Cooperation Centre for Scientific Research Relative to Tobacco

Agro-Phyto Information Collection Task Force

CORESTA Guide N° 16

Molecular Markers in Tobacco

March 2019

Authors:
Shengming Yang, Dandan Li, Anne Jack Fisher
University of Kentucky, U.S.A.
Sreepriya Pramod, Dongmei Xu
Altria Client Services, U.S.A.

CORESTA TECHNICAL GUIDE N° 16

Title: Molecular Markers in Tobacco

Status: Valid

Note: This document will be periodically reviewed by CORESTA

Document history:

Date of review	Information
March 2019	Version 1

TABLE OF CONTENTS

1. INTRODUCTION

Conventional plant breeding usually identifies genetic variation by visual or chemical selection. Certain important traits such as nutritional value, alkaloid levels and fatty acid composition, which lead to very subtle changes in morphological appearance, are always challenges for plant breeders. The process of developing new crop varieties can take up to 25 years. However, with the advancements in molecular biology, the duration has been considerably shortened to 7-10 years. One of the important techniques that makes it efficient for scientists to select plant traits is Marker-Assisted Selection (MAS).

Molecular technologies enable us to identify changes in DNA or gene sequences associated with the phenotype variation. Various types of molecular markers have been used to label and map these molecular changes through linkage analysis with segregating populations. Genetic linkage describes the fact that molecular markers and genes of interest proximal to each other are inherited together during meiosis. A desirable attribute for a user-friendly marker is close linkage with the gene corresponding to phenotype, which significantly precludes false-positive selection due to recombination (or cross-over) events. Other qualities expected by breeders for genetic markers are: (1) co-dominance; (2) representing high level of genetic polymorphism (can be easily discerned by users); (3) low cost during application; (4) high stability (able to be confirmed and duplicated by colleague researchers), and (5) ease of use.

In this review, we have surveyed the most commonly used molecular techniques in tobacco. These include RAPD (Random Amplified Polymorphic DNA), AFLP (Amplified Fragment Length Polymorphism), SSR (Simple Sequence Repeat), SCAR (Sequence Characterized Amplified Region), CAPS (Cleaved Amplified Polymorphic Sequences), dCAPS (derived CAPS), and KASP™ (Kompetitive Allele Specific PCR). The different types of molecular markers are explained in detail in the forthcoming sections. Further, the traits for which markers for tobacco are available are outlined in Table 1, and detailed information regarding primers and/or sequences are provided in supplementary Table S1. Quantitative Trait Loci (QTLs) and the markers used to map the QTL are summarized in Table 2, and detailed information is provided in supplementary Table S2.

2. TYPES OF MOLECULAR MARKERS

2.1 RFLP Markers

Restriction Fragment Length Polymorphism (RFLP) markers were first used to distinguish two ecotypes and to construct a genetic map in *Arabidopsis* (Chang et al. 1988); (Liu et al. 1996). Even though RFLP markers are co-dominant, meaning that distinct patterns are obtained for plants that are homozygous or heterozygous for the target alleles, the main disadvantage posed for the routine application of RFLP markers in plant breeding is that it requires relatively large amounts of DNA, which is subjected to digestion of restriction enzymes and southern blotting. RFLP markers are no longer commonly used, having been replaced by the more user-friendly PCR-based DNA markers.

2.2 PCR-based DNA Markers

PCR-based DNA markers are the most widely used because they need only a small amount of DNA sample (about 10-20 ng per reaction) and no time-consuming DNA blotting or hybridization is required. PCR-based DNA markers are grouped into two classes based on the type of information they provide; anonymous markers and Sequence-Tagged Site (STS) markers. The most commonly used anonymous markers include Amplified Fragment Length

Polymorphism (AFLP) markers (Vos et al. 1995) and Random Amplified Polymorphic DNA (RAPD) markers. STS markers include Sequence Characterized Amplified Region (SCAR) markers (Paran and Michelmore 1993), Simple Sequence Repeat (SSR) markers (Litt and Luty 1989) and CAPS and dCAPS markers based on Single Nucleotide Polymorphisms (SNPs) (Konieczny and Ausubel 1993; Neff et al. 1998).

2.2.1 Anonymous Markers

2.2.1.1 Amplified Fragment Length Polymorphism

The AFLP technique involves digestion of genomic DNA or cDNA with restriction enzymes, followed by ligation of adaptors to digested fragments. A subset of fragments is selected to be amplified under high stringency conditions. Hence, this technology combines the power of RFLP with the flexibility of PCR-based markers and provides a universal, multi-locus marker tool that can be applied to complex genomes from any source. In addition, based on the use of endonucleases, there are several modified AFLP techniques, such as single endonuclease AFLP (Boumedine and Rodolakis 1998), three endonuclease-AFLP (van Der Wurff et al. 2000) and second digestion AFLP (Knox and Ellis 2001).

Radio-active detection has been replaced with silver staining, fluorescent tag or agarose gels for single endonuclease AFLP. This method is remarkably reliable, consistent, and as a result, it has been used to specifically address genetic variation, phylogenetic relationships, homoplasy, etc.

2.2.1.2 Random Amplified Polymorphic DNA

RAPD technology exploits arbitrary primers of short synthetic oligonucleotides to amplify fragments. It is simple, time-saving, and needs only small amount of template DNA without the requirement of the sequencing. Due to these advantages, RAPD technology has been widely used in localization of target genes, genetic mapping and evolution genetics. However, RAPD markers are inherently unstable, which limits their wide application among laboratories.

2.2.2 Sequence Tagged Site markers

2.2.2.1 Sequence Characterized Amplified Region markers

To overcome the limitations associated with anonymous markers, RAPD and AFLP markers can be converted into SCAR markers (Paran and Michelmore, 1993). In general, the target bands are extracted from agarose gels, followed by ligation into sequencing vectors to obtain the full sequence of the insert. Eventually, primers are designed to specifically amplify the band of interest based on its sequence. Through this process, the RAPD and AFLP markers can be transformed into simple and robust PCR markers. This procedure also improves the reproducibility of anonymous markers and avoids the occurrence of non-homologous markers with equal molecular weight. The cloned and sequenced DNA fragments can then be used for the development of dominant markers, or co-dominant CAPS markers.

2.2.2.2 Simple Sequence Repeats

SSRs are tandem repeated sequences of one to six nucleotide long DNA motifs that occur frequently in all eukaryotic and prokaryotic genomes. SSR markers can be developed directly from genomic DNA libraries or from libraries enriched for specific microsatellites (Zane et al. 2002). SSR markers are characterized by high variability, reproducibility, co-dominance, locus specificity, amenability to automation and random

dispersion throughout most genomes; therefore SSR markers can be easily and reproducibly detected by PCR. The disadvantage of SSR markers is that they require a labor- and cost-intensive development process particularly when screening and sequencing genomic DNA libraries enriched by repeated motifs. Interestingly, once they have been developed for a plant species, SSR markers are almost always the first choice for genetic mapping because of low start-up costs.

2.2.2.3 Single Nucleotide Polymorphisms

Single Nucleotide Polymorphisms (SNPs) have emerged as the most abundant and universal form of genomic polymorphism. In breeding applications, utilization of SNPs is still in its infancy, mainly because other marker-related applications do not require the massively parallel, genome-wide genotyping necessary for SNPs. The availability of draft genome sequences of the tobacco varieties TN90 (Burley), K326 (flue-cured) and Basma Xanthi (Oriental) (Sierro et al. 2014), followed by optical mapping, enabled an improved assembly of the K326 genome (Edwards et al. 2017), which has laid the foundation for gene mapping and trait discovery in tobacco. The release of the SNP-based high density genetic map, *N. tabacum* 30k Infinium HD consensus map 2015 (Edwards et al. 2017) also provides the tobacco genetic research community with resources to fine map their trait of interest (https://solgenomics.net/cview/map.pl?map_version_id=178). The database can also be used to convert some of the previously used SSR, AFLP, SCAR markers etc., to closely linked SNP markers from the genetic map.

The use of SNPs for genotyping is based on fluorescent detection of SNP-specific hybridization probes on PCR products such as Molecular Beacons, Taqman™, KASP™ (Semagn et al. 2013), and Invader enhanced by array technology (Olivier 2005; Tapp et al. 2000). The end point detection of SNP genotypes requires either the PCR product to be read by a plate reader, or a real time PCR instrument. For simple and accurate genotyping to detect SNPs without sophisticated equipment, CAPS and dCAPS techniques provide a solution by using endonuclease. The CAPS assay consists of digestion of PCR amplicons with one or several restriction enzymes, and separation of the digested DNA on agarose or polyacrylamide gels. CAPS markers are co-dominant and locus specific and have been indispensable for fine mapping of target genes in map-based cloning. Even though SNPs can be converted into size polymorphisms through the CAPS assay, the disadvantage of this technique is that sometimes the SNP cannot be recognized by any endonuclease. To overcome this limitation, a variant of the CAPS assay, known as dCAPS, was proposed by Neff et al. (1998). Combined with dCAPS, CAPS can be used to detect virtually any SNP at a lower cost than that of other assays. The CAPS technique has been proved to be very useful for genotyping known point mutants in segregating populations. However, the process is time consuming and not efficient for large breeding programs. If the initial cost of instrumentation is not a barrier, then genotyping technologies based on end point detection, such as KASP™ are efficient, cost effective and reliable.

3. CONCLUSIONS

Our recommendations for tobacco breeding programs heavily weigh in favour of SNP markers. SNP or INDEL (Insertion Deletion) markers in the form of KASP™ and Taqman™ are co-dominant and reliable, and can be used for tracking trait-specific mutations or as closely linked markers. They are easy to use and efficient, especially for breeding programs handling a large volume of samples. The CAPS and dCAPS markers are also useful markers, because they are reliable, relatively inexpensive, and co-dominant markers that can be gene-specific. They can be used only when there are different PCR products from the two parents (e.g. they cannot be used where one allele results from a deletion). It is more time consuming to go from DNA to results, when compared to KASP™ or Taqman™, making it inefficient for breeding programs handling large volume of samples. The next most useful markers are closely-linked SCAR markers; they are reliable, but they are dominant markers and so less useful to breeders because they cannot identify heterozygotes. They are used when the sequence is known, but there is PCR product from only one parent. AFLP and RAPD markers are not ideal, but they can be used when the sequence information is unknown. RAPD markers are unreliable because they are not always repeatable. However, the procedure is simple and uses reagents and equipment routinely found in marker laboratories. AFLP markers are more reliable and give more repeatable results than RAPD markers, but they are not as convenient. The procedure is more expensive, takes longer and uses specialised equipment not routinely used in most marker laboratories.

Table 1: Available markers for major genes or traits in tobacco (*Nicotiana tabacum* L.)

Trait	Gene	Marker	Linkage	Dominance	Publication
Alkaloid locus B	*Nic2*		gene-specific	dominant	Shoji et al. (2010)
Black root rot resistance (ex *N. debneyi*) *Thielaviopsis basicola*		RAPD	closely-linked	dominant	Bai et al. (1995)
Black root rot resistance (ex *N. debneyi*) *Thielaviopsis basicola*		SCAR	1 cM	dominant	Julio et al. (2006)
Black root rot resistance (ex *N. debneyi*) *Thielaviopsis basicola SS192650*		CAPS	closely-linked	co-dominant	Qin et al. (2018)
Black shank race 0 resistance (ex *N. longiflora*) *Phytophthora nicotianae*	*Phl*	SCAR	closely-linked	dominant	Unpublished, personal communication
Black shank race 0 resistance (ex *N. plumbag*) *Phytophthora nicotianae*	*Php*	RAPD	closely-linked	dominant	Johnson et al. (2009)
Black shank race 0 resistance (ex *N. plumbag*) *Phytophthora nicotianae*	*Php*	RAPD	closely-linked	dominant	Johnson et al. (2009)
Black Shank Race 0 Coupling Phase	*Php*	SCAR	closely-linked	dominant	Unpublished, personal communication
Black Shank Race 0 Repulsion Phase	*Php*	SCAR	closely-linked	dominant	Unpublished, personal communication
Blue mold resistance *Peronospora tabacina*		SCAR	closely-linked	dominant	Milla et al.(2005)
Blue mold resistance *Peronospora tabacina*		SCAR	closely-linked	dominant	Julio et al. (2006)
Blue mold resistance *Peronospora tabacina*		SCAR	closely-linked	dominant	(Wu et al. 2015)
Brown spot (Alternaria) resistance *Alternaria alternata*		RAPD	closely-linked	dominant	Zhang et al. (2008)
Cis-abienol synthase; diterpene synthase	*abs cps2*		gene-specific	dominant	Sallaud et al. (2012)
Demethylase e10	*CYP82e10*	CAPS	gene-specific	co-dominant	Li et al. (2012)
Demethylase e4	*CYP82e4*	dCAPS	gene-specific	co-dominant	Li et al. (2012)
Demethylase e5	*CYP82e5*	dCAPS	gene-specific	co-dominant	Li et al. (2012)

Trait	Gene	Marker	Linkage	Dominance	Publication
Demethylase e10	*CYP82e10*	KASP	mutation specific	co-dominant	Unpublished, personal communication
Demethylase e4	*CYP82e4*	KASP	mutation specific	co-dominant	Unpublished, personal communication
Demethylase e5	*CYP82e5*	KASP	mutation specific	co-dominant	Unpublished, personal communication
N-Gene (*N. Glutinosa* Virus Resistance)	*N-gene*		gene specific	dominant	Whitham et al. (1994)
Potyvirus susceptibility	*eiF4E1.S*		gene specific	dominant	Dluge et al. (2018)
PVY resistance Potato virus Y	*va*	SCAR	5.1 cM	dominant	Julio et al. (2006)
PVY resistance Potato virus Y	*va*	RAPD	closely-linked	dominant	Lewis (2005)
PVY resistance Potato virus Y	*va (eIF4E)*		gene specific	dominant	Julio et al.(2015)
PVY RBV resistance	*eIF(iso)4E*	dCAPS	mutation specific	co-dominant	Takakura et al. (2018)
Powdery mildew resistance *Golovinomyces cichoracearum var. cichoracearum*	*NtMLO1 NtMLO2*	SSCP	gene specific	co-dominant	Fujimura et al. (2018)
Root-Knot resistance *Meloidogyne incognita*	*Rk*	RAPD	closely linked	dominant	Yi et al. (1998)
Tobacco mosaic virus (TMV) resistance	*N*	SCAR	gene-specific	dominant	Lewis (2005)
Tomato spotted wilt virus (TSWV) resistance		AFLP SCAR	<5cM	dominant	Moon et al. (2008)
Wildfire (race 0) resistance *Pseudomonas syringae* pv. tabaci tox-		RAPD	closely-linked	dominant	Yi et al. (1998)
Yellow Burley (Yb) trait Yb24-2	*YB1*	KASP	closely-linked	co-dominant	Edwards et al. (2017)
Yb24-4	*YB1*	KASP	closely-linked	co-dominant	Edwards et al. (2017)
Yb24-6	*YB1*	KASP	closely-linked	co-dominant	Edwards et al. (2017)
Yb24-7	*YB1*	KASP	closely-linked	co-dominant	Edwards et al. (2017)
EGY_2	*YB1*		gene specific	dominant	Edwards et al. (2017)

Trait	Gene	Marker	Linkage	Dominance	Publication
EGY_2kasp	*YB1*	KASP	gene and mutation specific	co-dominant	Edwards et al. (2017)
Yb5-1	*YB2*	KASP	closely-linked	co-dominant	Edwards et al. (2017)
Yb5-4	*YB2*	KASP	closely-linked	co-dominant	Edwards et al. (2017)

Table 2: QTLs and associated markers in *Nicotiana tabacum* L.

Trait	Marker	Linkage	Dominance	% Phenotypic variation	Publication
Bacterial wilt resistance	AFLP	closely-linked	dominant	0.3	Nishi et al. (2003)
Bacterial wilt resistance	SSR	LG 7, 6.6 cM	co-dominant	0.321	Drake-Stowe et al. (2017)
Black shank race 1 resistance	SSR	LG 4	co-dominant	0.25	Vontimitta and Lewis (2012)
Black shank resistance	SSR	LG 7, 6.6 cM	co-dominant	0.38	Drake-Stowe et al. (2017)
Increased leaf number	AFLP	3.1cM	dominant	increase 7-8 leaves	Lewis et al. (2007)
Delayed flowering time	AFLP	3.3cM	dominant	8 days	Lewis et al. (2007)
Black shank race 1 resistance	SSR	LG 8	co-dominant	0.2	Vontimitta and Lewis (2012)

4. REFERENCES

- Bai D, Reeleder R, Brandie JE (1995) Identification of two RAPD markers tightly linked with the Nicotiana debneyi gene for resistance to black root rot of tobacco. Theor Appl Genet 91:1184-1189
- Boumedine KS, Rodolakis A (1998) AFLP allows the identification of genomic markers of ruminant Chlamydia psittaci strains useful for typing and epidemiological studies. Res Microbiol 149:735-744
- Chang C, Bowman JL, DeJohn AW, Lander ES, Meyerowitz EM (1988) Restriction fragment length polymorphism linkage map for Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 85:6856-6860
- Dluge KL, Song Z, Wang B, Tyler Steede W, Xiao B, Liu Y, Dewey RE (2018) Characterization of Nicotiana tabacum genotypes possessing deletion mutations that affect potyvirus resistance and the production of trichome exudates. BMC Genomics 19:484
- Drake-Stowe K, Bakaher N, Goepfert S, Philippon B, Mark R, Peterson P, Lewis RS (2017) Multiple Disease Resistance Loci Affect Soilborne Disease Resistance in Tobacco (Nicotiana tabacum). Phytopathology 107:1055-1061
- Edwards KD, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans AD, Bombarely A, Allen F, Hurst R, White B, Kernodle SP, Bromley JR, Sanchez-Tamburrino JP, Lewis RS, Mueller LA (2017) A reference genome for Nicotiana tabacum enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. BMC Genomics 18:448
- Fujimura T, Sato S, Tajima T, Arai M (2018) Powdery mildew resistance in the Japanese domestic tobacco cultivar Kokubu is associated with aberrant splicing of MLO orthologues. Plant Pathology 65:1358-1365
- Johnson CS, Wernsman EA, LaMondia JA (2009) Effect of a Chromosome Segment Marked by the Php Gene for Resistance to Phytophthora nicotianae on Reproduction of Tobacco Cyst Nematodes. Plant Disease:309-315
- Julio E, Cotucheau J, Decorps C, Volpatti R, Sentenac C, Candresse T, Borne FDd (2015) A Eukaryotic Translation Initiation Factor 4E (eIF4E) is Responsible for the "va" Tobacco Recessive Resistance to Potyviruses. Plant Mol Biol Rep 33
- Julio E, Verrier JL, Dorlhac de Borne F (2006) Development of SCAR markers linked to three disease resistances based on AFLP within Nicotiana tabacum L. Theor Appl Genet 112:335-346
- Knox MR, Ellis TH (2001) Stability and inheritance of methylation states at PstI sites in Pisum. Molecular genetics and genomics : MGG 265:497-507
- Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. The Plant journal : for cell and molecular biology 4:403-410
- Lewis RS (2005) Transfer of resistance to potato virus Y (PVY) from Nicotiana africana to Nicotiana tabacum: possible influence of tissue culture on the rate of introgression. Theor Appl Genet 110:678-687

- Lewis RS, Milla SR, Kernodle SP (2007) Analysis of an introgressed Nicotiana tomentosa genomic region affecting leaf number and correlated traits in Nicotiana tabacum. Theor Appl Genet 114:841-854
- Li D, Lewis RS, Dewey RE, Bowen SW, Miller RD (2012) Development of CAPS and dCAPS markers for CYP82E4, CYP82E5v2 and CYP82E10 gene mutants reducing nicotine to nornicotine conversion in tobacco. Molecular Breeding 29:589-599
- Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American journal of human genetics 44:397-401
- Liu YG, Mitsukawa N, Lister C, Dean C, Whittier RF (1996) Isolation and mapping of a new set of 129 RFLP markers in Arabidopsis thaliana using recombinant inbred lines. The Plant journal : for cell and molecular biology 10:733-736
- Milla SR, Levin JS, Lewis RS, Rufty RC (2005) RAPD and SCAR Markers Linked to an Introgressed Gene Conditioning Resistance to D.B. Adam. in Tobacco. Crop Science 45:2346
- Moon HS, Nicholson JS, Lewis RS (2008) Use of transferable Nicotiana tabacum L. microsatellite markers for investigating genetic diversity in the genus Nicotiana. Genome 51:547-559
- Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. The Plant journal : for cell and molecular biology 14:387-392
- Nishi T, Tajima T, Noguchi S, Ajisaka H, Negishi H (2003) Identification of DNA markers of tobacco linked to bacterial wilt resistance. Theor Appl Genet 106:765-770
- Olivier M (2005) The Invader assay for SNP genotyping. Mutation research 573:103-110
- Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. TAG Theoretical and applied genetics Theoretische und angewandte Genetik 85:985-993
- Qin Q, Li Y, Ding N, Li D, Martinez N, Miller R, Zaitlin D, Yang S (2018) Development of user-friendly markers for disease resistance to black root rot of tobacco through genotyping by sequencing. Molecular Breeding 38
- Sallaud C, Giacalone C, Topfer R, Goepfert S, Bakaher N, Rosti S, Tissier A (2012) Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (Nicotiana tabacum) glandular trichomes. Plant J 72:1-17
- Semagn K, Babu R, Hearne S, Olsen M (2013) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Molecular Breeding 33:1-14
- Shoji T, Kajikawa M, Hashimoto T (2010) Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. Plant Cell 22:3390-3409
- Takakura Y, Udagawa H, Shinjo A, Koga K (2018) Mutation of a Nicotiana tabacum L. eukaryotic translation-initiation factor gene reduces susceptibility to a resistance-breaking strain of Potato virus Y. Mol Plant Pathol 19:2124-2133

- Tapp I, Malmberg L, Rennel E, Wik M, Syvanen AC (2000) Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan assay and Molecular Beacon probes. BioTechniques 28:732-738
- van Der Wurff AW, Chan YL, van Straalen NM, Schouten J (2000) TE-AFLP: combining rapidity and robustness in DNA fingerprinting. Nucleic acids research 28:E105
- Vontimitta V, Lewis RS (2012) Mapping of quantitative trait loci affecting resistance to Phytophthora nicotianae in tobacco (Nicotiana tabacum L.) line Beinhart-1000. Molecular Breeding 29:89-98
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic acids research 23:4407-4414
- Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B (1994) The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. Cell 78:1101-1115
- Wu X, Li D, Bao Y, Zaitlin D, Miller R, Yang S (2015) Genetic Dissection of Disease Resistance to the Blue Mold Pathogen, Peronospora tabacina, in Tobacco. agronomy 5:555-568
- Yi HY, Rufty RC, Wernsman EA, Conkling MC (1998) Mapping the Root-Knot Nematode Resistance Gene (Rk) in Tobacco with RAPD Markers. Plant Disease 82
- Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Molecular ecology 11:1-16
- Zhang HY, Yang YM, Li FS, He CS, Liu XZ (2008) Screening and characterization a RAPD marker of tobacco brown-spot resistant gene. African Journal of Biotechnology 7:2559-2561
- Yi YH, Rufty RC, Wernsman EA (1998). Identification of RAPD markers linked to the wildfire resistance gene of tobacco using bulked segregant analysis. Tobacco Sci. 42:52-57.

ANNEX A

Table S1: Molecular Markers in Tobacco – Marker Information for Major Genes

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
Alkaloid locus B	*Nic2*		gene-specific	dominant	Shoji, T.		55 °C		Forward: ATCCGAGCCCGAGGAGGGTGG Reverse: CGTGAATTATGACGTATCCCTG	Shoji, T., *et al*., 2010, The Plant Cell, 22:3390-3409
Black root rot resistance *(ex N. debneyi) Thielaviopsis basicola*		RAPD	closely-linked	dominant	Bai, D.		35 °C		UBC83: GGGCTCGTGG UBC418: GAGGAAGCTT	Bai, D., *et al*., 1995, Theor Appl Genet 91:1184-1189
Black root rot resistance *(ex N. debneyi) Thielaviopsis basicola*		SCAR	1 cM	dominant	Julio, E.		62 °C	temperature critical - bands for all genotypes at 62 °C, better separation at 59 °C	Forward: TAACAGCCTAACCCTATTCC Reverse: AATTCAAGGGGTAAAGCTATTC	Julio, E., *et al*., 2006, Theor Appl Genet 112:335-346
Black root rot resistance *(ex N. debneyi) Thielaviopsis basicola SS192650*	*BRR1*	CAPS	closely linked	co-dominant	Qin, Q.				Forward: GCATTAACTGATTAATGGGAAGGACCA Reverse: GTCCAACAAGTGAGTTCATCATTGGAT	Qin, Q., *et al.*, 2018, Mol Breed 38:76
Black shank race 0 resistance *(ex N. longiflora) Phytophthora nicotianae*	*Phl*	SCAR	closely-linked	dominant	Li, D.		55 °C		Forward: TTGGAGGACTTCGGAGAGTG Reverse: GAACATTCTGGGCCTCTTCA	unpublished, personal communication

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
Black shank race 0 resistance *(ex N. plumbag) Phytophthora nicotianae*	*Php*	RAPD	closely-linked	dominant	Johnson, E.S.		40 °C	linked in coupling phase	TCCCATGCTG	Johnson, E.S., *et al*., 2002, Plant Disease, 86:1303-1309
Black shank race 0 resistance *(ex N. plumbag) Phytophthora nicotianae*	*Php*	RAPD	closely-linked	dominant	Johnson, E.S.		40 °C	linked in repulsion phase	CCAGGAGGAC	Johnson, E.S., *et al*., 2002, Plant Disease, 86:1303-1309
Black Shank Race 0 Coupling Phase		SCAR	closely-linked		Kudithipudi, C. Adams, A.		65 °C		Forward: TCCCATGCTGTCCCAATCCCAC Reverse: TCCCATGCTGCGATGTTGGA	unpublished, personal communication
Black Shank Race 0 Repulson Phase		SCAR			Kudithipudi, C. Adams, A.		68 °C		Forward: TTCCCCCGCTCTTGTTAAGTGGTTA Reverse: TTCCCCCGCTTTGTGAAGATAGAGA	unpublished, personal communication
Brown spot (Alternaria) resistance *Alternaria alternata*		RAPD	closely-linked	dominant	Zhang, H.Y.		37 °C		S361: CATTCGAGCC	Zhang, H.Y., *et al*., 2008, African J of Biotech, 7(15):2559-2561

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
Cis-abienol synthase; diterpene synthase	*abs* *cps2*		gene-specific	dominant	Sallaud, C.				abs: Forward: CCATGGCGTTTCGACTTTTACGGA Reverse: GCCATGGCACAATGAGCAACGTCTGA cps2: Forward: GCGAATTCGAGGTCTTAATATGTGTTCAATCCAGA Reverse: CGGGATCCTTTCTAATTTAATTTTTGTTTTATTCTTC	Sallaud, C. *et al*.,2012, Plant J 2012, 72(1):1-17
Demethylase *e10*	*CYP82e10*	CAPS	gene-specific	co-dominant	Li, D.	*HphI*	54 °C		Forward: TGAAATATACAAGTAAGGTATAAAACACTA Reverse: TACCACTTCTATAGGATTTACCT	Li, D., *et al*., 2012, Molecular Breeding, 29:589-599
Demethylase *e4*	*CYP82e4*	dCAPS	gene-specific	co-dominant	Li, D.	*DdeI*	52 °C		Forward: AGCAGACACAGTTGCTCTTCACATAACTT Reverse: CAACACAATCTTCTACATTTTCGTG	Li, D., *et al*., 2012, Molecular Breeding, 29:589-599
Demethylase *e5*	*CYP82e5*	dCAPS	gene-specific	co-dominant	Li, D.	*HaeIII*	54 °C		Forward: AGAAAGCACAAGAAGAGATCGATAA Reverse: TCTCTCTGGATCAAACTTATCAGGATTTGG	Li, D., *et al*., 2012, Molecular Breeding, 29:589-599
Demethylase *e10*	*CYP82e10*	KASP	mutation specific	co-dominant	Pramod, S. & Adams, A.			KASP assays can be ordered from LGC genomics by providing the sequence	CTGGGAAAAATTATGAATCCGGTAAAGGAGATGAACAAGTGGAGAGATTTAGGAAAGCGTTTAAGGATTTTATAATTTTATCAATGGAGTTTGTGTTATG[G/A]GATGCTTTTCCAATTCCATTGTTCAAATGGGTGGATTTTCAAGGCCATGTTAAGGCCATGAAAAGGACATTTAAGGATATAGATTCTGTTTTTCAGAATT	unpublished, personal communication

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
Demethylase *e4*	*CYP82e4*	KASP	mutation specific	co-dominant	Pramod, S. & Adams, A.			KASP assays can be ordered from LGC genomics by providing the sequence	AAACAAGGTATGTGAATAATTGATATTCCTTTTTTAATTATTCTTTTTTCCAGAGTTTGGTCTTGGATGCAGCAGACACAGTTGCTCTTCACATAAATTG[G/A]GGAATGGCATTATTGATAAACAATCAAAAGGCCTTGACGAAAGCACAAGAAGAGATAGACACAAAAGTTGGTAAGGACAGATGGGTAGAAGAGAGTGATA	unpublished, personal communication
Demethylase *e5*	*CYP82e5*	KASP	mutation specific	co-dominant	Pramod, S. & Adams, A.			KASP assays can be ordered from LGC genomics by providing the sequence	ATTGTAGGACTAGTAACCCTTACACTTCTCTTCTACTTCCTATGGCCCAAAAAATTTCAAATACCTTCAAAACCATTACCACCGAAAATTCCCGGAGGGT[G/A]GCCGGTAATCGGCCATCTTTTCTACTTCGATGATGACGGCGACGACCGTCCATTAGCTCGAAAACTCGGAGACTTAGCTGACAAATACGGCCCGGTTTTC	unpublished, personal communication
EGY_2	*NtEGY2*		gene specific	dominant	Edwards, K.D., *et al.*			PCR EGY_2 region and then use EGY_2kasp to track the indel	EGY2-nF - GCAATCGTTGTCCAGTGTCTA EGY2-nR - TTTCCGACCTCTGTTACATCA	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448
EGY_2kasp	*NtEGY2*	KASP	Gene and mutation specific	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	GCTTAGTGTTGCAAATTGTACAGTCACTTAAAACATTCTGTTTGTGCTGCTATCAGCTTGGTGGACCTCT[T/]GTCACTTCCTTGGGGATTGTATGTGCTTATATGTCAGGTAATAAAAATACATTGCTTTATGAAGTATTATGCTTGTGTAATAATATCGCCGTCCTATTTCATGGTATA	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448
Potyvirus susceptibility	eiF4E1.S		gene specific	dominant	Dluge, K. L., et al.				eIF4E_C_F: GGATCCACGAAAATGGCAGAGGAAGC, eIF4E_C_R: GAGCTCGCTAATGTCTATAAACTTTCCAGTCCA	Dluge, K.L., *et al.,* 2018, BMC Genomics 19:484

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
PVY resistance Potato virus Y	*va*	SCAR	5.1 cM	dominant	Julio, E.		62 °C		Forward: TTAACAACAGCTTTTAGCAGACAC Reverse: ACAACTGGCAAGCTAAGCTCATT	Julio, E., *et al*., 2006, Theor Appl Genet 112:335-346
PVY resistance Potato virus Y	*va*	RAPD	closely-linked	dominant	Lewis, R.		40 °C		a cluster of RAPD primers	Lewis, R., *et al*., 2005, Theor Appl Genet 110:678-687
N-Gene (N. Glutinosa Virus Resistance)	*N-gene*		gene-specific	dominant	Adams, A., Kudithipudi, C.		60 °C		Forward: AACAAGTATAACTTTTTATGCTCAAATCAG Reverse: CAGTGTCATTGTTCTAGTTCTCAATATACA	Whitham, S., et al., 1994, Cell 78 (6), 1101-1115
TMV resistance Tobacco mosaic virus	*N*	SCAR	gene-specific	dominant	Lewis, R.		55 °C		Forward: ACCAGAATGATATGTTCCAC Reverse: GGACTCAACGTTAATTCTCTG	Lewis, R., *et al*., 2005, Crop Science, 45:2355-2362
TSWV resistance Tomato spotted wilt virus		AFLP SCAR	<5cM	dominant	Moon, H.		57 °C(172) 57 °C(268) 50 °C(169) 55 °C(228)		Four AFLP markers (AAC/CCC172,ACG/CCG169,AAG/CGA228, ACT/CTA268) SCAR markers: 172 Forward: AGCTTCTTTTCTCTTTCCATTTTT 172 Reverse: CAGAAGAAAAACTGCTGGAGCTAT 268 Forward: CTGATCGTTCCAGCAGGTTCTTAT 269 Reverse: GGAGCTATTTCCAGACACGAA 169 Forward: ACTTTTCACACCAAAAACTCACG 169 Reverse: GATGATGATAAAGATTGAAGAAAACAA 228 Forward: TAGATGTCATGAATGGAACTACGG 228 Reverse: TTTTGATCGAAAAACCCAACC	Moon, H. and Nicholson, J.S., 2007, Crop Science 47:1887-1894

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
Wildfire (race 0) resistance *Pseudomonas syringae* pv. t*abaci* tox-		RAPD	closely-linked	dominant	Yi, H.Y.		36 °C		OPW-10: TCGCATCCCT UBC-575: GGAGATGTAC	Yi, H.Y., *et al*., 1998, Tobacco Science, 42:52-57
Yellow Burley (Yb) trait Yb24-2	*YB1*	KASP	closely-linked	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	GAATCCTATGCTGGGATTAAGCCAAGCTG GTTTAGCTGCATCGGGATTTAATCCTTCGT T[T/C]GTGGGGATTGGTGCTGGTTATGATA TAAATAGCATTAATCCGGGCGTTTTAGGTA GTAGT	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448
Yb24-4	*YB1*	KASP	closely-linked	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	GTGGAATCTAAGGAGGGTAAAGTGTAAG CAAGTTTATCCTTGTCCTAGAAAGGTAGAG AT[A/G]TTGTTTCCGATAGATCCTTGGCTA AAAAAACGCAGAAAAGAAGCAGTGTCAAC AAGTAAT	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448
Yb24-6	*YB1*	KASP	closely-linked	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	GCAATATAAAAGTCAATACTAGTCTTAACG GTGTGAGGCCTTTTGGAAAAACTTGTACG G[T/G]CTTGGTCCAAAACAGACAATATCAC ACCATTTAAGAGTATCTTTAGGCCGTTTTA GTCCG	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448
Yb24-7	*YB1*	KASP	closely-linked	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	GACGGTATGGTTAAAGATGTATATACGGA AGACTTCTCGATCACCATCGGAGAGAAAT GC[T/G]TATTTTTATGTCTCTTTTGGCCCAA TGGAGAGCCCTTATTTATAAATGTAGAAG ATTGAA	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448

Trait	Gene	Marker	Linkage	Dominance	Developed By	Enzyme	Annealing Temp	Comments	Primer Sequence or Associated Sequence for KASP design	Publication
Yb5-1	*YB2*	KASP	closely-linked	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	TTAATAATAGTTCCTTTTTGACCGCGGGTACGAGGGCTAGCGTGAAATTGGCCACGAAGT[T/C]TGCCTAGATTTGAGACTTAATAGACGTTCGGGGTTGATACTCAATGTCGTACCTACTAAT	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448
Yb5-4	*YB2*	KASP	closely-linked	co-dominant	Edwards, K.D., *et al.*			KASP assays can be ordered from LGC genomics by providing the sequence	GACTACTATCATTGCCAGTCTCAACCTGTTCTGGTTCAAGCTTCAGTTCCACGTCCTCTC[T/C]ATCAGTCTCTGCAGGCTCAACTGTTTCGATCTCCTTGTGTTCTGGAACTTCTTTAGAATC	Edwards, K.D., *et al.*, 2017, BMC Genomics, 18:448

1. SNP or INDEL (**In**sertion **Del**etion) markers in the form of KASP™ and Taqman™ are co-dominant and reliable, and can be used for tracking trait specific mutations or as closely linked markers. They are easy to use and efficient, especially for breeding programs handling large volume of samples.
2. The CAPS and dCAPS markers are also useful markers, as they are reliable, and co-dominant markers that can be gene-specific. They can be used only when there are different PCR products from the two parents (e.g. they cannot be used where one allele results from a deletion). It is more time consuming to go from DNA to results, when compared to KASP™ or Taqman™, making it inefficient for breeding programs handing large volume of samples
3. The next most useful markers are closely-linked SCAR markers; they are reliable, but they are dominant markers and so less useful to breeders because they cannot identify heterozygotes. They are used when the sequence is known, but there is PCR product from only one parent.

RAPD and AFLP markers are not ideal markers, but they can be used when the sequence information is unknown.

4. RAPD markers are unreliable because they are not always repeatable. However, the procedure is simple and uses reagents and equipment routinely found in marker laboratories.
5. AFLP markers are more reliable and give more repeatable results than RAPD markers, but they are not as convenient. The procedure is more expensive, takes longer and uses specialised equipment not routinely used in most marker laboratories.

Table S2: Molecular Markers in Tobacco – QTLs

Trait	Marker	Linkage	Dominance	Author	% Phenotypic Variation	Annealing Temp	Primer Sequence	Publication
Bacterial wilt resistance	AFLP	closely-linked	dominant	Nishi, T.	30%		adaptor PstI: CTCGTAGACTGCGTACATGCA, CATCTGACGCATGT MseI: GACGATGAGTCCTGAG, TACTCAGGACTCAT amplification primer pp: GACTGCGTACATGCAGNNN, GATGAGTCCTGAGTAANNNN selective extensionof the amplication primer M17: PstI (AAA) MseI(CAGG); M84: PstI(CAG) MseI(CCTA)	Nishi, T., *et al.*, 2003, Theor Appl Genet, 106:765-770
Black shank race 1 resistance	SSR	LG 4	co-dominant	Vontimitta, V.	25%	55 °C	PT61373: Forward- GCGGGATAAACATGGGTAAA; Reverse- ACCCAAATAACCGCTCACAT PT51164: Forward- CAAGGACCACTCCAATGCTT; Reverse- TTCCAATGTTGTTTCTGTGTCTG	Vontimitta, V., *et al.*, 2012, Mol Breeding, 29:89-98
Increased leaf number	AFLP	3.1cM		Lewis R.	increase 7-8 leaves		M1 = AFLP marker EACT/MCCT-130	Lewis et al. 2007b, Theor Appl Genet,114:842-854
Delayed flowering time	AFLP	3.3cM		Lewis R.	8 days		M1 = AFLP marker EACT/MCCT-130	Lewis *et al.* 2007b, Theor Appl Genet, 114:842-854
Black shank race 1 resistance	SSR	LG 8	co-dominant	Vontimitta, V.	20%	55 °C	PT61472: Forward- TCCAATACCTTTAATGCATCTCC; Reverse- GCATGACATGTTGAAGTGGG PT30174: Forward- TGGTCGATCAACATGACAAA; Reverse- TCTAAATCACGCTGCATTGG	Vontimitta, V., *et al.*, 2012, Mol Breeding, 29:89-98