

The Splits in the Neighborhood of a Tree

David Bryant

McGill Centre for Bioinformatics, 3775 University, Montréal, Québec H3A 2B4, Canada
bryant@mcb.mcgill.ca

Received April 17, 2003

AMS Subject Classification: 68R10, 05C05, 68Q25, 92D15

Abstract. A *phylogenetic tree* represents historical evolutionary relationships between different species or organisms. The space of possible phylogenetic trees is both complex and exponentially large. Here we study combinatorial features of neighbourhoods within this space, with respect to four standard tree metrics. We focus on the *splits* of a tree: the bipartitions induced by removing a single edge from the tree. We characterize those splits appearing in trees that are within a given distance of the original tree, demonstrating close connections between these splits, the Whitney number of a tree, and the binary characters with a given parsimony length.

Keywords: phylogenetic tree, splits, tree metric, Whitney numbers, parsimony

References

1. B.L. Allen and M. Steel, Subtree transfer operations and their induced metrics on evolutionary trees, *Ann. Combin.* **5** (1) (2001) 1–15.
2. L.J. Billera, S.P. Holmes, and K. Vogtman, Geometry of the space of phylogenetic trees, *Adv. Appl. Math.* **27** (4) (2001) 733–767.
3. D. Bryant, Hunting for trees in binary character sets, *J. Comput. Biol.* **3** (2) (1996) 275–288.
4. D. Bryant, Building trees, hunting for trees and comparing trees, Ph.D. Thesis, Department of Mathematics, University of Canterbury, 1997.
5. D. Bryant and M. Steel, Constructing optimal trees from quartets, *J. Algorithms* **38** (1) (2001) 237–259, Tenth Annual ACM-SIAM Symposium on Discrete Algorithms, Baltimore, MD, 1999.
6. P. Buneman, The recovery of trees from measures of dissimilarity, In: *Mathematics in the Archaeological and Historical Sciences*, F.R. Hodson, D.G. Kendall, and P. Tautu, Eds., Edinburgh University Press, Edinburgh, 1971, pp. 387–395.
7. M. Carter, M. Hendy, D. Penny, L.A. Székely, and N.C. Wormald, On the distributions of lengths of evolutionary trees, *SIAM J. Discrete Math.* **3** (1990) 38–47.
8. M. Charleston and M. Steel, Five surprising properties of parsimoniously colored trees, *Bull. Math. Biol.* **57** (2) (1993) 367–375.

9. B. DasGupta, X. He, T. Jiang, M. Li, J. Tromp, and L. Zhang, On computing the nearest neighbor interchange distance, In: *Discrete Mathematical Problems with Medical Applications*, New Brunswick, NJ, 1999, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Vol. 55, Amer. Math. Soc., Providence, RI, 2000, pp. 125–143.
10. W.H.E. Day, Optimal algorithms for comparing trees with labeled leaves, *J. Classification* **2** (1985) 7–28.
11. J. Hein, T. Jiang, L. Wang, and K. Zhang, On the complexity of comparing evolutionary trees, *Discrete Appl. Math.* **71** (1-3) (1996) 153–169.
12. R.E. Jamison, Alternating Whitney sums and matchings in trees, I, *Discrete Math.* **67** (2) (1987) 177–189.
13. D.F. Robinson, Comparison of labeled trees with valency three, *J. Combin. Theory, Ser. B* **11** (1971) 105–119.
14. D.F. Robinson and L.R. Foulds, Comparison of weighted labelled trees, In: *Combinatorial Mathematics, VI, Proc. Sixth Austral. Conf., Univ. New England, Armidale, 1978*, Lecture Notes in Mathematics, Vol. 748, Springer-Verlag, Berlin, 1979, pp. 119–126.
15. D.F. Robinson and L.R. Foulds, Comparison of phylogenetic trees, *Math. Biosci.* **53** (1981) 131–147.
16. C. Semple and M. Steel, *Phylogenetics*, Oxford University Press, 2003.