



SCIENTIFIC METHOD / SCIENCE & EXPLORATION

Pines punched a “one way ticket toward genome obesity”

So much junk that it may push the limits of stuffing DNA into a chromosome.

by John Timmer - May 22 2013, 2:20pm EDT

LIFE SCIENCES

78



Relatives of the Norway spruce are some of the oldest living things on the planet. They haven't used all that time to tidy up their genomes, though.

 National Park Service

Last week we heard about the genome of a plant that pushed the limits of compacting its DNA: [the bladderwort](#) seems to have done away with most of the genetic material that typically makes plant and animal genomes so large without any apparent ill effects. This week, the genome of a different plant is in the spotlight: the Norway spruce (*Picea abies*), which also seems to suffer no ill effects, even though it has picked up an enormous amount of DNA. Each one of its chromosomes is nearly the size of the entire human genome—and it has a dozen of them. When researchers looked at what all that extra DNA might be doing, they came up with a simple answer: probably not anything useful.

If you're aware of the Norway spruce, it's probably because you have been shopping for a Christmas tree. But conifers (technically Gymnosperms, although the group includes ginkgoes and a few other species) are some of the most phenomenally successful organisms on Earth. They've dominated forests for over 200 million years, and members of the group include the tallest, heaviest, and oldest things currently alive. All of them seem to have managed this despite having a staggeringly inefficient genome management style.

Unlike many groups that vary widely in the number of chromosomes their species carry, pretty much all the Gymnosperms have a dozen pairs of chromosomes. And pretty much all of these chromosomes are up in the area of two billion bases long, or a bit smaller than the human genome. That size is so consistent, in fact, that the authors think the trees might be pushing up against the limits of how much stuff you can put in a chromosome and still get it copied and shared between two cells when they divide. In other words, if firs wanted to carry any more DNA than they already do, they'd have to start making new chromosomes.

From an evolutionary fitness perspective, would the plants actually want more DNA? Probably not, if the new genome is anything to go by. Despite all the extra DNA, the Norway spruce has almost exactly the same number of genes—28,354 in total—that the bladderwort does, even though the latter has about 1/250th the DNA. But it has plenty of dead copies of genes that have been inactivated by mutation. All told, these pseudogenes take up over seven times as much space in the genome as the working genes do.

However, the pseudogenes are a small contributor to the size of the genome compared to mobile genetic parasites called transposons. The transposons have hopped into all sorts of places in the genome—within the non-coding introns of genes, in between genes—and just stayed there. In fact, the Norway Spruce has an unusually high number of large introns simply because so many of them have picked up one or more transposons. Based on looking at a number of other Gymnosperms, these transposons have just been slowly accumulating throughout the group's history and have just never gone away, "possibly owing to the lack of an efficient elimination mechanism."

Inaccurate recombination between chromosomes can sometimes create deletions, which might get rid of some of the excess DNA once it's present. But the conifers don't undergo recombination very often in the areas where that DNA resides—instead, the exchange of DNA mostly happens where the genes are. All told, the authors call this a "one way ticket toward genome obesity."

Incidentally, all this stuff made sequencing the genome a nightmare. Normally, software is used to recognize when two stretches of sequence partly overlap because the sequence is identical, and it uses further overlaps to build ever-larger sequences. In this case, the frequency of transposons meant that there were nearly identical sequences scattered everywhere in the genome. Imagine trying to build a city map where every road that ran north-to-south had a name, but everything east-west was simply called "street." To cope with this, the team separated out chunks of the chromosome a few hundred thousand bases long, figured out the sequence of the chunk, and then looked for places where the chunks overlapped. This method got the job done, but there are still plenty of gaps and missing sequences.

There are a few other draft conifer genomes in the works and all of them pretty much look like this, although the exact details of which transposons are present and where they're located differ somewhat among the species. So far, the genomes only tell us a little about the origin of the features we commonly associate with trees. But they definitely tell us that a group of species don't have to be neat freaks in order to be phenomenally successful.