and the amount of scattering determines only the spatial extent of the modes (and thereby the cavity volume). For 3D structures, like powders or porous networks, light can also be localized, but the amount of scattering has to exceed a certain critical value. Although Sapienza *et al*. did not reach the strong coupling regime, they did see a large effect on the emission of the dots, in the form of a strong emission enhancement. This so-called Purcell effect is a precursor to strong coupling and shows that the strategy is very promising.

The benefits of using disordered materials can be realized both in fabrication and device operation. Disordered structures are much easier and cheaper to make than extremely precise nanoscale cavities. It sounds almost trivial, but disordered structures are robust against disorder. We cannot predict in advance which source will couple to which mode, but the end result will be a broad set of sources in a range of frequencies that are all coupled to cavity modes. For example, it may be possible to build a single-photon source that is also broadband, meaning that it emits not at just one frequency but over a range of frequencies. The emitters could then be controlled by secondary light beams at frequencies outside the localization band.

The approach of Sapienza *et al*. also raises interesting possibilities in quantum information processing. When Anderson-localized modes occur at the same or nearby frequencies, they can couple to each other to form so-called necklace states. A series of localized modes can exchange photons and thereby share the quantum information of the atoms or dots located in each mode. It might be much simpler to construct a large-size quantum memory or information processor this way than with perfect, periodic coupled nanocavities (see the second figure). The information contained in such a random structure can be written, read, and erased by external light beams that are also multiply scattered, but not localized, and carry within their speckle pattern the information stored in the quantum network. This is but one of many possibilities that may arise from the use of disordered photonic structures.

References and Notes

1. M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, UK, 2000).
2. T. Wilk, S. C. Webster, A. Kuhn, G. Rempe, *Science* **317**, 488 (2007).
3. L. Sapienza *et al*., *Science* **327**, 1352 (2010).
4. For an inspiring alternative view on beauty, see (*5*).
5. T. Itoh, *Wabi Sabi Suki: The Essence of Japanese Beauty* (Mazda Motor Corporation, Hiroshima, 1993).
6. J. D. Joannopoulos, R. D. Meade, J. N. Winn, *Photonic Crystals: Molding the Flow of Light* (Princeton Univ. Press, Princeton, NJ, 1995).
7. K. Sakoda, *Optical Properties of Photonic Crystals* (Springer-Verlag, Berlin, 2001).
8. A. F. Koenderink *et al*., *Phys. Rev. B* **72**, 153102 (2005).
9. A. Lagendijk, B. van Tiggelen, D. S. Wiersma, *Phys. Today* **62**, 24 (2009).
10. D. P. DiVincenzo, *Fortschr. Phys.* **48**, 771 (2000).
11. H. J. Kimble, *Nature* **453**, 1023 (2008).
12. D. S. Wiersma, *Nat. Phys.* **4**, 359 (2008).

10.1126/science.1187084

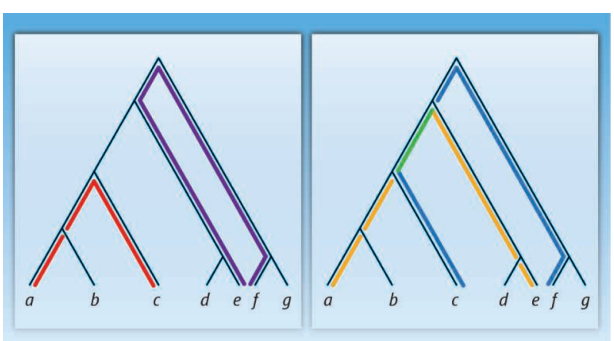
EVOLUTION

Trees, Fast and Accurate

A scalable and fast method for building very large evolutionary trees achieves greater accuracy than previously thought.

Elizabeth S. Allman and John A. Rhodes

Inferring evolutionary relationships from DNA or protein sequence data is now routine in biological investigations. Although techniques have improved for using such information to generate phylogenetic trees that represent these histories (*1*), challenges remain. In particular, the computational demands of preferred statistical approaches make them infeasible for finding large phylogenies. However, on page 1376 of this issue (*2*), Roch theorizes that a less comprehensive analysis might be surprisingly reliable.



Distance estimates. Shown is an example of a "true" phylogenetic tree representing the evolutionary history of seven taxa (*a*, *b*, *c*, *d*, *e*, *f*, and *g*). Pairwise distances between taxa are computed from orthologous gene sequences. Distances are random variables reflecting path lengths in the tree. (**Left**) The paths between *a* and *c* (red) and between *e* and *f* (purple) do not share a common history, so the random variables [distances $d(a, c)$ and $d(e, f)$] are independent of each other. (**Right**) The paths between *a* and *e* and between *c* and *f* share a common history (green), so independence is lost; $d(a, e)$ (yellow and green) and $d(c, f)$ (blue and green) are correlated.

For a moderate number of taxa, the now-dominant methods adopt a probabilistic model of sequence evolution and then, using either maximum likelihood (ML) or Bayesian frameworks, determine the trees that best fit the sequence data. Such analyses are well grounded in statistical tradition and, if the model approximates reality, can be expected to perform well. However, they are computationally intensive: ML leads to a complex optimization problem not only over many numerical parameters, but also over all trees that might relate the taxa. Bayesian approaches involve a random walk over these trees. For n taxa there are $1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n - 5)$ trees to be considered, so that when n is large, the size of tree space is a fundamental impediment. This explosion in the number of trees results in unacceptably long run times of computer analyses, so searches for optimal trees may be incomplete.

An alternative faster approach is to algorithmically construct a single tree by successively grouping closely related taxa. The sequences for each pair of taxa *a*, *b* are compared, and a single number $d(a, b)$ is computed to express their distance (relationship) along the unknown tree. The sequences are then discarded, as all decisions on joining taxa are based solely on these distances. Two well-known examples of such agglomerative algorithms are the unweighted pair group method with arithmetic mean and the more reliable neighbor-joining (NJ) method.

The apparent drawback of distance-based techniques lies in only comparing sequences pairwise—the full information that could be extracted from n-way comparisons is not used. Thus, distance methods do not perform as well as more complete statistical

Department of Mathematics and Statistics, University of Alaska, P.O. Box 756660, Fairbanks, AK 99775, USA. E-mail: e.allman@alaska.edu; j.rhodes@alaska.edu

approaches. However, for large trees, only distance methods are practical.

Roch shows, through a theoretical examination, that agglomerative construction of trees from distances can perform better than one might expect. The basic insight is to take advantage of correlations among the distances that result from a shared evolutionary history (see the figure). Agglomeration joins groups of taxa iteratively, so that at any stage, partial knowledge of the tree allows an algorithm to make use of such correlations. Using additional ideas—in particular, a "blindfolded cherry-picking" algorithm that allows mistaken groupings to be dissolved (*3*)—Roch provides an agglomerative algorithm that effectively exploits this extra information in the distances.

This new method is surprisingly efficient, as measured by the length of sequences required to return the correct tree with high probability. Roch determined that the required sequence length had the same dependence on key tree features—the number of taxa, the depth of the tree, and the shortest edge length—as that conjectured for ML, up to an undetermined constant factor. Although the sequence length requirement for a different, nonagglomerative distance method (*4*) matches Roch's if some tree branches are long, it is notable that agglomeration need not weaken performance.

Another surprise of the analysis by Roch is that when all tree branches are short, the sequence length requirement is independent of the depth of the tree. It has long been known that deep divergences in trees can be difficult to infer, as more recent mutations obscure the signal of older ones. However, short branches imply that more bifurcations occur in the tree, giving additional sources of information. A "phase transition" occurs at a critical value of branch length, below which depth is irrelevant to the required sequence length. This sort of behavior holds for a more elaborate statistical method of phylogenetic reconstruction (*3*), but now we learn that even distance methods can achieve it.

Unfortunately, theoretical results on sequence length requirements for ML are lacking, as are precise values of the constants in Roch's result, preventing a full comparison of the methods. Moreover, although ML simultaneously fits numerical model parameters and trees, Roch's method assumes that the mutation rate parameters are known, so in practice they would have to be inferred by some other procedure. Additionally, the models Roch considers lack mutation rate variation among sites along the sequence. Still, all models are likely to be deficient under some circumstances, and the potential shown for both speed and accuracy, even under restrictive assumptions, is tantalizing.

Most current practical agglomerative algorithms, including NJ, were not designed to account for distance correlations. A notable exception is the BIONJ algorithm (*5*), which does consider distance covariances to some extent, and through simulation appears to offer improved performance. Roch's algorithm, while ingenious, is complex and designed for the theoretical analysis of what is possible. The next challenge is to translate the understanding that Roch provides into fast, practical algorithms that more fully exploit the distance correlations. New software must strike a balance between extracting more information from distances and possible speed reductions over simpler agglomerative methods. Although Roch provides a roadmap for one approach, practical trade-offs remain to be seen.

References and Notes

1. J. Felsenstein, *Inferring Phylogenies* (Sinauer, Sunderland, MA, 2004).
2. S. Roch, *Science* **327**, 1376 (2010).
3. C. Daskalakis, E. Mossel, S. Roch, *Probab. Theory Relat. Fields* 10.1007/s00440-009-0246-2 (2009).
4. P. L. Erdős, M. A. Steel, L. A. Székely, T. J. Warnow, *Random Struct. Alg.* **14**, 153 (1999).
5. O. Gascuel, *Mol. Biol. Evol.* **14**, 685 (1997).
6. Supported by NSF grant DMS 0714830.

10.1126/science.1187797

BIOPHYSICS

Enforcing Order on Signaling

Mechanical forces constrain the spatial organization of a cell surface receptor, thereby altering its signaling function and cell behavior.

Matthew Paszek[1] and Valerie Weaver[1,2,3,4]

Mechanical forces provide important regulatory information that directs development. Even throughout adulthood, tissue homeostasis remains tightly linked to tensional homeostasis, the perturbation of which often leads to chronic conditions such as cardiovascular disease, arthritis, and cancer (*1*, *2*). Environmental cues affect cell behavior by triggering signal transduction networks, but how cells actually integrate mechanical cues with these biochemical networks remains largely unresolved. On page 1380 in this issue, Salaita *et al*. describe how mechanical force, spatial organization of large clusters of cell surface receptors, and receptor-mediated signal transduction are coupled (*3*). Disruption of this mechanical-coupling mechanism in tumor cells may explain the invasive characteristics of aggressive, metastatic cancers.

Most models of cellular "mechanotransduction" incorporate the idea that proteins subjected to force undergo conformational changes that alter their function. One example is illustrated by the force-dependent assembly of integrin adhesion structures called focal adhesions. An integrin is a cell surface transmembrane receptor that interacts with the extracellular matrix outside the cell and with cytoskeletal and signaling molecules through its intracellular domain. In a "protein-centric" model of mechanotransduction, cytoskeletal tension on integrins and associated focal adhesion proteins drives the unfolding of key adaptor proteins that are linked to the "tensed" integrin. These conformational changes reveal binding sites that permit association with a network of proteins essential for signal transduction (*4*).

In contrast to this classic protein-based mechanotransduction model, Salaita *et al*. postulate that mechanical cues can impose spatial patterning on cell surface receptors that alter their signaling function. Many transmembrane receptors at the cell surface assemble into dimers, trimers, or higher-order oligomers to initiate intracellular signaling. In a "membrane-centric" mechanotransduction mechanism (*5*), Salaita *et al*. propose that either applied force or cellular tension can affect lateral receptor movement within the membrane to change the size (degree of oligomerization) of receptor clusters, or alter their spatial organization throughout the

[1]Center for Bioengineering and Tissue Regeneration, Department of Surgery, University of California, San Francisco (UCSF), San Francisco, CA 94143, USA. [2]Department of Anatomy, UCSF, San Francisco, CA 94143, USA. [3]Department of Bioengineering and Therapeutic Sciences, Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, and Helen Diller Comprehensive Cancer Center, UCSF, San Francisco, CA 94143, USA. [4]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA. E-mail: valerie.weaver@ucsfmedctr.org