

REPORT: Isolation of Peanut Resistance Gene Analogs, and Evaluation of their Role in Disease and Insect Resistance

Principal Investigator: Andrew H. Paterson, Professor & Director, Plant Genome Mapping Laboratory, Univ. Georgia, 111 Riverbend Rd Rm. 228, Athens, GA 30602. Tel. 1-706-583-0162; FAX 1-706-583-0160; Email paterson@uga.edu.

Objectives: We will use 'consensus sequences' from genes that confer resistance to key viral, bacterial, and fungal pathogens and insects in other plants, to isolate corresponding gene families in peanut. The peanut gene families will be placed on the integrated genetic-physical map toward determining their specific role(s) in controlling peanut susceptibility to diseases and insects, and prioritizing them for transformation.

Rationale and Economic Significance: Disease and pest management is an important challenge to peanut production. Development of superior cultivars with intrinsic genetic resistance can contribute directly to assisting the peanut industry and growers by decreasing costs of production per ton. The identification and manipulation of endogenous peanut genes to improve resistance enables one to use genes that are known to function well in the target crop, and does not suffer from the stigmas attached to 'genetically-modified organisms.'

Due to its large amount of DNA, low level of variation, and complicated life cycle, isolation of peanut genes conferring disease and pest resistance has lagged behind other more favorable models. It has been established for several years from research in botanical models and other crops that many disease and pest resistance genes in diverse plants share common features at the DNA level, and our preliminary studies prove that this is also true in peanut. We will take advantage of these common features to isolate large numbers of disease and pest resistance genes from peanut, and determine their locations and copy numbers in the genome. This provides a foundation for rapid screening to identify peanut genes that are responsible for response to both existing and new biotic stresses (diseases and pests).

Progress: In year 1 of the funding period (2004), we cloned and sequenced a total of 1028 amplicons generated from *Arachis hypogaea* genomic DNA with six different primer sets (Table 1, below), corresponding to the six families of host plant resistance genes known. A subset of 234 (22.7%) representing 5 of the 6 gene classes, contain open reading frames uninterrupted by a 'stop' codon, meaning that they may potentially encode a gene product (Table 2) thus may be true resistance genes [the remainder are probably 'pseudogenes', nonfunctional versions of gene-like junk DNA sequences that are common in plant genomes].

Further analysis in year 2 (2005) concentrated on these 234 potentially functional genes, which appear to cluster (based on sequence similarities) into 47 different 'gene families'. A summary of relationships among the gene families, and also their relationships to resistance genes known from other organism, is shown (Figure 1). Next, we determined the physical distributions of these genes/families with respect to one another and to the peanut chromosomes. Specifically, we applied synthetic DNA probes specific to each gene family, to a 'BAC' (bacterial artificial chromosome) library that represents each nucleotide in the peanut genome and average of six times. We identified a total of 1,028 BACs. The BACs were then further screened to remove false positives, and 'fingerprinted' (to determine which were from the same places in the genome and which were from different places). In total, we have estimated that the peanut genome contains about 250 loci that each contain one or more genes with features known to be diagnostic of disease-resistance genes.

Among these estimated 250 loci, some showed much stronger evidence of association with disease resistance than others. In particular, two peanut BACs each hybridized to 11 different hybridization probes, representing at least four clades ('super-families') of resistance genes. **These unusual regions of the genome will be the focal point of our continuation proposal.**

Table 1 The primer sequences used for the amplification of peanut resistance gene analogs (RGA)

Primer Id	Motif	Motif sequence	Primer sequence (5'-3')
PLTR-fwd ^a	P-loop	GMGGVGKTT	GGNATGGGNGTNGGNAARACNACN
PLTR-rev ^a	GLPL	GLPLALKVLG	NCANCARAANGGNTGNGGNGGGTANGG
PNTR-fwd ^b	P-loop	GGVGKTT	GGNGGNGTNGGNAANACNAC
PNTR-rev ^b	RNBS-D	CFLYCALFP	CGRAANARNSHRCARTANVNRAARC
PCRE-fwd ^c	Kinase-2	LILDDVW	TGATACTGGATGATGTCTGG
PCRE-rev ^c	EGF	EGFIRNT	GTGCTTCTTATGAACCCTTC
PPTO-fwd ^d	<i>Pto</i> kinase 1		GCATTGGAACAAGGTGAA
PPTO-rev ^d	<i>Pto</i> kinase 2		AGGGGGACCACCACGTAG
PCf-fwd ^e	LRR	SNKLGPI	WSNAAYAARYTNCA YGGNCCNAT
PCf-rev ^e	LRR	GEIPQQLA	GCNARYTGTCKNGGNATYTCNCC
PRGA-fwd ^f	P-loop	GVGKTT	GGNGGNGTNGGNAANACNAC
PRGA-rev ^f	GLPL	GLPLAL	ARNGCTARNGGNARNCC

^{a, b, c, d, e, f} Originally suggested by (Aarts et al. 1998), (Penuela et al. 2002), (Lagudah et al. 1998), (Leung et al. 1998), (Ohmori et al. 1998), and (Kanazin et al. 1996) respectively

Degenerate code: N=A, G, C or T, R=A or G, H=A, C or T, S=C or G, V=A, C or G W=A or T, Y=C or T, K=G or T

Table 2 *Arachis hypogaea* amplicons amplified with degenerate primers from conserved domains of several classes of cloned resistance genes

Sequence Id ^a	No. of clones sequenced	No. of RGA ^c matches	No. of RGAS with ORF	Non-RGA	No match
PLTR	179	138 (77%)	95	6	35
PNTR	189	80 (45%)	37	18	91
PRGA ^b	171	129 (75%)	77	10	32
PCf	129	6 (5%)	1	107	16
PCRE	169	78 (46%)	20	18	73
PPTO	191	0 (0%)	0	80	107
Total	1028		234		

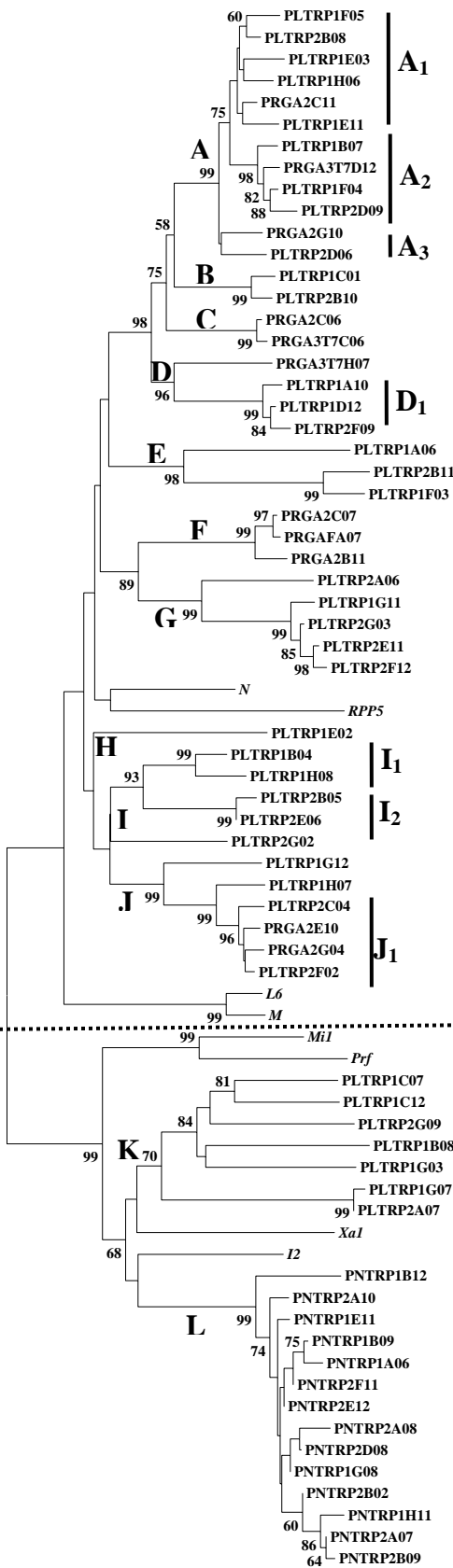
^a The sequence Ids are same as the primer Ids used in table 1

^b J. Ballester and A. H. Paterson, unpublished data

^c The minimum E-value was 1×10^{-7} in blasting against the plant protein database

Next page:

Figure 1: Phyletic relationships between peanut and other plant resistance genes. The tree is based on CLUSTAL_X alignment. The confidence of the nodes was tested by bootstrapping of 1000 replicates; bootstrap test results varied between 65% and 100%. The branch lengths are proportional to the average number of amino acid substitutions per site as indicated by the scale. TIR class genes are indicated in bold characters, and the species origin for each sequence is denoted in parenthesis.



<u>Overgos</u>	<u>Clade</u>	<u>Shared</u>	<u>Unique(%)</u>
PRGOV44 ¹	A1	64	15 (19)
PRGOV52	A1&A2	48	0 (0)
PRGOV45	A2	27	64 (70)
PRGOV49	A3	20	1 (5)
PRGOV01 ²	B	14	29 (67)
PRGOV07 ³	PRGA3T7H07	53	14 (21)
PRGOV04 ⁴	D1	28	9 (24)
PRGOV12	PLTRP1A06	0	10 (100)
PRGOV14 ⁵	E1	15	56 (79)
PRGOV36 ⁶	F	58	103 (64)
PRGOV34 ⁷	G	7	8 (53)
PRGOV16 ⁸	PLTRP1E02	5	35 (88)
PRGOV19 ⁹	I1	15	76 (84)
PRGOV22 ¹⁰	I2	0	18 (100)
PRGOV50	PLTRP2G02	3	0 (0)
PRGOV25 ¹¹	PLTRP1G12	11	12 (52)
PRGOV28 ¹²	PLTRP1H07	25	10 (29)
PRGOV31 ¹³	J1	24	28 (54)

PRGOV40	PLTRP1C07	1	13 (93)
PRGOV39	PLTRP2G09	1	35 (97)
PRGOV53	PLTRP1G07	0	6 (100)
PRGOV41 ¹⁴		8	11 (58)

TIR

NonTIR

0.2