# **First draft: EEOB563 Final project - Evolution of cultivated wheat**

Ann Murithi

GIT Files

EEOB\_Final Project\_first draft

Chloroplast \_neighbor

Chloroplast\_ml.raxml.bestTree.

**Introduction**

Cultivated wheat belongs to the genus Triticum L. which includes cultivated and wild species. Triticum is made up of six species; *Tritucum monococcum* (AA genome), *Triticum urartu* (AA genome), *Triticum Turgidum* (AABB genome), *Triticum timopheevii* (AAGG genome)*, Triticum aestivum* (AABBDD genome) and *Triticum zhukovskyi* (AAAAGG genome). The species are further grouped into those belonging to diploid species (monococcon), tetraploid (Dicoccoidea) and Triticum (consisting of hexaploid species). The hexaploid arose under cultivation after the domestication of diploid and tetraploid species in the last 10,000 years. *T. aestuvim* arose from the hybridization between cultivated *T. turgidum* and diploid goatgrass Aegilop tauschii with DD genome, while *T. zhukovskyi* originated from hybridization of *T. monococcum* a diploid with *T.timopheevii*. The two hexaploid make up two lineages of polyploid wheat; one, Emmer lineage that consists of *T. aestivum* and *T. turgidum, while T.timopheevii* and *T. zhukovskyi* make up the Timopheevii lineage.

*Triticum urartu* with AA genome is believed to be the male parent contributing the A genome in both lineages, while *Aegilops* is the female donor believed to have contributed the remaining two genomes of the hexaploid genomes. From the work of several Japanese wheat geneticists, *Aegilops* were divided into three major genomic groups, C, D, and S. The C-genome group included two species; the D-genome group included four species; and the S-genome group consisted of three species of the*Sitopsis*section*: Ae. longissima (including Ae.sharonensis), Ae. bicornis (and Ae. speltoides Tausch.* Within the S-genome, current taxonomy recognizes five diploid species carrying the S-genome: Ae. speltoides including ssp. ligustica (Savign.) Fiori (SS) and ssp. speltoides Boiss., Ae. bicornis (SbSb), Ae. searsii (SsSs), Ae. sharonensis (SshSsh), and Ae. longissima (SlSl).

The knowledge of the sources of the genomic constitution of wheat is crucial to wheat improvement. This is mostly due to the ability of wheat genome to pair either within across genome of distant relatives, creating a wide genetic pool for sources of genetic variation for agronomic important traits such as pest and disease resistant and grain quality. Although many agronomically useful genes have already been transferred from *Aegilops* to common wheat varieties or breeding lines, their genetic potential in broadening genetic diversity of wheat is not fully exploited. Utilization of gene pool of *Aegilops* requires good knowledge of genetics and genomics of these species, including their genome and distribution of their genomes across the two lineages of Triticum.

Despite the wide knowledge of genome organization of the *Aegilops* debates over the origin of the B genome and therefore the cytoplasm of *T. turgidum* have spanned over decades with several hypotheses of the origin proposed. In one hypothesis, *Aegilops* is proposed as a possible donor, in which B and G genomes could have been derived from different genotypes of *Aegilops*. This is possible due to its diverse plasmon and outcrossing nature of *Aegilops.* Although the second hypothesis was contradicted by molecular and morphological data, it postulates the origin of the B genome to be in the *Sitopsis* section of *Aegilops*. And lastly, it is also possible that the donor of B genome could be extinct or has yet to be collected. In a bid to decipher the source of the female genome, this paper extracted and utilized the chloroplast genome of 20 genotypes to construct a phylogenetic tree to show the divergence of the Emmer and Timopheevii lineage and the sources of each genome that define these species.

**METHOD**

**Material**

To replicate the work that had been done, I data of the materials used in this analysis from Genbank using the accession numbers provided in the paper.

Table 1: Description of materials making up the taxa for the phylogenetic analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Organisms - Species** | **Cultivar** | **Gene Bank accession** | **Common name** | **Ploidy** | **Lineage** |
| Triticum aestivum | Chinese Spring | KJ614396.1 | Bread wheat | Hexaploid (AABBDD) | Emmer |
| Triticum aestivum | spleta | KJ614403.1 | Bread wheat | Hexaploid (AABBDD) | Emmer |
| Triticum turgidum | TA2836 | KJ614397.1 | ssp. carthlicum | Tetraploid (AABB genome) | Emmer |
| Triticum turgidum | TA2801 | KJ614399.1 | ssp. carthlicum | Tetraploid (AABB genome) | Emmer |
| Triticum turgidum | PI520121 | KJ614398.1 | ssp. durum | Tetraploid (AABB genome) | Emmer |
| Triticum turgidum | TA0073 | KJ614400.1 | ssp. dicoccoides | Tetraploid (AABB genome) | Emmer |
| Triticum turgidum | TA0060 | KJ614401.1 | ssp. dicoccoides | Tetraploid (AABB genome) | Emmer |
| Aegilops speltoides | AE918 | KJ614404.1 | ssp. ligustica | Tetraploid (AABB genome) |  |
| Aegilops speltoides | PI487232 | KJ614406.1 | ssp. ligustica | Diploid (DD genome) |  |
| Aegilops speltoides | TA1796 | KJ614405.1 | ssp. ligustica | Diploid (DD genome) |  |
| Triticum timopheevii | TA0941 | KJ614407.1 | ssp. armeniacum | Tetraploid (AAGG genome) | Timopheevii |
| Triticum timopheevii | TA944 | KJ614409.1 | ssp. armeniacum | Tetraploid (AAGG genome) | Timopheevii |
| Triticum timopheevii | TA1485 | KJ614408.1 | ssp. armeniacum | Tetraploid (AAGG genome) | Timopheevii |
| Aegilops bicornis | Clae57 | KJ614418.1 | Goat grass | Diplod (SbSb Genome) |  |
| Aegilops searsii | TA1926 | KJ614413.1 | Goat grass | Diplod (SsSs Genome) |  |
| Aegilops searsii | TA1837 | KJ614414.1 | Goat grass | Diplod (SsSs Genome) |  |
| Aegilops searsii | TA1841 | KJ614415.1 | Goat grass | Diplod (SsSs Genome) |  |
| Aegilops sharonensis | TA1995 | KJ614419.1 | Goat grass | Diplod (SshSsh genome) |  |
| Aegilops sharonensis | TA1996 | KJ614417.1 | Goat grass | Diplod (SshSsh genome) |  |
| Aegilops longissima | TA1924 | KJ614416.1 | Goat grass | Diplod (SlSl genome) |  |
| Aegilops kotschyi | TA1980 | KJ614420.1 | Goat grass |  |  |
| Triticum urartu | PI428335 | KJ614411.1 |  | Diploid (AA genome) |  |
| Aegilops tauschii | AL8/78 | KJ614412.1 | Goat grass | Diplod (DD genome) |  |
| Hordeum vulgare | Morex | EF115541.1 | Barley |  |  |
| Triticum Zhukovskyi\*\* |  |  |  | Hexaploid (AAAAGG) | Timopheevii |

**2. Sequence alignment and phylogeny analysis**

Nucleotide sequence of whole chloroplast genome were extracted from Genbank in FASTA format to note pad. The sequences were aligned using MAFFT program installed in HPC class. The Phylogenetic analysis was performed using Phylip. Neighbor joining was done based on Jukes-Cantor distance substitution. Bootstrap values were calculated using default setting, at 100 replicates. A majority rule maximum likelihood tree in RAxML using GTR-G model was also generated.

I will also build a majority rule tree after partitioning the nucleotides based on codon. Finally, I will estimate the best nucleotide substitution model using bayesian information criterion in Mr. Bayes.

**Results**

**References**

[Piotr Gornicki](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Gornicki%2C+Piotr), [Huilan Zhu](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Zhu%2C+Huilan), [Junwei Wang](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Wang%2C+Junwei), [Ghana S. Challa](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Challa%2C+Ghana+S), [Zhengzhi Zhang](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Zhang%2C+Zhengzhi), [Bikram S. Gill](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Gill%2C+Bikram+S), [Wanlong Li](https://nph.onlinelibrary.wiley.com/action/doSearch?ContribAuthorStored=Li%2C+Wanlong), The chloroplast view of the evolution of polyploid wheat (2014)***,*** *New Phytologist, 204(3), 74-714*