

Ion Torrents for Forest Trees

The Norwegian Forest and Landscape Institute has recently set up a DNA sequencing platform based on Ion Torrent technology at their lab in Ås to study diseases of forest trees (Figure 1) and to develop markers for improved characteristics (for example, wood quality and stress tolerance). Given that the genome size of Norway spruce (20 Gb) is nearly 10-fold greater than in humans, the challenges involved in this research are formidable (1).

Several projects in Norway are using Ion Torrent to sequence DNA. According to an article appearing in Nature in 2011 (2), individual sequences are attached to beads in separate wells of a multi-well semiconductor chip. Solutions containing DNA polymerase plus individual dNTPs are then sequentially injected into the chip at a frequency of approximately 20 seconds per cycle. Each time a nucleotide is incorporated into a growing DNA chain, a proton is released and this ionic change in an individual well is detected electronically. Repeated cycles with individual nucleotides plus chip washes, combined with electro-

nic detection of nucleotide incorporation in a multi-well chip generate data on large numbers of DNA sequences. A typical 2 hour run using an ion chip with 1.2 million wells (sensors) generates approximately 25 million bases (25 Mb) of DNA sequence information. Ion Torrent machines are produced by Life Technologies, Inc. USA. Their most advanced chips have a capacity for 165 million wells (3).

Adam Vivian-Smith, a researcher trained in Australia, coordinates Ion Torrent sequencing at the Norwegian Forest and Landscape Institute (Fig. 2). Since obtaining their machine in 2012, they have sequenced gigabases of DNA from various sources and have established collaborations



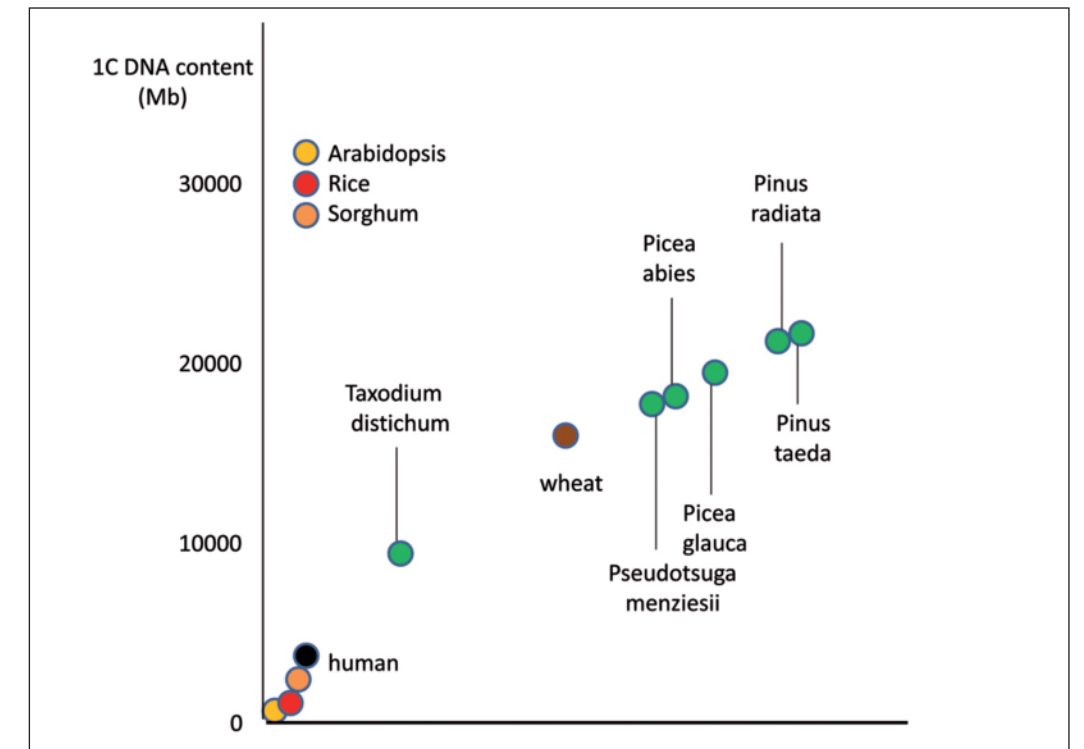
Norway spruce

with groups in South Africa, the United Kingdom (UK) and British Columbia. Vivian-Smith says that powerful sequencing technologies and access to well-characterized genetic diversity can lead to significant advances in tree science during the next 10 years.

A project in collaboration with researchers in the UK involves the die back disease of ash (*Fraxinus*) caused by species of the fungus *Hymenoscyphus*. Even though use of ashes as landscape trees in Norway is limited, the ash die back disease is a serious problem in Europe,

generally, with devastating potential. To identify determinants involved in disease causation, the research group at Ås has sequenced several isolates of both virulent and non-virulent strain of the fungus. They are also busy characterizing genome diversity in 16 different ash populations in Norway.

In contrast to ash, Norway spruce (*Picea abies*) is a major forest tree in Norway with substantial economic impact. According to Vivian-Smith, over the years the Norwegian Forest and Landscape Institute has collected spruce tree material varying in wood quality, stress tolerance etc. The group has



Genome sizes of conifers and selected angiosperms with complete- or partially-sequenced genomes. The largest genomes presented belong to *Pseudotsuga menziesii* (Douglas fir), *Picea abies* (Norway spruce, gran), *Picea glauca* (white spruce, hvit gran), *Pinus radiata* (Monterey pine, Momtereyfur), *Pinus taeda* (Loblolly pine, virakfuru) and *Pinus pinaster* (Maritime pine, strandfuru). *Taxodium distichum* is Bald Cypress, *Sumpsypruss*. The smallest genome in the Figure is *Arabidopsis* with a 0.135 Gb genome size. Adapted from Mackay et al 2012.

already generated considerable sequence information on these trees.

Internationally, there are several groups involved in sequencing the Norway spruce genome. A draft sequence was published this year in Nature (4, 5). To give an idea of the magnitude of the challenge involved, this article reported that the genome size in Norway spruce is 19.6 Gb with 70% of the total sequence involving high-copy repeats, complicating assembly of individual chromosome double strands ($n = 12$) which are 1.6 Gb average length. Not surprisingly, the 2013 Nature paper could only assemble 4.3 Gb in sequences >10 kb with the longest continuous sequence being <20 kb.

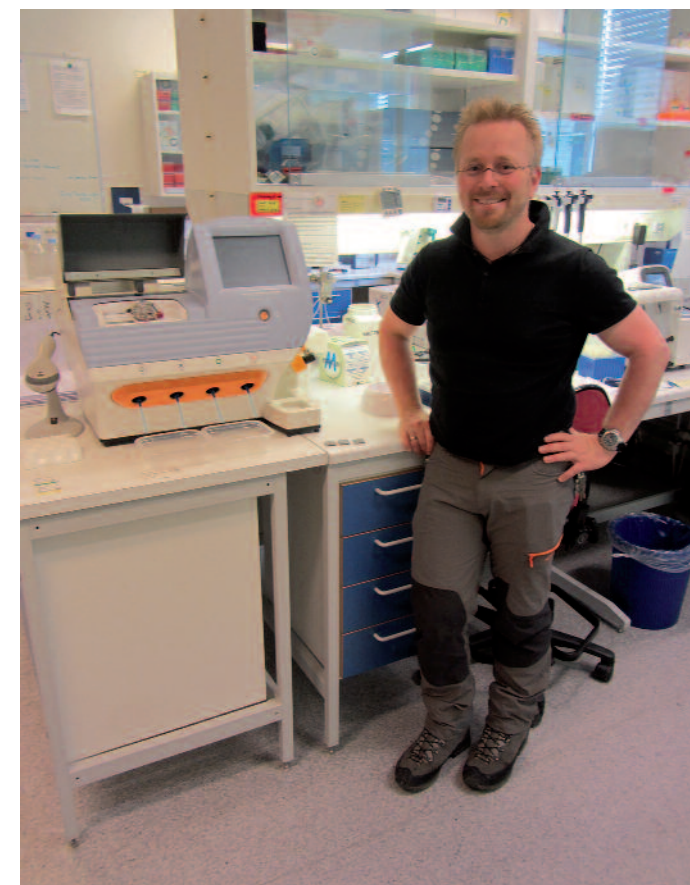
When asked by NBS-Nytt about the future of DNA sequencing of forest trees at Ås, Adam Vivian-Smith said, 'Having completed nearly 2 years of high-throughput sequence so far, we're looking forward to good progress in the years to come.'

References

1. Mackay J et al.: Towards decoding the conifer giga-genome. *Plant Mol Biol* 80 (2012) 555-569.
2. Rothberg JM et al.: An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475 (2011) 348-352.
3. Merriman B et al.: Progress in Ion Torrent semiconductor chip based sequencing. *Electrophoresis* 33 (2012) 3397-3417.
4. Sederoff R: A spruce sequence. *Nature* 497 (2013) 569-570.
5. Nystedt B: The Norway spruce genome sequence and conifer genome evolution. *Nature* 497 (2013) 579-584.



FORFATTER: JOHN EINSET, UNIVERSITET FOR MILJØ- OG BIOVITENSKAP, ÅS.
E-post: john.einset@umb.no



Adam Vivian-Smith