.224	0.221	0.25	0.183	0.133	0.266	0.227	0.216	0.341	0.38	0.391	1	0.345	0.296	0.333	0.385	0.213	0.198
Rathuh	0.11	0.14	0.141	0.155	0.171	0.116	0.133	0.153	0.171	0.198	0.144	0.171	0.155	0.124	0.135	0.164	0.216	0.164	0.139	0.146	0.159	0.194	0.137	0.147	0.152	0.102	0.329	0.289	0.313	0.233	0.217	0.282	0.233	0.278	0.217	0.181	0.303	0.262	0.266	0.205	0.311	0.276	0.293	0.364	0.371	0.455	0.345	1	0.41	0.484	0.397	0.169	0.27
chembava	0.176	0.161	0.137	0.16	0.157	0.126	0.143	0.136	0.121	0.154	0.13	0.167	0.135	0.107	0.132	0.152	0.205	0.106	0.136	0.113	0.147	0.156	0.125	0.128	0.148	0.098	0.328	0.333	0.263	0.255	0.256	0.333	0.244	0.275	0.311	0.221	0.25	0.275	0.28	0.191	0.257	0.257	0.244	0.34	0.397	0.333	0.296	0.41	1	0.354	0.292	0.194	0.267
Murunga	0.147	0.162	0.162	0.153	0.205	0.149	0.138	0.158	0.138	0.216	0.151	0.177	0.177	0.116	0.119	0.153	0.187	0.141	0.138	0.181	0.157	0.2	0.167	0.168	0.167	0.12	0.319	0.297	0.372	0.194	0.295	0.325	0.215	0.343	0.256	0.19	0.312	0.286	0.259	0.234	0.269	0.253	0.241	0.375	0.403	0.439	0.333	0.484	0.354	1	0.429	0.179	0.369
Chemba	0.162	0.126	0.159	0.127	0.183	0.153	0.135	0.169	0.158	0.161	0.119	0.164	0.174	0.126	0.108	0.149	0.184	0.124	0.127	0.147	0.179	0.153	0.116	0.141	0.163	0.152	0.358	0.267	0.304	0.276	0.31	0.324	0.313	0.343	0.222	0.231	0.311	0.333	0.225	0.22	0.284	0.234	0.284	0.377	0.361	0.418	0.385	0.397	0.292	0.429	1	0.158	0.261
hygroryza	0.101	0.079	0.083	0.069	0.079	0.089	0.059	0.077	0.096	0.112	0.067	0.116	0.072	0.067	0.067	0.083	0.124	0.123	0.062	0.071	0.088	0.096	0.092	0.055	0.089	0.076	0.163	0.126	0.193	0.124	0.204	0.163	0.121	0.136	0.148	0.101	0.138	0.178	0.124	0.107	0.115	0.141	0.116	0.19	0.133	0.212	0.213	0.169	0.194	0.179	0.158	1	0.165
Pokkali	0.203	0.2	0.237	0.198	0.208	0.226	0.197	0.223	0.213	0.256	0.194	0.201	0.206	0.177	0.157	0.218	0.2	0.175	0.172	0.207	0.166	0.253	0.181	0.258	0.19	0.15	0.221	0.241	0.347	0.21	0.298	0.309	0.218	0.264	0.216	0.198	0.245	0.252	0.218	0.254	0.188	0.237	0.24	0.273	0.241	0.272	0.198	0.27	0.267	0.369	0.261	0.165	1

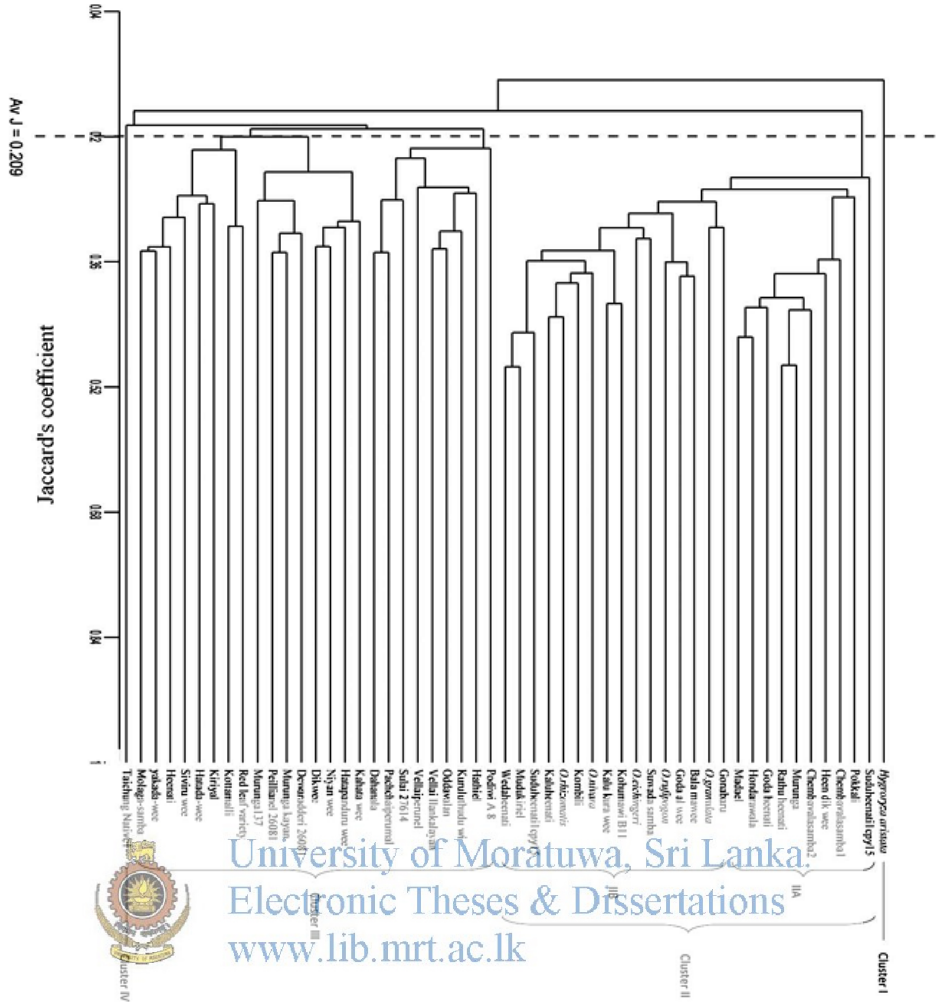

FIGURE 1: The UPGMA dendrogram showing genetic diversity among Sri Lankan traditional rice varieties