

Loblolly pine's immense genome conquered

March 20 2014



The loblolly pine -- whose genome is the largest ever sequenced -- is the most commercially important tree species in the United States and the source of most American paper products. Credit: Ron Billings/Texas A&M Forest Service

The massive genome sequence of the loblolly pine—the most commercially important tree species in the United States and the source of most American paper products—has been completed by a nationwide research team, led by a UC Davis scientist.

The draft [genome](#)—approximately seven times bigger than the human

genome—is the largest genome sequenced to date and the most complete conifer genome sequence ever published. The sequencing was accomplished by using, for the first time, a faster and more efficient analytical process. The achievement is described in two papers in the March 2014 issue of *Genetics* and in one paper in the open access journal *Genome Biology*.

The genome sequence will help scientists breed improved varieties of the loblolly pine, which also is being developed as a feedstock for biofuel. The newly sequenced genome also provides a better understanding of the evolution and diversity of plants.

"It's a huge genome. But the challenge isn't just collecting all the sequence data. The problem is assembling that sequence into order," said David Neale, a professor of plant sciences at the University of California, Davis, who led the loblolly pine genome project and is an author on the *Genetics* and *Genome Biology* articles.

To tackle the enormous size of the loblolly pine's genome, which until recently has been an obstacle to sequencing efforts, the research team used a new method that can speed up genome assembly by compressing the raw sequence data 100-fold.

Modern [genome sequencing](#) methods make it relatively easy to read the individual "letters" in DNA, but only in short fragments. In the case of the loblolly, 16 billion separate fragments had to be fit back together—a computational puzzle called genome assembly.

"We were able to assemble the human genome, but that was close to the limit of our ability; seven times bigger was just too much," said Steven Salzberg, professor of medicine and biostatistics at Johns Hopkins University, one of the directors of the loblolly genome assembly team and an author on the papers.

The key to the solution was using a new method, developed by researchers at the University of Maryland, which pre-processes the sequence data, eliminates redundancies and yields 100 times less sequence data. This approach, tested for the first time in this study, allowed the team to assemble a much more complete [genome sequence](#) than the draft assemblies of two other conifer species reported last year.

"The size of the pieces of consecutive sequence that we assembled are orders of magnitude larger than what's been previously published," said Neale, noting that the loblolly now provides a high-quality "reference" genome that considerably speeds along future conifer genome projects.

The loblolly genome research was conducted in an open-access manner, benefiting the research community even before the genome sequencing effort was completed and published. Data have been freely available throughout the project, with three public releases starting in June 2012.

The new sequencing confirmed that 82 percent of the loblolly genome is made up of invasive DNA elements and other DNA fragments that copied themselves around the genome. The genome sequencing also revealed the location of genes that may be involved in fighting off pathogens, which will help scientists understand more about disease resistance in pines.

For example, researchers from the Forest Service Southern Institute for Forest Genetics identified an important candidate gene for resistance to fusiform rust, the most damaging disease of southern pines. A molecular understanding of genetic resistance is a valuable tool for forest managers as they select trees that will develop into healthy stands.

"The fusiform rust mapping that our scientists did as part of this project provides significant information for land managers, since more than 500 million loblolly pine seedlings with these resistance genes are planted

every year," said Dana Nelson, the institute's project leader. "The group selected loblolly pine for sequencing because of the relatively long history of genetic research from the institute and others on the loblolly's complex traits such as disease resistance," she said.

Sonny Ramaswamy, director of USDA's National Institute of Food and Agriculture, which funded the research, noted that the loblolly pine plays an important role in American forestry.

"Now that we've unlocked its genetic secrets, loblolly pine will take on even greater importance as we look for new sources of biomass to drive our nation's bio-economy, and ways to increase carbon sequestration and mitigate climate change," Ramaswamy said.

The loblolly genome project was led by a UC Davis team, and the assembly stages were led by Johns Hopkins University and the University of Maryland. Other collaborating institutions include Indiana University, Bloomington; Texas A&M University; Children's Hospital Oakland Research Institute; and Washington State University.

More information: A. Zimin, K.A. Stevens, M. Crepeau, A. Holtz-Morris, M. Koriabine, G. Marçais, D. Puiu, M. Roberts, J.L. Wegrzyn, P.J. de Jong, D.B. Neale, S.L. Salzberg, J.A. Yorke, and C.H. Langley. Sequencing and assembly of the 22-Gb loblolly pine genome. *Genetics* March 2014 196: 875-890; doi: 10.1534/genetics.113.159715
www.genetics.org/content/196/3/875

J.L. Wegrzyn, J.D. Liechty, K.A. Stevens, L. Wu, C.A. Loopstra, H. Vasquez-Gross, W.M. Dougherty, B.Y. Lin, J.J. Zieve, P.J. Martínez-García, C. Holt, M. Yandell, A. Zimin, J.A. Yorke, M. Crepeau, D. Puiu, S.L. Salzberg, P.J. de Jong, K. Mockaitis, D. Main, C.H. Langley, and D.B. Neale. Unique Features of the Loblolly Pine (*Pinus taeda* L.) Megagenome Revealed Through Sequence Annotation. *Genetics* March

2014 196: 891-909; doi: 10.1534/genetics.113.159996

www.genetics.org/content/196/3/891

D.B. Neale, J.L. Wegrzyn, K.A. Stevens, A.V. Zimin, D. Puiu, M.W. Crepeau, C. Cardeno, M. Koriabine, A.E. Holtz-Morris, J.D. Liechty, P.J. Martínez-García, H.A. Vasquez-Gross, B.Y. Lin, J.J. Zieve, W.M. Dougherty, S. Fuentes-Soriano, L. Wu, D. Gilbert, G. Marçais, M. Roberts, C. Holt, M. Yandell, J.M. Davis, K. Smith, J.F.D. Dean, W.W. Lorenz, R.W. Whetten, R. Sederoff, N. Wheeler, P.E. McGuire, D. Main, C.A. Loopstra, K. Mockaitis, P.J. deJong, J.A. Yorke, S.L. Salzberg, and C.H. Langley Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies *Genome Biology* 2014, 15:R59 genomebiology.com/2014/15/3/R59

Provided by UC Davis

Citation: Loblolly pine's immense genome conquered (2014, March 20) retrieved 26 April 2024 from <https://phys.org/news/2014-03-loblolly-immense-genome-conquered.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.