The genomics of peanut and its wild relatives.

David J. Bertioli[1], Bruna Vidigal[1,2], Soraya C.M. Leal-Bertioli[2], Lutz Froenicke[3], Christopher Beitel[3], Stephan Nielen[2,4], Millind Ratnaparkhe[5], Ana C. Brasileiro[2], Patricia M. Guimarães[2], Trude Schwarzacher[6], Richard Michelmore[3], Andrew H. Paterson[5], Pat Heslop-Harison[6], Ana C.G. Araujo[2].

[1]University of Brasilia, Department of Genetics, Campus Universitário, 70910-900 Brasília DF, Brazil.

[2]Embrapa Genetic Resources and Biotechnology, Brasilia, DF, Brazil.

[3]UC Davis Genome Center, Davis, CA, USA

[4]Now at: Plant Breeding and Genetics Section, Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, Vienna, Austria.

[5]Plant Genome Mapping Laboratory, The University of Georgia, Athens 30605, USA.

[6]Department of Biology, University of Leicester, Leicester, UK.

Cultivated peanut (*A. hypogaea* L. is a very important tropical oilseed, valued both for subsistence and as a cash crop. It is an allotetraploid of recent origin with an AB type genome (2n = 4x = 40) and has very low DNA polymorphism, a characteristic that has hampered genetic studies. The two sub-genomes diverged from a common ancestor about three and a half million years ago, more recently than the sub-genomes of cotton or soya. The genome of peanut is estimated at about 2.8 Gb (only slightly smaller than the human genome) and has a high repetitive DNA content. During meiosis, peanut chromosomes pairing is almost entirely bivalent, an indication of genetic divergence of the A and B chromosomes. A significant divergence is also indicated by *in situ* hybridization analyses, which show that the repetitive contents of the A and B genomes are substantially different. Interestingly, evidence regarding the low copy fraction of the genome is converse: strong colinearity is observed between the A and B genome genetic maps, and indeed in many genome regions, gene order appears to have changed little over the last 55 million years. These diverging lines of evidence point to an intriguing question in the evolution of genome structure: how can the evolutionary flux of the structurally predominant repetitive DNA be reconciled with evolutionary conservation in low copy number DNA? Analysis of genomic sequences from the A and B genomes indicated that a large proportion of the sequence space is accounted for by relatively few transposons that have been active since the evolutionary divergence of the two subgenomes. The activity of these transposons has been a very significant driver of the erosion of A-B genome sequence similarity. Notably, this erosion is not evenly distributed; it is concentrated in intergenic regions.