

TEACHER'S GUIDE

DNA to Darwin Case Study

Who's the daddy?

Polyploidy in plant evolution

Version 1.1

James Clarkson

Royal Botanic Gardens, Kew

Dean Madden (Ed.)

NCBE, University of Reading



Charles Darwin

Polyploidy in plant evolution

The evolutionary biologist Sean Carroll has called DNA the 'forensic evidence' for evolution. That's certainly true here, because without DNA sequence data it would be difficult if not impossible to unravel the relationships between the species of tobacco plants that form the basis of this *Case Study*.

Students use data from both the nuclear and chloroplast genomes of *Nicotiana* species to determine 'Who's the daddy' (and 'Who's the Mummy' too). In so doing, they are introduced to the concept of polyploidy, which is of great importance in plant speciation and in agriculture.

Polyploidy in speciation and agriculture

Some estimates suggest that between 30 and 80% of living plant species are polyploid: that is, they have multiple sets of chromosomes. Polyploidy plays a vital role in the evolution of new plant species and the creation of novel varieties in agriculture.

Polyploid plants can arise spontaneously in nature. For example, ancient polyploidy events coincided with increases in the number of species of flowering plants and ferns. Because many lineages show evidence of ancient polyploidy (paleopolyploidy), it might be more accurate to think of most plants as polyploids with some having undergone additional rounds of polyploidy more recently.

Most polyploids exhibit 'hybrid vigour' (heterosis) relative to their parental species, and they may display novel variations that contribute to the processes of speciation. The causes of novel variation in polyploids include gene dosage effects (resulting from numerous copies of genome), alteration of gene regulation and chromosomal rearrangements, all of which may alter gene content and/or expression levels. These rapid changes may contribute to reproductive isolation and speciation.

The four most widely-cultivated crop plants in the world (wheat, rice, soybeans and maize) are polyploids. In fact, 15 of the 21 most important crops, as measured by area of cultivation, are polyploids.

Polyploids can occur by several mechanisms, including meiotic or mitotic failures and fusion of unreduced ($n=2$) gametes. Both *autopolyploids* (where all sets of genes come from one species) and *allopolyploids* (where the chromosomes come from two species) can be found among both wild and domesticated plant species. This *Case Study* examines the formation of two polyploid tobacco species from diploid relatives.

Outline of the activity

There are four main parts to the activity:

- A document containing the DNA sequences from five diploid *Nicotiana* species and one species to be used as an 'outgroup' is provided but the two polyploid sequences need to be added. An online DNA sequence database, known as *GenBank*, is used as a source for the sequences. Alternatively, if internet access is not available, the files can be given to students;
- Once all the sequences are loaded into *Geneious*, an alignment needs to be performed so that their homologous regions can be compared;
- The aligned sequence data is used to build phylogenetic trees using the neighbour-joining method;
- Trees need to be interpreted to find the chloroplast donors of these two polyploids. The diploid sister species of the two allotetraploids are the maternal progenitors. Examining the evolutionary distances (the lengths of the branches) gives an indication of the age of polyploids.

General reading

Botany: A blooming history (2011) BBC4 television series. Three 60-minute programmes. Occasionally available on BBC iPlayer in the UK.

Episode 3, 'Hidden world', gives examples of the development of new varieties of crops in plant breeding, including the hybrid wheat varieties in the 1960's 'Green revolution'.

Reading the story in DNA: A beginner's guide to molecular evolution by Lindell Bromham (2008) Oxford University Press (Paperback) ISBN: 978 0199290918. An engaging textbook on molecular evolution, which assumes no specialist mathematical knowledge and takes the reader from first principles. Although it's aimed at undergraduates, this superb book contains sufficient detail for PhD students, yet parts will appeal equally to 16–19 year-olds.

River out of Eden: A Darwinian view of life by Richard Dawkins (2001) Phoenix. ISBN: 978 1857994056. The 'Mitochondrial Eve' story was popularised by Richard Dawkins this book; in the book, Dawkins describes how mitochondrial evolution was used to pinpoint the maternal ancestor of all living humans.

The seven daughters of Eve by Bryan Sykes (2004) Corgi. ISBN: 978 0552152181. Another popular book that explains how mitochondrial DNA was used to trace human origins.

Scientific publications

Both of these publications require payment to access them. They are concerned with another 'model organism', *Arabidopsis thaliana*.

Bowers, J.E. *et al* (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature*, **422** (6930) 433–438. doi: 10.1038/nature01521

Vision, T.J. *et al* (2000) The origins of genomic duplications in *Arabidopsis*. *Science*, **290** (5499) 2114–2117. doi: 10.1126/science.290.5499.2114

Requirements

Software

The software required, *Geneious*, can be downloaded free-of-charge from: www.geneious.com. The software is available for Windows, Macintosh and Linux operating systems. Only the free, 'basic' version of the software is required for this activity.

DNA sequence data

Students will need the *Geneious* document containing DNA sequences from diploid *Nicotiana* species: **Nicotiana.geneious**. Students will also require the DNA sequences from the two tetraploid species, *N. clevelandii* and *N. tabacum*, which can either be downloaded from *GenBank* within the *Geneious* software, or given to the students. The *GenBank* Accession Codes are AJ585925 and AJ585931, respectively.

Students' worksheets

Students will require copies of Student's Guide, pages 2–13, a paper print-out of the results on page 7 of this document, and a ruler.

Presentations

Teachers may find the *QuickTime* animations in the *Introductory Activities*, which are in a separate document, useful here. *QuickTime* may be downloaded free-of-charge from the Apple web site: www.apple.com/quicktime. The *PowerPoint* or *Keynote* presentations should be helpful for introducing this exercise.

Educational aims

The activity reinforces students' understanding of DNA structure. It introduces the principle of alignment, generating an evolutionary tree from molecular data and using branch lengths to date such trees.

It provides an introduction to the topic of polyploidy, its evolutionary and agricultural significance, and the use of 'model organisms' in biology.

Prerequisite knowledge

Students will need to understand the structure of DNA (*i.e.*, that DNA is comprised of nucleotides). They will need to know that chloroplasts have their own DNA and to be taught the principle of sequence alignment (the *Introductory Activities*, which are in a separate document, will be useful.)

It would be helpful, although not essential, if they were familiar with the concept of polyploidy and the basic means by which it arises.

Answers to the questions on the worksheets

Page 6

- a. *Symonanthus aromaticus* has been used as the outgroup. This Australian genus is related to *Nicotiana*.
- b. The diploid *Nicotiana* species are: *N. attenuata*, *N. undulata*, *N. sylvestris*, *N. obtusifolia* and *N. tomentosiformis* (n=12).
- c. *N. clevelandii* and *N. tabacum* are allotetraploids (n=24).
- d. There are two groups (clades) on the tree. Each of the tetraploid species occurs once in the two groups, clustered with one of the haploid species.
- e. The two tetraploid species have two different copies of the gene, one from each 'parent' species.
- f. i. The 'parents' of *N. clevelandii* are *N. attenuata* and *N. obtusifolia*.
ii. The 'parents' of *N. tabacum* are *N. sylvestris* and *N. tomentosiformis*.

Page 13

- g. *N. clevelandii* is grouped with *N. obtusifolia*; *N. tabacum* is grouped with *N. sylvestris*.
- h. See table below:

Tetraploid species	Male parent	Female parent
<i>N. clevelandii</i>	<i>N. attenuata</i>	<i>N. obtusifolia</i>
<i>N. tabacum</i>	<i>N. tomentosiformis</i>	<i>N. sylvestris</i>

- i. The actual lengths measured will vary according to the size of the print-out, but it should be found that the branch leading to *N. clevelandii* is three times longer than that leading to *N. tabacum*. If the 'Molecular clock' hypothesis is correct in this context (i.e., that base substitutions occur at a steady rate over evolutionary time), then *N. tabacum* appears to have evolved more recently and is a third of the age of *N. clevelandii*.

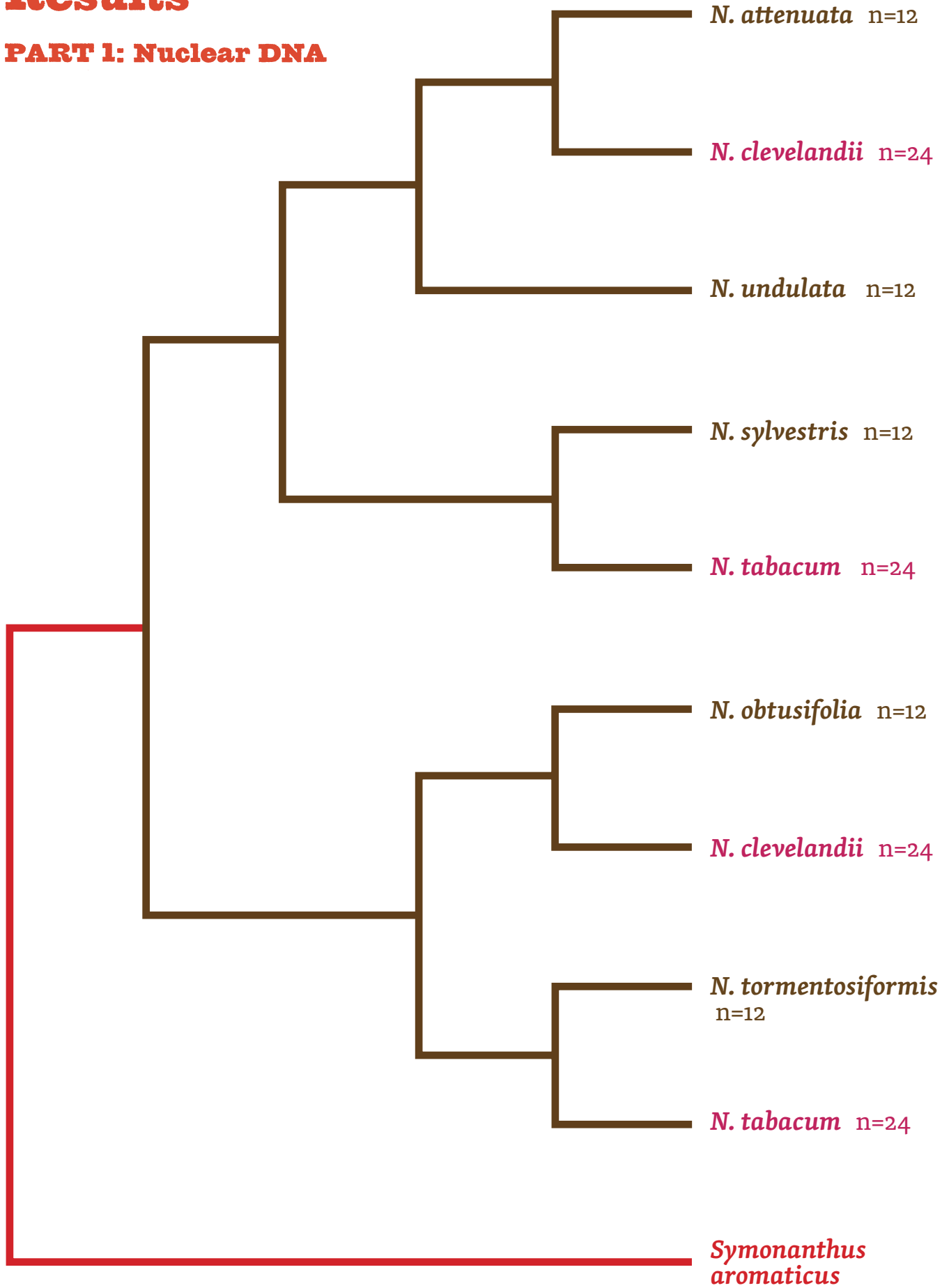


PHOTO FROM: James Clarkson.

The formation of *N. tabacum* from *N. sylvestris* x *N. tomentosiformis*. Notice that the flowers of the allotetraploid are approximately intermediate between its two progenitor diploid types.

Results

PART I: Nuclear DNA



Results

PART 2: Chloroplast DNA

