Spring 2007

# Studies of verticillium wilt and characterization of candidate verticillium wilt resistance genes in the mint species Mentha longifolia (L) Huds

Kelly Jean Vining
*University of New Hampshire, Durham*

# STUDIES OF VERTICILLIUM WILT AND CHARACTERIZATION OF CANDIDATE VERTICILLIUM WILT RESISTANCE GENES IN THE MINT SPECIES *MENTHA LONGIFOLIA* (L.) HUDS.

BY

KELLY JEAN VINING

Bachelor of Arts, University of Southern Maine, 1996

Master of Science, University of New Hampshire, 1999

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

in

Genetics

May, 2007

This dissertation has been examined and approved.

Dissertation Director, Thomas M. Davis
Professor of Plant Biology and Genetics

Estelle M. Hrabak
Associate Professor of Plant Biology
and Genetics

Louis S. Tisa
Associate Professor of Microbiology
and Genetics

Cheryl Smith
Extension Professor of Plant Biology

Robert O. Blanchard
Professor Emeritus of Plant Biology

Dennis Mathews
Research Assistant Professor
of Plant Biology

5/3/2007
Date

## ACKNOWLEDGEMENTS

Like all things, the completion of this dissertation was dependent on many conditions. I wish to acknowledge the opportunities I have been afforded, and the people whose work facilitated and complemented my own efforts.

Scientific research cannot proceed without funding. I would like to thank the Mint Industry Research Council for funding for UNH mint research at the beginning of my thesis work in 2001, and at its completion in 2006. My work was also supported by assistantships from the UNH Genetics Program, by UNH Graduate School Summer Teaching Fellowships, and by Hatch grant NH00433.

I have been fortunate to have a diverse dissertation advisory committee, composed of one Professor, two Associate Professors, one Extension Professor, one Research Assistant Professor, and one Professor Emeritus. They have been attentive and helpful committee members, and they are all pleasant and interesting human beings as well. Special thanks must be given to my dissertation director, Thomas M. Davis, for allowing my work to proceed despite long funding droughts, and for patiently tolerating the extremes sometimes exhibited by graduate students experiencing the frustrations and triumphs of molecular biology research.

Speaking of patience and tolerance, I am extremely grateful to my husband, Kevin Ronkko, who has lived with me for my entire graduate student career, and continues to do so.

My thesis work was accomplished with the help of a number of undergraduate assistants, all of whom have been especially bright, capable and willing in the face of tedious tasks. Sarah Olson-Prochazka and Elisabeth Antanavich, with whom I have worked most closely, deserve special recognition. Other lab colleagues also must be thanked: Zhang Qian is a co-author of the publications from which chapters I and II of this dissertation are drawn. She has always been concerned with my progress and has offered helpful advice. Melanie Shields has been an astute proofreader of this dissertation and other writings. All of these people are good friends.

I would like to acknowledge my predecessors and contemporaries in mint verticillium wilt research, some of whom I have had the good fortune to meet. The field work of Raymond Nelson, Ralph Green, M.L. Lacy, C.E. Horner and Fred Crowe provided the foundation for my verticillium wilt resistance trials. Dennis Johnson has shared useful information, offered helpful manuscript criticism, and continues to be a collaborator.

Finally, I must express my deep appreciation for the genus *Mentha*, and for my *Mentha longifolia* study subjects in particular. I have been incredibly lucky to have had the opportunity to do my dissertation work with mint plants which are so beautiful in their diversity, so stimulating in their aromaticity, and which continue to be a source of inspiration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

STUDIES OF VERTICILLIUM WILT AND CHARACTERIZATION OF CANDIDATE VERTICILLIUM WILT RESISTANCE GENES IN THE MINT SPECIES *MENTHA LONGIFOLIA* (L.) HUDS.

by

Kelly Jean Vining

University of New Hampshire, May, 2007

To investigate the genetic basis of verticillium wilt resistance in mint (*Mentha* L., Lamiaceae), wild-collected germplasm obtained from the United States Department of Agriculture was employed to develop breeding populations for wilt resistance screening and molecular genetic study, including cloning of candidate verticillium wilt resistance genes.

A collection of fourteen *Mentha longifolia* accessions from Europe, Asia and South Africa was analyzed for morphological traits, oil composition, and verticillium wilt resistance. In addition, a preliminary molecular diversity assessment was conducted utilizing randomly amplified polymorphic DNA (RAPD) markers. The accessions were found to be diverse regarding all observed traits and the South African accessions in particular were shown to

possess unique features. Most importantly, highly wilt-resistant and highly wilt-susceptible accessions were observed.

A collection of twenty-seven resistance gene analogs (RGAs) was isolated from *M. longifolia* accessions using a PCR-based approach with primers targeting the conserved nucleotide binding site (NBS) domain found in most plant disease resistance genes. The mint RGAs shared predicted amino acid sequence similarity with disease resistance genes and RGAs from various other plant species, and were grouped into seven distinct families based on DNA and predicted amino acid sequence similarity. In addition to the NBS-related RGAs, a fragment of a verticillium wilt resistance gene candidate was isolated from a verticillium-resistant *M. longifolia* accession using a combination of PCR-based approaches that exploited known sequences of tomato *Ve* (*Verticillium* resistance) genes.

Finally, the complete coding region of the mint verticillium wilt resistance candidate gene, *mVe1*, was cloned and sequenced. Alleles of *mVe1* were compared among four *M. longifolia* accessions used as crossing parents. These seven alleles were highly similar to each other (96.2-99.6% nucleotide identity) and had ~50% predicted amino acid sequence identity to the tomato *Ve* genes. F1 and F2 populations were genotyped with respect to *mVe1* alleles, and individuals from these populations were screened for wilt resistance. No correlation was found between any *mVe* allele and resistance or susceptibility to verticillium wilt in plants in the studied populations. However, this result does not discount the possibility that an *mVe1* gene product plays a role in mint verticillium wilt resistance.

# INTRODUCTION

This project was initiated to address a serious issue of commercial mint production in the United States: verticillium wilt disease. Peppermint and spearmint, specialty crops cultivated for their aromatic oils, are vegetatively-propagated perennials. The causal organism of mint verticillium wilt, *Verticillium dahliae*, is a soil-borne vascular wilt fungus that occurs worldwide and affects a large number of crop plants. Mint fields infested with *V. dahliae* suffer progressive decreases in annual oil yields and are eventually abandoned.

Genetic studies of commercial mints are hindered because of polyploidy and low fertility. In addition, because mints are vegetatively-propagated clones, little genetic diversity exists among the few commercial cultivars. Therefore, the approach taken to address the mint-verticillium wilt problem was to exploit a wild, diploid relative of the cultivated mints: *Mentha longifolia*. This species has a broad geographic distribution in Eurasia and Africa, is genetically diverse, and has proven to have many useful features in exploratory studies.

The disease organism, *V. dahliae,* disperses via two types of vegetative spores: conidia and microsclerotia. The latter can persist in soil for long periods of time. When spores germinate in the vicinity of host plants, fungal hyphae enter plant roots, grow through root tissue and reach the vascular system. The fungus invades systemically by producing waves of conidia that are carried upward in the transpiration stream. In mint, verticillium wilt has historically been spread via

infected stolons being planted in fields. The combination of the perennial growth habit of mint with the persistence of *V. dahliae* in soil makes verticillium wilt of mint a particularly difficult problem, as infestations worsen from year to year. While modern cultural practices include planting only certified disease-free stock, the fungus has already spread through most mint-growing regions of the United States.

This research project has generated substantial knowledge about the investigated species, *M. longifolia*, about verticillium wilt disease in mint, and about candidate resistance genes. Many valuable germplasm and genomic resources have been generated as well. This dissertation is presented as a series of related projects designed to demonstrate the usefulness of *M. longifolia* as a model for mint genetic studies and for plant pathology research. While each chapter is intended to stand alone, they collectively aim to provide a foundation on which future investigations may elaborate.

As a starting point for these investigations, a diverse collection of *M. longifolia* accessions was obtained from the United States Department of Agriculture (USDA) National Clonal Germplasm Repository in Corvallis, Oregon. These fourteen accessions represented plants that had been collected from Asia, Europe and South Africa. The diversity encompassed by this germplasm set is described in Chapter I, which was previously published in the journal HortScience, volume 40, issue 5, 2005. Three morphological traits—flower color, leaf shape and general growth habit—were described. Oil constituents were measured by collaborator Arthur Tucker at the University of Delaware using a

gas chromatograph, and principal constituents (>5% of total oil content) were reported. A protocol was developed for consistent screening of mint clonal plants for resistance vs. susceptibility to verticillium wilt, and a scale of disease symptom ratings used in all subsequent studies was introduced. The responses of the *M. longifolia* accessions to *Verticillium dahliae* inoculation were rated, and these data were used to select parents for sexual hybridizations which generated segregating populations used for studies in Chapter III of this dissertation. Examples of verticillium wilt disease symptoms are shown in photos in Appendix A. In addition to gross phenotypic character descriptions, Chapter I includes a molecular genetic diversity analysis using Randomly Amplified Polymorphic DNA (RAPD) molecular markers. Finally, Chapter I reports genome sizes (C values) for the two USDA *M. longifolia* accessions from South Africa; these values were determined by cooperating investigators at Kew Royal Botanic Gardens. The data from the various diversity analyses conducted in Chapter I reinforced the distinctiveness of the two South African accessions, and their value as study subjects with respect to many interesting genetic traits, in particular, oil quality and verticillium wilt resistance.

Chapter II introduces a collection of DNA sequences obtained in an initial exploration of the *M. longifolia* genome for putative disease resistance-related genes. The content of this chapter has been accepted for publication by the Journal for the American Society for Horticultural Science (JASHS). The mint DNA sequences presented therein were isolated via PCR-based approaches from six of the *M. longifolia* accessions evaluated in Chapter I. One type of

sequence, Resistance Gene Analogs (RGAs), are short (289-680 bp) sequences containing motifs common to most known disease resistance genes. Chapter II also reports the isolation of a verticillium-resistance-like (*Ve*-like) DNA sequence fragment (445 bp) from a verticillium wilt-resistant *M. longifolia* accession. The mint *Ve*-like fragment was obtained as a result of specific targeting of a homolog of genes (*Ve1* and *Ve2)* known to confer verticillium wilt resistance in tomato. The RGA sequences were submitted to the National Center for Biotechnology Information (NCBI) GenBank® biological sequence database. In sum, Chapter II represents the first disease-resistance-focused molecular genetic investigation of any mint species, and, in addition, describes the initial steps in the isolation of a candidate gene that could be a functional component of a mint verticillium wilt-specific resistance mechanism.

Chapter III describes the complete sequencing of the *mVe1*, a mint homolog of the tomato *Ve* genes. As a first step in assessing molecular genetic diversity of this gene in *M. longifolia*, *mVe1* sequences were isolated and sequenced from two resistant and two susceptible *M. longifolia* accessions that were used as parents of F1 and F2 populations. Plants from these F1 and F2 populations were genotyped with respect to *mVe1* alleles, and were screened for *Verticillium* inoculation response using the protocol developed in Chapter I. The hypothesis of association between an *mVe1* genotype and a *Verticillium* response phenotype was tested. No significant association between genotype and phenotype was detected. While this result does not discount the possibility that *mVe1* encodes a functional gene contributing to resistance, it does indicate

that the observed verticillium wilt susceptibility in these segregating *M. longifolia* populations could not be attributed to an identified allele of *mVe1*.

Appendix B describes the construction of a genomic library for a verticillium wilt-resistant *M. longifolia* accession. The library was intended to be a tool for isolation and cloning of *mVe1*, but *mVe1* was cloned by other means described in Chapters II and III. Nevertheless, the *M. longifolia* genomic library is a resource that was developed during the course of this work, and is available for use in future investigations.

CHAPTER I.

# *MENTHA LONGIFOLIA* (L.) HUDS.: A MODEL SPECIES FOR MINT GENETIC RESEARCH[1]


## Abstract

Fourteen accessions of *Mentha longifolia* from Europe, Asia and southern Africa were evaluated as a diploid resource for genetic research relevant to economically important traits in the polyploid mints of commerce. This readily available germplasm, obtained from the US Department of Agriculture (USDA), Agricultural Research Service (ARS), National Clonal Germplasm Repository (NCGR), was highly diverse with respect to oil composition, verticillium wilt resistance, plant architecture and other aspects of morphology. Trans-piperitone oxide was the primary oil component of accessions CMEN 17 and CMEN 18, while pulegone was most abundant in CMEN 20, CMEN 500, CMEN 501 and CMEN 585. CMEN 584 was the only carvone chemotype, and CMEN 682 was the only accession with high menthol content. CMEN 585, CMEN 17, CMEN 501 and CMEN81 were resistant to verticillium wilt, while CMEN 584 and CMEN 516 were highly susceptible. Molecular genetic diversity was evaluated. Sixty-three


[1] This chapter is revised from an article published in the journal HortScience ©2005 American Society for Horticultural Science Press. Vining, K.J., Q. Zhang, A. O. Tucker, C. Smith, T.M. Davis. *Mentha longifolia* (L.) L.: A model species for mint genetic research. Vol. 40, no. 5, pp. 1225-1229.

informative Random Amplified Polymorphic DNA (RAPD) marker bands provided data used to calculate pairwise similarity coefficients and to produce a Unweighted Pair-Group Method of Analysis (UPGMA) tree. CMEN 585 and CMEN 584 shared the greatest number of bands (16), and formed a distinct cluster in the UPGMA tree. Seven pairs of accessions had no bands in common, emphasizing the high degree of diversity represented by these accessions. *M. longifolia* has a comparatively small genome size (400-500 Mb) and diploid (2n=2x=24) genome constitution. These genomic features, along with its self-fertility, fecundity, and diversity with respect to economically relevant traits, make *M. longifolia* a potentially useful model system with relevance to the polyploid, cultivated mints, and also for the broader study of plant-microbe interactions and disease resistance mechanisms.

## Introduction

The principal *Mentha* (Lamiaceae = mint family) species of commerce in the United States are vegetatively propagated polyploids that, as such, present problems for transmission genetic analysis and conventional breeding. Native spearmint (*Mentha* ×*villosonervata* Opiz) is triploid (2*n*=3*x*=36), although morphologically similar to fertile, tetraploid spearmint (*M. spicata* L., 2*n*=4*x*=48). Scotch spearmint (*M.* ×*gracilis* Sole) is heptaploid (2*n*=7*x*=84). 'Mitcham' peppermint (*M.* ×*piperita* L.) is hexaploid (2*n*=6*x*=72) (Tucker and Naczi, 2005; Tucker et al., 1990; Udo et al., 1962). Polyploidy increases composite genome

size and allelic complexity, hampering structural and functional genomics studies, and may be accompanied by poor fertility. Not surprisingly, no genetic linkage maps have been constructed for *Mentha*. Other than an extensive literature on the genetics of oil quality, both classical (Hefendehl and Murray, 1976; Hendriks et al., 1976) and molecular (Croteau and Gershenzon, 1994), few traits have been characterized genetically, and few genomic resources have been developed. Gene identification in *Mentha* has been limited to genes encoding enzymes involved in menthol biosynthesis. These genes have been extensively characterized, and genetic manipulation of peppermint oil biosynthesis has begun (Mahmoud et al., 2004; Burke et al., 2004; Mahmoud and Croteau, 2001).

A diverse and widely distributed *Mentha* germplasm base has been documented (Tucker and Naczi, 2005). As of December 2004, the NCGR in Corvallis, OR maintained 441 *Mentha* accessions as vegetative clones and 52 as seed, representing 20 species and a diversity of interspecific hybrids (http://www.ars-grin.gov/cor/mentha/meninfo.html). Twenty-one accessions are *M. longifolia*, and six are listed as *M. longifolia* × *M. longifolia* hybrids. In addition, of the 67 accessions listed as simply '*Mentha* hybrid', 30 include *M. longifolia* in the known or inferred pedigree. The USDA collection of *M. longifolia* accessions represents a wide range of geographic, phenotypic, and genetic diversity. The USDA National Plant Germplasm System Plant Information (PI) numbers for each accession are provided (Table 1.1).

*M. longifolia* has the widest natural geographic distribution of any *Mentha* species, from western Europe to central Asia and in southern Africa. It may encompass 22 subspecies (Tucker and Naczi, 2005). Almost all are diploid ($2n = 2x = 24$), but some tetraploid ($2n = 4x = 48$) forms have also been described (Chambers and Hummer, 1994). The size of the *M. longifolia* genome was reported as 1C = 385 Mbp (Bennett and Leitch, 2005), and in the range of 2C = 0.84-0.99 pg (Gobert et al., 2002), or 1C = 405-477 Mbp. The *M. longifolia* C value is relatively small among those of cultivated plants, being comparable to that of rice (C = 400-466 Mbp) and about half that of tomato (C = 980 Mbp) (Bennett and Leitch, 2005). Phylogenetic analysis of *Mentha* indicates that *M. longifolia* is an ancestor of *M. spicata*, and may be the latter's organelle genome source (Bunsawat et al., 2004). In turn, *M. spicata* is a parent of *M.* ×*gracilis* and of *M.* ×*piperita* (Tucker and Naczi, 2005; Tucker et al., 1991; Tucker et al., 1990).

We have examined a set of *M. longifolia* accessions maintained by the NCGR, with particular attention to two traits of economic relevance: oil composition and resistance to verticillium wilt, an important disease of peppermint. This paper documents the phenotypic and genetic diversity among these *M. longifolia* accessions and reviews the features that make *M. longifolia* a potentially useful model species for *Mentha* genetic and genomic research.

## Materials and methods

### Germplasm

Fourteen accessions initially identified as *M. longifolia* (Table 1.1), including four subspecies (subsp. *capensis*, *hymalaiensis*, *polyadenia*, *typhoides*) were obtained as rooted plants or rhizomes from the NCGR. Plants were maintained in 22 cm pots in a greenhouse at the University of New Hampshire, Durham, NH, and were propagated vegetatively. Observations of morphology were made by direct visual examination and by light microscopy.

### Oil composition

Oils from whole flowering plants were distilled with a neo-Clevenger of Moritz after Kaiser and Lang with the modification of Hefendehl (Kaiser and Lang, 1951; von Rudloff, 1969). Mass spectra were recorded with a 5970 Hewlett-Packard Mass Selective detector coupled to a HP 5890 GC using a HP 50 m x 0.2 mm fused silica column coated with 0.33 mm FFAP (crosslinked). The GC was operated under the following conditions: injector temp. 250 °C; oven temp. programmed: 60 °C held for one min to 115°C at 2.5 °C per min, then to 210°C at 1.0°C per min and held for 30 min; injection size: 1mL (~50% solution in spectroscopy grade n-pentane) split 1:10. The MSD EI was operated at electron impact source 70 eV, 250 °C. Identifications were made by Kovats Indices and library searches of our volatile oil library supplemented with those of NBS, NIST, and Wiley.

**Verticillium wilt resistance screening**

Qualitative assessment of verticillium wilt resistance in all 14 accessions was conducted with a wild type *Verticillium dahliae* strain provided by Dennis Johnson at Washington State University. Based on the outcome of these initial trials, a subset of resistant and susceptible accessions was chosen for closer examination and for use as crossing parents for future genetic studies. The latter trials differed from the initial assessments in that a quantitative rating scale was employed, and a *V. dahliae* strain which was transformed with green fluorescent protein (GFP), provided by Linda Ciuffetti at Oregon State University, was used instead of the wild type strain. Both *V. dahliae* cultures were maintained in petri plates on Czapek-Dox medium, which was supplemented with 45 µg/ml hygromycin for the GFP strain.

*M. longifolia* cuttings of uniform size were rooted in 1206 cell packs with soilless Metro Mix 360 (The Scotts Co., Maryville, OH). They were maintained for 2 weeks in a growth chamber with fluorescent lighting (15 ± 3 µmol/$m^2$/sec), cool temperatures and a short-day cycle (22 °C, 10 hour light/ 20 °C, 14 hour dark) to minimize growth and prevent flowering.

Screening was performed using a modified root-dip inoculation technique based on that of Green and Simon (1996). An appropriate volume of *V. dahliae* in liquid Czapek-Dox medium was incubated for 1-2 weeks at room temperature on a shaker. The volume of liquid culture used depended on the number of cuttings

to be inoculated. Liquid cultures were strained through a single layer of Miracloth® (Calbiochem, San Diego, CA) in order to separate conidia from hyphae. The filtrate was centrifuged @ 10,000 x g for 5 min to pellet conidia. Pellets were re-suspended in 100 ml distilled $H_2O$. This step was performed to eliminate any residual medium as well as any substances secreted by the fungus. A hemacytometer was used to count conidia with a compound microscope (20x objective). The conidial suspension was diluted with distilled $H_2O$ to ~$10^7$ conidia/ml.

Screening trials had 12 replicate cuttings randomized within each treatment (control, inoculated). Control plants and inoculated plants were processed simultaneously. Cuttings were uprooted, soil was shaken from roots, and plants were placed in 50 ml beakers containing ~20 ml conidial suspension for 5 min. Plants were then replanted in new flats in the same soil type in which they were originally rooted. After inoculation, plants were kept in the growth chamber with minimal watering, continuing the aforementioned light and temperature conditions. After two weeks, plants were moved to the greenhouse under natural light and ambient temperature.

Observations were taken 8 weeks post-inoculation. For plants with dead primary stems, stem sections were surface-sterilized and plated on water agar to confirm the presence of *Verticillium dahliae*. Degree of plant stunting, horizontal leaf curling ("crescent leaf"), and tissue chlorosis and necrosis were all observed in order to assign an overall symptom severity rating from zero (0) to four (4) (Appendix A). A zero rating indicated lack of visible symptoms; a four rating

described a dead plant. Intermediate ratings described plants with mild to severe symptoms. Ratings data for inoculated plants were processed in Systat® v.10.0 (SPSS, Madison, WI) by ANOVA as a completely randomized design. Pairwise comparisons were made with Tukey's tests.

**DNA extraction**

DNA was extracted from fresh, unexpanded leaf tissue using the CTAB miniprep method of Torres et al. (1993), with the following modifications: sodium bisulphate was not used in the grinding buffer, chloroform/octanol was not added to the grinding slurry prior to 65 °C incubation, and the ethanol wash utilized 70% ethanol without sodium acetate. DNA concentration was measured fluorometrically, then DNA was diluted with sterile $H_2O$ to a standard stock concentration of 40 ng/µl.

**Polymerase chain reaction**

Fourteen oligonucleotide primers (Operon Technologies, Alameda, CA) were used individually in PCR to detect RAPD polymorphisms. DNA was amplified in 25 µl reactions using 100 ng template DNA, 0.1U Taq DNA polymerase (Eppendorf), 2.5 mM each of dNTP (Promega, Madison, WI) and 0.8 uM primer. The reactions were performed in a thermal cycler programmed for one cycle of 2 min at 94 °C followed by 39 cycles of 1 min at 94 °C, 2 min 30 s at 35 °C, 30 s at 45 °C, and a final elongation step of 10 min at 72 °C.

## Separation and visualization of amplification products

PCR products were separated on gels containing 1% NuSieve GTG agarose (FMC Bioproducts, Rockland, ME) and 1% IBI agarose (Shelton Scientific, Shelton, CT) electrophoresed in 1X TBE, pH 8.0 at 90V for at least 3.5 hours at 4 °C. Gels were stained with ethidium bromide and photographed under UV light.

## RAPD marker diversity analysis

Only informative markers (bands that were present in at least two accessions and absent in at least one) were included in the analysis. A total of 63 bands were treated in a binary format and scored as one (band present) or zero (band absent). Pairwise genetic similarities were calculated using the Jaccard similarity coefficient (a/(a+b+c)) (Jaccard, 1908). Additionally, a phenetic analysis was conducted using Paup*4.0b10. A dendrogram was generated using the UPGMA method with 1000 UPGMA replicates.

## Genome size determination

Root tips were harvested from accessions CMEN 585, CMEN 584, and CMEN 17, which were viewed as likely candidates for future use in genome library construction. Root tips were fixed in Farmer's solution (three parts ethanol:one part glacial acetic acid) and sent to Dr. I.J. Leitch, Jodrell Laboratory, Royal Botanic Gardens, Kew, where the C values were measured by Feulgen microdensitometry (Bennett and Leitch, 2005).

## Results

Among the morphological characters showing variation were flower color and leaf shape (Table 1.2). The two accessions from South Africa, CMEN 584 and CMEN 585, were markedly different in appearance from the other accessions. Besides differing in flower color from all but one of the other accessions, CMEN 584 and CMEN 585 had lanceolate leaves, while the others' leaves were ovate. Under the growth conditions in the UNH greenhouse, CMEN 584 and CMEN 585 had a tall upright growth habit, reaching a height of ~100 cm at flowering. CMEN 682 and CMEN 34 had a moderately upright growth habit, but only reached 50%-75% of the height of CMEN 584 and CMEN 585. The other accessions had a shorter upright growth habit.

Oil composition was highly variable among the accessions (Table 1.2). Pulegone was the principal oil component of CMEN 20, CMEN 500, CMEN 501 and CMEN 585. These accessions, along with CMEN 682 and CMEN81, contained moderate levels of menthone. CMEN 17, CMEN 18 and CMEN 635 had high levels of *cis*- or *trans*-piperitone oxide. CMEN 584 was the only accession for which the principal oil component was carvone.

The *M. longifolia* germplasm showed diversity in response to inoculation with *V. dahliae*. Symptoms first became apparent two to four weeks post-inoculation. Symptoms ranged from mild horizontal curling of apical leaves to complete necrosis. Nine accessions, selected as representing extremes of inoculation response and other traits of interest, were then screened more

rigorously, using the quantitative rating scale of zero to four (Table 1.2). Results of screenings conducted with the wild type *Verticillium dahliae* strain were highly consistent with those obtained using the GFP strain: the four accessions given zero to one ratings in the second trial had all received R qualitative ratings in the first trial, while accessions given ratings two and above in the second trial all had S ratings in the first trial (Table 1.2). Overall, CMEN 585 and CMEN 17 were the most resistant diploid accessions, and CMEN 584 and CMEN 516 were the most consistently susceptible diploid accessions.

Symptom development varied considerably among susceptible accessions. For example, by four to six weeks post-inoculation, CMEN 516 exhibited overall chlorosis of leaf tissue, mild to moderate crescent leaf and little or no stunting, while CMEN 584 was consistently stunted >50% compared to controls and had substantial crescent leaf symptoms. Both CMEN 516 and CMEN 584 primary stems had died by the time final observations were recorded. However, asymptomatic shoot growth was sometimes seen emerging at the soil surface after complete death of primary stems above the soil line, indicating that at least some portion of these plants survived and escaped or recovered from fungal infection.

Similarly, wilt-resistant accessions showed differences in response to fungal inoculation. CMEN 585 occasionally had mild to moderate horizontal curling of apical leaves ~four weeks post-inoculation, followed by production of asymptomatic leaves. CMEN 17 commonly displayed shortened internodes and

mild horizontal leaf curl ~four weeks post-inoculation, followed by apparent recovery.

For RAPD analysis (Fig. 1.1), 14 oligonucleotide primers produced a total of 63 informative bands. The number of bands shared by any pair of accessions ranged from 16 to zero: for example, CMEN 584 had 16 bands in common with CMEN 585 and none in common with five of the accessions (Fig. 1.2). The Jaccard similarity indices ranged from a high of 0.7619 (CMEN 584 vs. CMEN 585) to a low of zero (e.g., CMEN 584 vs. CMEN 682) (Fig. 1.2). A UPGMA tree had five nodes with bootstrap support of 50% or better (Fig. 1.3). CMEN 585 and CMEN 584 formed a group that was highly distinct from, and sister to, the other accessions.

The genome sizes of *M. longifolia* diploid accessions CMEN 584 and CMEN 585 were reported to be 4C = 1.75 pg (1C = 440 Mbp) and 4C = 1.64 pg (1C = 410 Mbp), respectively (Hanson, pers. comm.). The genome size of CMEN 17 was estimated to be 4C = 1.57 pg (1C = 385 Mbp) (Bennett and Leitch, 2005).

## Discussion

Our examination of 14 clonally propagated NCGR accessions of *M. longifolia* detected considerable phenotypic and genetic variation. Plant height, flower color and leaf shape were obviously variable among the accessions. Variation was noted but not systematically examined in other morphological features such as leaf color, leaf margin type and stem thickness. Although

*Mentha* species are distinguished primarily by their essential oil contents, the range of morphological variation in *M. longifolia* points to its potential for development as an ornamental species as well as a genetic model system.

*M. longifolia* has been a subject of numerous oil composition studies (Shaiq et al., 2002; Ghoulami et al., 2001; Venskutonis, 1996; Kokkini et al., 1995; Kokkini and Papageorgiou, 1988; Hefendehl, 1977). *M. longifolia* oil composition has attracted recent attention due to its potential for antimicrobial and antifungal activity (Mimica-Dukic et al., 2003; Abou-jawdah et al., 2002). The present paper adds data on 14 NCGR accessions to the substantial body of knowledge about *M. longifolia* oil composition. CMEN 584 is the only carvone chemotype in the NCGR *M. longifolia* collection; however, other *M. longifolia* carvone chemotypes have been reported (Kokkini et al., 1995; Hefendehl, 1977).

One major focus of our research with *M. longifolia* is the identification of plants with differential responses to the fungal pathogen *V. dahliae.* Toward that end, all *M. longifolia* accessions were initially screened with a wild type strain of the fungus. When a GFP-transformed strain became available, it was used for subsequent screenings of selected accessions, and of F1 and F2 populations developed from resistant x susceptible crosses (results presented in Chapter III). Trials conducted with the GFP strain of *V. dahliae* produced results equivalent to those performed with the wild type strain. The GFP strain is of interest as a potentially useful tool for the study of the early events of fungal penetration of a plant host (Lorang et al., 2001).

The disease resistance screening showed that some accessions are highly susceptible to verticillium wilt and others are highly resistant. The most susceptible *M. longifolia* accessions, CMEN 516 and CMEN 584, showed differences in symptom development, although the eventual outcome for both was primary stem death. Both genotypes occasionally exhibited secondary growth after death of primary stems, indicating that even when primary stems were completely dead above the soil line, some stem tissue survived and was capable of regeneration. It is possible that part of the disease resistance response in these plants involves blockage of part of the root vascular system in order to sequester the invading fungus. In the field, such a response could allow the plants to escape verticillium wilt disease by growing via secondarily-produced shoots and stolons to a non-infested area. A strategy for outgrowing soil-borne pathogens is especially important for a perennial species with a primarily asexual mode of reproduction.

*M. longifolia* is an interesting species from a phylogenetic perspective, because it is recognized as the most phenotypically diverse species of the taxonomically complex *Mentha* genus (Gobert et al., 2002). These investigators used AFLP markers to analyze 62 *Mentha* accessions, 6 of which are genotypes in the present study. They found that *M. longifolia* groups as a distinct taxon from other *Mentha* species, and is most closely related to *M. spicata* and *M. suaveolens*. The present study, which was aimed only at assessing genetic diversity in *M. longifolia*, demonstrated substantial molecular diversity as detected using RAPD markers. In pairwise comparisons of RAPD markers, only

two pairs of accessions (CMEN 17 vs. CMEN 19, and CMEN 585 vs. CMEN 584) had >50% of phylogenetically informative markers in common.

The two South African accessions, CMEN 584 and CMEN 585, are remarkably different in appearance from the others. Both have a tall upright growth habit and lanceolate leaves. In addition, the RAPD marker data set these two accessions apart (Figs. 1.2, 1.3). However, despite their morphological similarity and the high number of shared RAPD markers, these two accessions were very different from each other in oil composition and verticillium wilt resistance. Our initial results indicate a need to expand the available germplasm collection to include a broader sampling of the South African representatives of *M. longifolia*.

*M. longifolia* is a suitable and valuable species to serve as a diploid model species for mint genetics for several reasons. The *M. longifolia* genome size in the 400-500 Mbp range is relatively small, making it a favorable subject for structural and functional genomics studies. The C values we obtained for CMEN 585 and CMEN 584 are the first reported for South African genotypes of *M. longifolia*. They are comparable to previously published C value measurements of other NCGR *M. longifolia* accessions (Gobert et al., 2002). Because of the abundant genetic/phenotypic diversity apparent in the species, crosses between appropriately chosen representatives could be used to study the genetic basis for variation in numerous characters of economic relevance. Examples of trait diversity documented here include plant morphology, disease resistance and oil composition. Given the broad geographic range of *M. longifolia*, the species is

likely to contain considerable variation for responses to environmental stress factors as well.

*M. longifolia* is also an intriguing subject for the study of host-pathogen interactions because of its perennial, mostly vegetative growth habit and stem morphology. Replication of screening experiments is facilitated because large numbers of cuttings (clones) may be quickly generated from a single plant. Plants can be maintained in a perpetual vegetative growth state under short-day light regimens, minimizing variation due to hormonal differences between flowering and vegetative growth stages. *M. longifolia* is particularly useful for the study of vascular wilt pathogens because of stem morphology: stems are square, and each stem possesses exactly four vascular bundles—one at each corner—making it possible to observe localized disease symptoms and correlate them to pathogen invasion of particular vascular bundles. In addition, *M. longifolia* has a relatively small (C ≈ 400Mbp) genome, which makes this mint species a good candidate for genomic library construction and map-based gene cloning. Thus, the many favorable features of *M. longifolia* make this species potentially valuable as a diploid model for studies of *Mentha* genetics and for plant pathology genetic studies in particular.

Table 1.1. List of *Mentha longifolia* accessions maintained at NCGR, Corvallis, OR.

| Accession | Status | $2n^z$ | Collected from |
|---|---|---|---|
| CMEN 17 (PI 557755) | Breeding material | 24 | Unknown European country |
| CMEN 18 (PI 557756) | Wild material | 24 | Netherlands |
| CMEN 19 (PI 557757) | Wild material | 24 | France |
| CMEN 20 (PI 557770) | Wild material | 24 | Syria |
| CMEN 34 (PI 557758) | Wild material | ___[y] | India |
| CMEN 500 (PI 212313) | Wild material | 48 | Afghanistan |
| CMEN 501 (PI 212314) | Cultivated material | 48 | Afghanistan |
| CMEN 516 (PI 557760) | Cultivated material | ___[y] | Italy |
| CMEN 584 (PI 557769) | Uncertain improvement status | 24 | South Africa |
| CMEN 585 (PI 557767) | Uncertain improvement status | 24 | South Africa |
| CMEN 592 (PI 557766) | Wild material | 24 | Uzbekistan |
| CMEN 635 (PI 557768) | Wild material | 24 | Nepal |
| CMEN 682 (PI617491) | Cultivar: 'Velvet' | ___[y] | Russia |
| CMEN 81 (PI 557759) | Probable hybrid of *M. longifolia* x *M. spicata* | ___[y] | United States |

[z] Chambers and Hummer (1994).
[y]undetermined chromosome number.

Table 1.2. Phenotypes of *M. longifolia* accessions. Flower color is white (W) or purple (P). Verticillium wilt resistance qualitative ratings are from initial screenings conducted with wild type *Verticillium dahliae* before a numerical rating system was implemented. Qualitative ratings are R=resistant, I=intermediate, S=susceptible. Quantitative ratings are from subsequent screenings of plants used to make crosses. These trials were conducted with a GFP-transformed *V. dahliae* strain. Ratings are average scores for total numbers of plants screened for each genotype. The rating system is: 0=no visible symptoms; 0.5-2.5= mild to moderate symptoms; 3-3.5=severe symptoms; 4=dead. Ratings followed by the same letter are not significantly different from one another (p=0.05). Ratings with different letters are highly significantly different (p<0.01) according to a Tukey's test. Only principal oil compounds (>5%) are listed.

| Accession | Leaf shape | Flower color | Verticillium qualitative | Verticillium quantitative | Oil composition |
|---|---|---|---|---|---|
| CMEN 585 | Lanceolate | W | R | 0[a] | 32.8% pulegone<br>24.3% menthone<br>11.3% 1,8-cineole |
| CMEN 501 | Ovate | P | R | 0[a] | 30.4% pulegone<br>25.3% menthone<br>11.0% menthol<br>5.0% limonene |
| CMEN 81 | Ovate | P | R | 0[a] | 39.2% menthone<br>22.5% iso-menthone<br>8.1% 1,8-cineole |
| CMEN 17 | Ovate | P | R | 0.3[a] | 43.4% trans-piperitone oxide<br>19.7% cis-piperitone oxide<br>7.0% 1,8-cineole |
| CMEN 635 | Ovate | P | R | 1[b] | 45.6% cis-piperitone oxide<br>26.6% piperitenone oxide<br>5.0% trans-piperitone oxide |
| CMEN 34 | Ovate | W | S | 2.0[c] | 14.9% piperitenone oxide<br>6.97% limonene |
| CMEN 682 | Ovate | P | S | 2.6[d] | 56.5% menthol<br>14.8% menthone |
| CMEN 516 | Ovate | P | S | 3.5[e] | 21.9% germacrene D<br>18.6% trans-piperitone oxide<br>11.7% limonene<br>8.0% (Z)-B-ocimene |
| CMEN 584 | Lanceolate | W | S | 3.8[e] | 59.6% carvone<br>12.3% limonene |
| CMEN 18 | Ovate | P | I | | 56.4% trans-piperitone oxide<br>7.2% cis-piperitone oxide<br>5.8% 1,8-cineole |
| CMEN 19 | Ovate | P | S | | |
| CMEN 20 | Ovate | P | R | | 13.5% pulegone<br>11.7% nonanal<br>7.8% menthone<br>7.0% trans-piperitone oxide<br>6.6% limonene |
| CMEN 500 | Ovate | P | R | | 34.6% pulegone<br>17.0% menthone<br>14.2% sabinene<br>6.1% limonene |
| CMEN 592 | Ovate | P | S | | 22.4% (E)-B-farnesene<br>16.0% iimonene<br>12.7% nonanal<br>11.0% B- caryphyllene<br>7.4% gamma-muurolene |

Figure 1.1. Example of electrophoretic banding pattern of PCR-amplified DNA fragments produced from RAPD primer OPO20. Lane 1 is a molecular size marker. Lanes 2-14 are CMEN 17, CMEN 18, CMEN 19, CMEN 20, CMEN 34, CMEN 516, CMEN 592, CMEN 500, CMEN 501, CMEN 635, CMEN 682, CMEN 585 and CMEN 584, respectively.

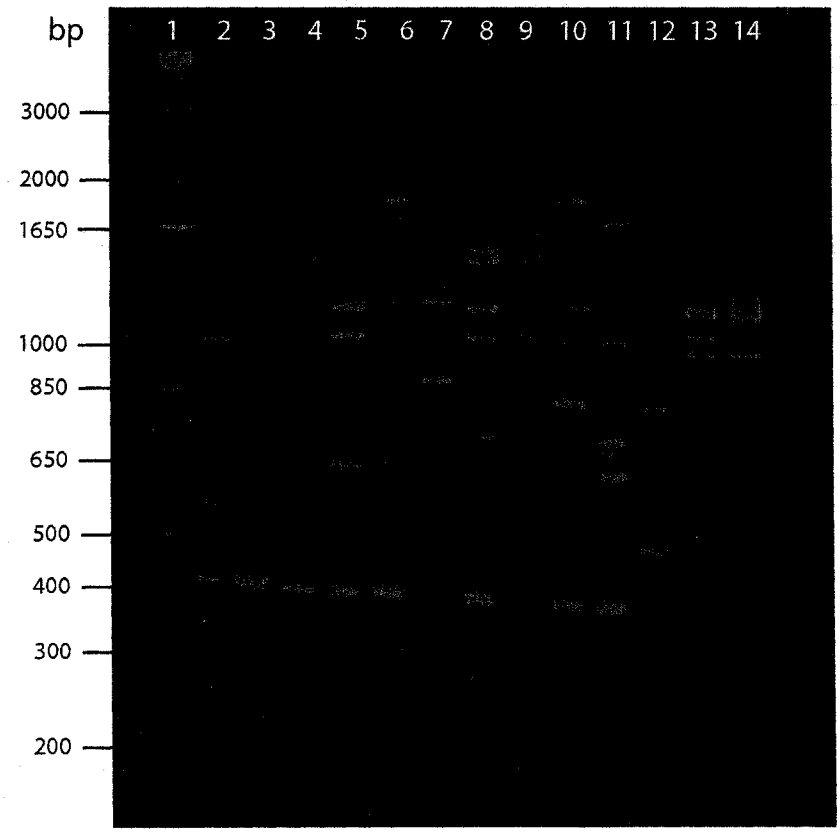
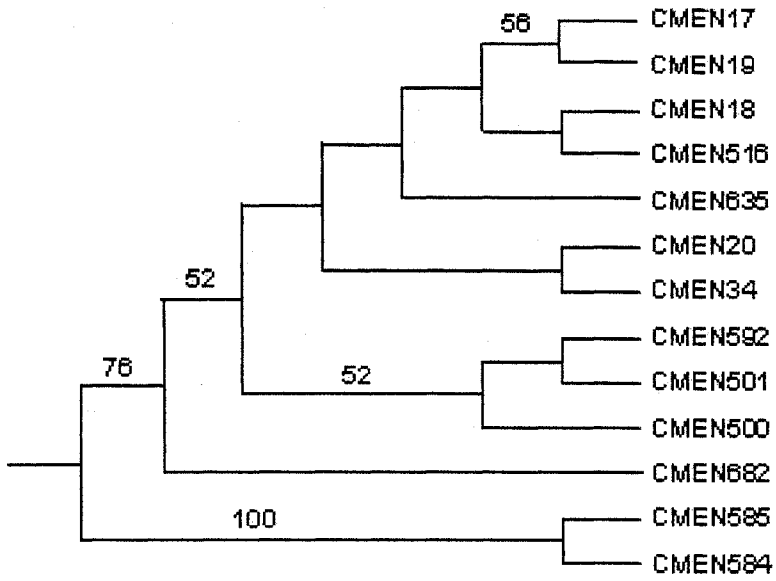


Figure 1.2. Jaccard similarity coefficient matrix. Values above the diagonal are the number of bands shared by each pair of accessions. Values below the diagonal are Jaccard similarity coefficients.

| | CMEN 17 | CMEN 18 | CMEN 19 | CMEN 20 | CMEN 34 | CMEN 516 | CMEN 592 | CMEN 500 | CMEN 501 | CMEN 635 | CMEN 682 | CMEN 585 | CMEN 584 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CMEN 17** | ___ | 9 | 12 | 7 | 7 | 9 | 8 | 3 | 4 | 7 | 3 | 2 | 0 |
| **CMEN 18** | 0.4286 | ___ | 8 | 7 | 5 | 9 | 8 | 2 | 2 | 7 | 1 | 1 | 0 |
| **CMEN 19** | 0.5217 | 0.3478 | ___ | 11 | 6 | 9 | 10 | 4 | 3 | 8 | 3 | 1 | 0 |
| **CMEN 20** | 0.2414 | 0.2800 | 0.4231 | ___ | 7 | 7 | 9 | 3 | 1 | 6 | 0 | 4 | 2 |
| **CMEN 34** | 0.3043 | 0.2381 | 0.2400 | 0.2800 | ___ | 5 | 7 | 2 | 4 | 3 | 1 | 1 | 1 |
| **CMEN 516** | 0.3750 | 0.4500 | 0.3600 | 0.2500 | 0.2083 | ___ | 8 | 4 | 3 | 5 | 1 | 2 | 1 |
| **CMEN 592** | 0.2857 | 0.3226 | 0.2727 | 0.2414 | 0.2581 | 0.2857 | ___ | 9 | 10 | 5 | 2 | 2 | 1 |
| **CMEN 500** | 0.0968 | 0.0714 | 0.1290 | 0.0909 | 0.0714 | 0.1379 | 0.2903 | ___ | 7 | 2 | 1 | 1 | 2 |
| **CMEN 501** | 0.1290 | 0.0690 | 0.0909 | 0.0278 | 0.1481 | 0.0968 | 0.3226 | 0.2500 | ___ | 4 | 3 | 1 | 2 |
| **CMEN 635** | 0.3043 | 0.3684 | 0.3478 | 0.2308 | 0.1304 | 0.2083 | 0.1613 | 0.0714 | 0.1481 | ___ | 2 | 1 | 0 |
| **CMEN 682** | 0.1500 | 0.0556 | 0.1429 | 0.0000 | 0.0556 | 0.0476 | 0.0741 | 0.0455 | 0.1429 | 0.1176 | ___ | 0 | 0 |
| **CMEN 585** | 0.0588 | 0.0323 | 0.0278 | 0.1176 | 0.0323 | 0.0606 | 0.0500 | 0.0286 | 0.0278 | 0.0667 | 0.0000 | ___ | 16 |
| **CMEN 584** | .0000 | 0.0000 | 0.0000 | 0.0571 | 0.0333 | 0.0303 | 0.0250 | 0.0606 | 0.0588 | 0.0333 | 0.0000 | 0.7619 | ___ |

Figure 1.3. RAPD marker-based, midpoint-rooted UPGMA dendrogram showing relationships of *Mentha longifolia* accessions. Numbers above branches are bootstrap values.



CHAPTER II.

# IDENTIFICATION OF RESISTANCE GENE ANALOGS AND VERTICILLIUM WILT RESISTANCE-LIKE SEQUENCES IN *MENTHA LONGIFOLIA* (L.) HUDS[1].


## Abstract

Resistance gene analog (RGA) sequences were obtained from four *Mentha longifolia* (L.) Huds. accessions using degenerate PCR primers targeting the conserved nucleotide binding site (NBS) domain found in many plant disease resistance genes. Seven distinct RGA families were identified. All *M. longifolia* RGAs showed similarity to sequences of the non-TIR (Toll-Interleukin 1 Receptor) *R* gene class. In addition, degenerate PCR primers based on the tomato (*Solanum lycopersicum* L.) verticillium wilt resistance (*Ve*) genes were used to amplify a 445 base pair (bp) *Ve*-like sequence from *M. longifolia* that had 56% to 57% predicted amino acid identity with *Ve*. Mint-specific primers based on the original mint *Ve* sequence were used to obtain mint-specific *Ve* sequences from four *M. longifolia* accessions, and from peppermint (*Mentha ×piperita* L.) cultivar Black Mitcham that had 95% to 100% predicted amino acid identity to the original mint *Ve* sequence. Inverse PCR was then utilized to obtain flanking mint *Ve* sequence from one *M. longifolia* accession, extending the mint *Ve* sequence to 1077 bp. This chapter


[1] This chapter is revised from an article accepted for publication by the Journal of the American Society for Horticultural Science. Vining, K.J., Q. Zhang, C. Smith, T.M. Davis. Identification of resistance gene analogs and verticillium wilt resistance-like sequences in *Mentha longifolia* (L.) Huds.

is the first report of RGA sequences in the Lamiaceae, and the first report of *Ve*-like sequences obtained with degenerate PCR primers.

## Introduction

We are using *M. longifolia* (Lamiaceae) as a diploid model species with relevance to the polyploid commercial mints (*Mentha* L. sp.) and to the study of plant resistance to vascular wilt diseases. Verticillium wilt, caused by the fungus *Verticillium dahliae* Kleb., is the most damaging disease of commercial peppermint (Green, 1951; Lacy and Horner, 1965). With the goal of isolating genetic determinants of verticillium wilt resistance, we have identified highly resistant and susceptible *M. longifolia* accessions from the USDA National Clonal Germplasm Repository (NCGR) collection (Vining et al., 2005), and have used crosses between them to develop progeny populations segregating for resistance vs. susceptibility. As reported here, we have also exploited the availability of conserved motifs from resistance genes and resistance gene-like sequences from other species to isolate candidate resistance gene sequences in *M. longifolia*.

Our approach exploits the rapidly growing knowledge of disease resistance (*R*) genes and their commonalities in many crop and model plant species. The known or predicted protein products of most *R* genes described in plants to date encode one or two characteristic domains: a highly conserved ~300 amino acid nucleotide binding site (NBS), and/or a highly variable leucine-rich repeat (LRR) region. The NBS has been implicated in defense-related cell signaling, while the latter is thought to be involved in pathogen recognition. Plant *R* genes are classified

on the basis of their various combinations of NBS and LRR with receptor kinase, transmembrane, and other domains. While many LRR or LRR-receptor kinase proteins are predicted or demonstrated to have roles other than plant defense, the NBS-LRR proteins have been implicated solely in plant disease resistance (Meyers et al., 2005). The *Arabidopsis thaliana* (L.) Heynh. genome includes 149 NBS-LRR sequences (Meyers et al., 2005).

The NBS-LRR superclass can be divided into two classes based on motifs within the ~200 amino acid region upstream of the NBS: the Toll-Interleukin Receptor-like (TIR-NBS-LRR) class, and the non-TIR class, which has a coiled-coil domain (CC-NBS-LRR). The TIR domain has homology to the *Drosophila melanogaster* Toll receptor and the mammalian interleukin-1 receptor, both of which are involved in immune responses (Jebanathirajah et al., 2002). TIR and non-TIR classes of *R* genes can be distinguished by amino acids within the NBS. The TIR class has the amino acid motifs RNBS-A-TIR (LQKKLLSKLL) and RNBS-D-TIR (FLHIACFF), while the non-TIR class has RNBS-A-non-TIR (FDLxAWVCVSQxF) and RNBS-D-non-TIR (CFLYCALFPED) (Meyers et al., 1999). In addition, the presence of tryptophan (W) or aspartic acid (D) at the final amino acid position of the Kin-2 motif distinguishes the non-TIR and TIR classes, respectively, with 95% accuracy (Meyers et al., 1999).

The NBS itself, a protein domain represented in both prokaryotes and eukaryotes, contains highly conserved motifs. The most common NBS motif is the 'P-loop/Kinase 1a' domain, involved in binding ATP or GTP (Saraste et al., 1990; Traut, 1994). NBS domains of plant *R* genes include the P-loop, as well as Kin-2,

RNBS-A-D, and GLPL motifs (Meyers et al., 1999). *R* gene analogs (RGAs) are sequences sharing these conserved NBS motifs. Such NBS motifs were first exploited as target sites for PCR with degenerate primers for the purpose of isolating families of RGAs by Kanazin et al. (1996), Leister et al. (1996), and Yu et al. (1996). Degenerate primers have since been used extensively to identify RGA sequences and candidate *R* genes in numerous angiosperms (Brotman et al., 2002; Cordero and Skinner, 2002; Irigoyen et al., 2004; Lopez et al., 2003; Maleki et al., 2003; Martinez Zamora et al., 2004; Noir et al., 2001; Shen et al., 1998; Soriano et al., 2004; Yaish et al., 2004), gymnosperms in the *Pinus* L. genus (Diaz and Ferrer, 2003; Liu and Ekramoddoullah, 2003), and even the bryophyte *Physcomitrella patens* (Hedw.) B.S.G. (Akita and Valkonen, 2002). RGAs are one of two categories of mint candidate gene sequences acquired using degenerate primers in the present study; the second category is putative homologues of the tomato *Ve* (verticillium wilt resistance) gene, as described below.

*V. dahliae* infects a large number of important crop plants, including tomato, potato (*Solanum tuberosum* L.), pepper (*Capsicum* L.), cotton (*Gossypium* L.), strawberry (*Fragaria* L.), lettuce (*Lactuca* L.), and melon (*Cucumis* L.) (Bhat and Sabbarao, 1999) as well as trees and woody and herbaceous ornamental crops (Sinclair et al., 1987; Smith and Neely, 1979). Verticillium wilt results from fungal invasion of the plant vascular system. For some plant hosts, variability in resistance has been identified in germplasm, but monogenic determinants of resistance have not been reported. The single exception is tomato, for which single-gene verticillium wilt resistance was identified in classical genetics experiments. The gene was

designated *Ve* (Schaible et al., 1951). This resistance has been bred into many modern tomato cultivars and has proven durable. The *Ve* locus was eventually mapped to tomato linkage group IX (Diwan et al., 1999). Positional cloning then determined that this locus consisted of two functional genes, encoding products with 84% amino acid identity, designated *Ve*1 and *Ve*2 (Kawchuk et al., 2001).

The *S. lycopersicum* verticillium wilt resistance genes *Ve1* and *Ve2* (to be collectively referred to as *Ve*) are predicted to encode cell surface receptor-like proteins that lack an NBS, but have 38 LRRs, an N-terminal leucine-zipper motif, and endocytosis-like signals (Kawchuk et al., 2001). Transfer of *Ve* transgenes to susceptible tomato and potato lines has conferred resistance to race 1 of *V. dahliae* (Kawchuk et al., 2001). Library probes and PCR primers specific to the tomato *Ve* sequences have been used to identify homologous sequences from other solanaceous species (Chai et al., 2003; Simko et al., 2004a,). Outside the Solanaceae, only a few sequences with homology to *Ve1* and *Ve2* have been listed in the National Center for Biotechnology (NCBI) GenBank® database, none of which have been shown to play a role in verticillium wilt resistance. However, the cloning of the tomato *Ve* genes has created the opportunity to use their sequences and those of possible homologues in other species as a basis for design of degenerate primers aimed at isolating *Ve* homologues in mint.

Here, we report the use of degenerate primer PCR for identification of two types of disease resistance-related sequences in *M. longifolia*: i) RGAs, which may represent a variety of disease specificities; and ii) sequences homologous to the known verticillium wilt resistance (*Ve*) genes of tomato.

## Materials and methods

### Plant and fungal material

*M. longifolia* accessions were obtained as stolons or rooted cuttings from the U.S. Department of Agriculture, National Clonal Germplasm Repository (USDA-NCGR), Corvallis, OR. Plants were maintained in a greenhouse at the University of New Hampshire (Durham) in 22-cm-diameter pots and propagated vegetatively. Accessions CMEN17.001 (PI 557755), CMEN585.001 (PI 557767), CMEN501.001 (PI 212314) and CMEN81.001 (PI 557759) were previously classified as verticillium wilt-resistant, while accessions CMEN584.001 (PI 557769) and CMEN516.001 (PI 557760) were classified as wilt-susceptible (Vining et al., 2005). For convenience, the ".001" suffix common to all of these CMEN accession numbers is dropped throughout the subsequent text. Peppermint cultivar Black Mitcham, classified as verticillium wilt-susceptible (Douhan and Johnson, 2001), was also utilized. The accessions used in this study are diploid (2n=2x=24), with the exceptions of CMEN501 (tetraploid: 2n=4x=48), CMEN81 (unknown ploidy), and 'Black Mitcham' (hexaploid: 2n=6x=72) (Chambers and Hummer, 1994). The *V. dahliae* culture used to generate inoculum was provided by D. Johnson at Washington State University (Pullman, WA).

### Molecular techniques

Genomic DNA was extracted from fresh, unexpanded leaf tissue using a modified CTAB miniprep method as described in Vining et al. (2005). For reverse transcription-polymerase chain reaction (RT-PCR), template RNA was obtained from

roots of plants that had been incubated in either water or an aqueous *V. dahliae* spore suspension (~$10^6$ spores/mL) for 16-24 h. Total RNA was isolated from 100-200 mg fresh or frozen roots using Trizol® reagent according to manufacturer's instructions (GIBCO, Carlsbad, CA). Messenger RNA (mRNA) was separated from total RNA using oligo d(T) cellulose columns (GIBCO).

PCR conditions are listed in Table 2.1. Reaction volumes of 25 µl contained 100 ng template DNA, 0.1 U *Taq* DNA polymerase (Sigma, St. Louis, MO), 2.5 mM each dNTP (Promega Corporation, Madison, WI) and either 4.8 µmol of each degenerate primer or 0.8 µmol of each specific primer. RT-PCR was performed with the ProSTAR® HF Single-Tube RT-PCR System (Stratagene, La Jolla, CA). Inverse PCR was used to obtain sequences flanking the mint *Ve* segment initially identified by degenerate primer PCR. One microgram of total genomic DNA from CMEN 585 was digested overnight at 37 °C in a 20 µL reaction with 10 U of the restriction enzyme *Mse*I (New England Biolabs, Ipswich, MA.). Two microliters of the digest were used directly in a ligation reaction with 400 U T4 DNA ligase (New England Biolabs) to circularize the genomic restriction fragments. The ligation reaction was incubated for 2.5 h at room temperature, then stored at 4 °C overnight. PCR was performed in 25 µL reactions using 5 µL of the ligation products as template, 0.1 U *Taq* polymerase (Sigma), 2.5 mM of each dNTP (Promega), and 0.8 µmol each inverse PCR primer.

Widely employed degenerate PCR primers targeted to the P-loop motif (forward primer s2) and hydrophobic domain (reverse primers as2 or as3) (Leister et al., 1996) were used for initial amplification of RGAs from genomic DNA (gDNA) and

from cDNA produced by RT-PCR (Table 2.2). Degenerate primers intended to amplify *Ve* homologues in mint were designed manually, on the basis of tomato *Ve*1 amino acid and nucleotide sequences aligned with the most similar sequences from *A. thaliana* and *Oryza sativa* L., obtained by searching GenBank®'s non-redundant (nr) database using the blastx algorithm (Altschul et al. 1990), submitting the tomato *Ve*1 amino acid sequence as a search query. Subsequently, specific (non-degenerate) primers intended to amplify segments of putative mint *Ve* homologues, or a 425 bp segment of the tomato *Ve*1 gene itself (as a positive control), were designed using PrimerSelect software version 5.53 (DNASTAR Inc., Madison, WI). The tomato-specific forward and reverse primers, Ve5F19 and Ve5R23, bracketed *Ve*1 nucleotides 1776-1794 and 2200-2178, respectively (Table 2.2).

PCR, RT-PCR and IPCR products were electrophoretically separated on 2% agarose gels in 1X TBE, pH 8.0 at 4.5 V/cm at room temperature. Gels were stained with ethidium bromide and photographed over UV light. Amplification products were recovered from gel slices using the Geneclean® Spin kit (QBiogene, Irvine, CA) and cloned with the TOPO TA Cloning® kit (Invitrogen, Carlsbad, CA). Recombinant plasmids were isolated using the Wizard® SV Plasmid Purification kit (Promega).

Cycle sequencing reactions were performed following the University of New Hampshire's Hubbard Center for Genome Studies protocol (Hubbard Center for Genome Studies, 2002) using DYEnamic™ Terminator ET cycle sequencing premix (Amersham Biosciences, Piscataway, NJ) and standard plasmid vector primers

M13F, M13R, T3 and T7. Reaction products were analyzed using an ABI PRISM® 377 DNA sequencer (Applied Biosystems, Foster City, CA).

## Sequence analysis

DNA sequence chromatograms were viewed and assessed for base-calling accuracy using SeqEd version 1.0.3 (DNASTAR, Madison, WI). Similarity searches of the GenBank® nr database were conducted using blastx and tblastx algorithms (Altschul et al., 1990). Clustal X (Thompson et al., 1997) was used to perform sequence alignments, and to generate an amino acid sequence identity matrix of the RGAs, including top non-mint RGA matches from the GenBank® nr database (RGA blast hits). A Clustal X alignment of translated mint RGAs and top blastx hits was used to construct a neighbor-joining tree with 1000 bootstrap replicates.

## Genome library construction

A genomic library was constructed in a fosmid vector using DNA from wilt-resistant accession CMEN 585. Methods for library construction are described in Appendix B.

# Results

## RGAs

The insert sizes of 52 PCR product clones obtained from genomic DNA of four verticillium wilt-resistant *M. longifolia* accessions using degenerate RGA primers (s2 + as2, or s2 + as3) ranged from 289-680 bp. When these sequences were

subjected to blastx and tblastx searches of the GenBank® nr database, 20 of them had best matches (i.e. blast hits) that were *R* genes and/or RGAs. These *M. longifolia* genomic RGA sequences will be referred to as gRGAs (Table 2.3).
Of the 20 gRGA sequences acquired (seven from CMEN585, six from CMEN17, three from CMEN81, and one from CMEN501), 17 were found to be unique after discounting sequence redundancy. One gRGA sequence was represented by three identical clones: two from CMEN585 (clones 5856 and 58531), and one from CMEN17 (clone 1744). Another gRGA was represented by two identical clones, 1712 and 1725, both from CMEN17.

All but three of the unique gRGA sequences consisted of uninterrupted open reading frames (ORFs). Clone 5857 had an RGA reading frame that was interrupted by one stop codon [position 271-273 (TAG)], and clone 171 had an RGA reading frame interrupted by two stop codons [positions 229-231 (TAA) and 358-360 (TAA)]; otherwise both of these interrupted RGA sequences were continuous ORFs. gRGA clone 173 had an RGA sequence that shifted reading frame forward one nucleotide at a point 245 bases into the sequence, dividing two nearly continuous ORFs that were each interrupted by a single stop codon [positions 79-81 (TGA), 386-388 (TGA)].

The s2 + as2 degenerate primer combination was also used in RT-PCR with mRNA from both inoculated and non-inoculated roots. Following RT-PCR, the most densely staining area of the gel lane (in the 250-500 bp range), or a discrete gel band, if present, was excised for cDNA recovery. Among 23 sequenced cDNA clones, seven RGAs were identified (Table 2.3). These cDNA-derived RGAs will be

referred to as cRGAs. All seven cRGAs were comprised of uninterrupted ORFs. Of the cRGAs from non-inoculated root tissue, one, two, and three cRGAs, respectively, were from CMEN17, CMEN585, and CMEN81. Clone 17B9 from CMEN17 was the only cRGA recovered from inoculated roots. The sequence of cRGA clone 1719 was identical to that of gRGA clone 58513 from CMEN585.

Of 31 genomic and 16 cDNA non-RGA sequences obtained using the degenerate RGA primer pairs, one cDNA and four genomic sequences had ABC transporter sequences as best blast hits to the nr database, and three cDNA and two genomic sequences' best nr database blast hits were gag-pol or Ty3/gypsy retroelements. Among the other cDNA sequences that had uninterrupted ORFs, blast matches included one chlorophyll a/b binding protein, one L29 ribosomal protein, and two different 'senescence-associated proteins'. The rest of the non-RGA genomic and cDNA clones lacked uninterrupted ORFs and had no informative blast hits.

Among a total of 27 gRGA and cRGA sequences acquired using the degenerate RGA primer pairs there were 23 unique RGA sequences, all of which have been submitted to GenBank® (under accession numbers shown in Table 2.3). All 23 unique RGAs contained the RNBS-A non-TIR motif, and 20 of them had tryptophan in the terminal position of the Kin-2 motif, which is characteristic of the non-TIR RGAs. When each of the 23 unique mint RGAs was used in blastx and tblastx searches of the GenBank® nr database, the best blast hits were always mint RGAs from the present data set; but many very good non-mint blast hits were also

obtained. In many instances, the same best non-mint blast hit was obtained for two or more mint RGAs.

The 23 unique mint RGAs and a selection of 14 GenBank® sequences chosen to represent the diversity of top non-mint blast hits were subjected to a Clustal X analysis that generated an amino acid sequence alignment (Fig. 2.1) and identity matrix (not shown). Percent amino acid sequence identities of pairs of unique mint RGAs ranged from 18% to 96%. From the Clustal X alignment, a neighbor-joining tree was generated which organized the 23 mint RGA sequences and best non-mint blast hits into three major groups (I, II, III) that were further subdivided into seven distinct families (A through G) (Fig. 2.2). Percent amino acid identities of mint sequences within families ranged from 48% to 96%. The two most similar pairs of sequences were in family E, where gRGA sequences 501AS38A and cRGA 81CRGC4 were 96% identical, and in family B, where gRGA sequences 5856 and 178 were 89% identical. Mint RGAs in family G were most similar to *Phytophthora infestans* (Mont.) de Bary *R* genes from *Solanum* L. species, while mint RGAs in families C and D were most similar to potato and pepper *R* genes with known pathogen specificities (Table 2.4). For mint RGAs in families B, E and F, representatives from several angiosperm species were included among the top non-mint blast hits (Table 2.4). Top non-mint blast hits for mint family A were RGAs from *Coffea* L. species.

***Ve***

When a blastx search of the GenBank® non-redundant data base was performed using the tomato *Ve1* amino acid sequence as a search query, the most similar sequences (30% to 40% amino acid identities) were those from *A. thaliana* (NP_187712, NP_198058, NP_180861, NP_187217 ) and *O. sativa* (BAA96753, BAB08209). Using both amino acid and nucleotide alignments of the tomato *Ve1* and *Ve*2 sequence with the aforementioned *A. thaliana* and *O. sativa* sequences as a basis for identifying conserved regions, degenerate primers Ve2398-2416F and Ve2791-2810R (Table 2.2) were designed manually. These primers, which targeted sites approximately corresponding to codons 800-805 and 942-948 of tomato *Ve1*, amplified a 445 bp product from CMEN585, the sequence of which we designate the "original mint *Ve*". Excluding from consideration the amino acids coinciding with the primer sites, the original mint *Ve* predicted amino acid sequence shared 57.1% and 56.5% amino acid identity with the corresponding regions of the tomato *Ve1* and *Ve2* sequences, respectively (Fig. 2.3).

On the basis of the original mint *Ve* sequence, mint-specific primers (Ve4-24F and Ve403-386R, Table 2.2) internal to the degenerate primers sites were designed and tested on four *M. longifolia* accessions (CMEN numbers 17, 516, 584, 585) and peppermint cultivar Black Mitcham, in each case producing "second generation" mint *Ve* sequences of about ~400 bp that had 94.7-100% predicted amino acid identities to the original mint *Ve*. The predicted amino acid sequence of the second generation mint *Ve* product from CMEN585 is shown (as "Mint 2") aligned with the original mint *Ve* (Fig. 2.3).

Also on the basis of the original mint *Ve* sequence, inverse PCR primers (IPCR138-153F, IPCR325-304R, Table 2.2) were designed and used to obtain extended *Ve* sequence from CMEN585 that included and flanked the original degenerate primer sites. The IPCR results extended the original *Ve* sequence by adding 202 and 431 bases, respectively, to its 5' and 3' ends (Mint IPCR, Fig. 2.3). Examination of the IPCR sequence revealed that the actual mint sequence at the site of the initial *Ve* degenerate primers was perfectly matched by one permutation of each of the original *Ve* degenerate PCR primer sequences, and that the second generation "mint specific" primer Ve4-24F had a two-base mismatch with the actual CMEN585 target sequence (Table 2.2). The extended mint *Ve* open reading frame ended with three sequential stop codons [TAA, TGA, TGA] at nucleotides 1013-1021. Nucleotides 1022-1077 were presumed to be part of the 3' untranslated region.

Finally, the extended mint *Ve* sequences provided a basis for design of specific PCR primers (Ve111-133F and Ve974-951R, Table 2.2), which bracketed 865 nucleotides of the extended mint *Ve* sequence. The predicted amino acid sequence of the extended mint *Ve* aligned with the C terminus of the tomato sequences, spanning amino acids 730-1053 and 728-1079 of *Ve1* and *Ve2*, respectively. This section of the tomato sequences included the following regions defined by Kawchuk et al. (2001): LRRs 28-38 (end of region B), negatively-charged extracytoplasmic (C), membrane-spanning (D), and positively-charged cytoplasmic (E). While the extended mint *Ve* sequence shared 56% to 57% amino acid identity over the corresponding B region of the tomato *Ve* sequences, the mint sequence

shared only 7/21 and 8/21 amino acid identities with the D regions of *Ve*1 and *Ve*2, respectively. The C and E regions of the mint and tomato *Ve* sequences could not be meaningfully aligned.

## Discussion

### RGAs

We used degenerate primers targeting the NBS motif in genomic PCR and RT-PCR to identify a set of 23 unique RGA sequences from *M. longifolia*, 20 of which had uninterrupted ORFs. Best blast hits for these RGAs included angiosperm *R* genes with a variety of known disease specificities. The 40% to 50% range of amino acid identities to known *R* genes compares to that seen for RGAs identified using the same degenerate primers in other plant taxa, such as strawberry (Martinez Zamora et al., 2004) and coffee (*Coffea arabica* L.) (Noir et al., 2001). The presence of uninterrupted ORFs in 20 of the 23 unique RGAs is suggestive that at least some of these mint RGAs are derived from functional genes. One gRGA clone (1719) from CMEN17 was identical to a cRGA (clone 58513) from CMEN585, indicating that this particular RGA is both highly conserved and is expressed, strongly suggesting that it may encode a functional product involved in disease resistance.

Two of the 23 *M. longifolia* gRGA sequences had stop codons in otherwise uninterrupted ORFs, and one gRGA contained a frameshift that separated two nearly continuous ORFs. Similarly, Martinez Zamora et al. (2004) identified 51 RGAs in strawberry, of which 28 had uninterrupted ORFs and 23 had stop codons or

frameshifts. In pine (*Pinus monticola* Dougl. ex. D. Don.), Liu and Ekramoddoullah (2003) found that 18 out of 46 genomic RGA clones and 14 out of 21 cDNA RGA clones had at least one stop codon or frameshift. RGAs with interrupted ORFs may be parts of pseudogenes. Plant *R* genes have been shown to occur in complex clusters of related functional genes, pseudogenes, gene fragments, and transposable elements (Lescot et al., 2004; Wei et al., 2002). Whether pseudogenes are evolutionary remnants – once-active genes rendered inactive by accumulation of mutations – or whether they have some function, regulatory or otherwise, remains to be determined on a case-by-case basis.

In two instances, identical RGA sequences were amplified from different accessions. The gRGA clone 1744 from CMEN17 was identical to gRGA clones 5856 and 58513 from CMEN585, while gRGA clone 1719 from CMEN17 was identical to cRGA clone 58513 from CMEN585. The amplification of identical RGAs from different accessions suggests that the respective pairs of sequences come from orthologous genes. Four of the mint gRGAs obtained using degenerate primers were represented by multiple clones from the same accession. One pair of identical clones (1725, 1712) was from CMEN17, and another (5856, 58531) was from CMEN585. These identities are probably the consequence of redundant sampling of the same gene. The recoveries of non-identical RGA sequences from a single accession may result from sampling of different genes or different alleles of the same gene. The closest nucleotide sequence match between non-identical gRGAs within an accession was 79.8% (clones 8120 and 8116), which is low enough to suggest that these could be different genes.

NBS-LRR genes comprise significant proportions of plant genomes. Whole genome sequences are available for only two plants: *A. thaliana* ecotype Col-0 and *O. sativa* ssp. *japonica* L. cv. Nipponbare. The *A. thaliana* genome of ~157 Mb (Bennett and Leitch, 2005) contains 149 NBS-LRR genes (55 non-TIR, 94 TIR), representing ~1% of the genome (Meyers et al., 2005).The *O. sativa* haploid genome size is estimated to be 389 Mb (International Rice Genome Sequencing Project, 2005). *O. sativa* has ~500 NBS-containing RGA sequences, all of which are of the non-TIR type, and which, like *A. thaliana*, represent ~1% of the genome (Monosi et al., 2004). The *M. longifolia* genome size of ~400 Mb is comparable to that of rice (*Oryza sativa* L.) (Bennett and Leitch, 2005). If the total number of *M. longifolia* RGAs parallels the ~500 of rice, then the 23 RGAs reported here represent <5% of the potential *M. longifolia* RGAs. The low level of redundancy encountered among the *M. longifolia* RGAs is also indicative that these sequences represent a small subset of the RGAs in the mint genome.

Studies of resistance gene clusters have revealed evidence of duplication, recombination and rearrangement events leading to the expansion of clusters, growth of multigene families, and the potential for evolution of new resistance specificities (Chin et al., 2001; Graham et al., 2002; Kruijt et al., 2004). The 23 *M. longifolia* RGAs grouped into seven distinct families, each of which had identities with different RGAs and *R* genes from other plant species. Particular amino acid motifs conserved within mint RGA families may indicate some functional specificity. However, closely related *R* genes do not necessarily confer resistance to the same pathogen. The potato *Gpa2* and *Rx1* genes are in the same cluster on potato

chromosome 12 and are considered to be highly homologous, but have very different specificities; *Gpa2* is a nematode (*Caenorhabditis elegans* Miers) resistance gene, while *Rx1* confers resistance to a virus (van der Vossen et al., 2000). The *M. longifolia* RGA families C and D each have significant similarity to *Gpa2* and *Rx1*. It would be interesting to determine whether family C and D mint RGAs map to a common genomic region or cluster, as their apparent counterparts do in potato.

All the *M. longifolia* RGAs amplified with degenerate primers based on the P-loop and GLPL domains, or amplified with mint RGA-specific primers, had motifs consistent with the non-TIR class. This result does not lead to a conclusion that TIR-class RGAs are absent from the *M. longifolia* genome. In general, dicot species have been found to contain varying ratios of TIR and non-TIR classes of RGAs, while in monocot species only non-TIR RGA sequences have been observed (Cannon et al., 2002). However, dicot RGA sequences obtained from degenerate primer PCR have not given any indication that non-TIR and TIR classes are amplified in proportions that reflect their presence in plant genomes. The same P-loop and GLPL degenerate primers used in the present study, or very similar primers, amplified both TIR and non-TIR RGAs in alfalfa (*Medicago* L.) (Cordero and Skinner, 2002), but only non-TIR sequences from coffee (Noir et al., 2001). In soybean [*Glycine max* (L.) Merr.], where Kanazin et al. (1996) found only TIR-class RGAs using P-loop and GLPL-based primers, Peñuela et al. (2002) amplified non-TIR RGAs with a set of degenerate primers pairing a P-loop forward primer with a reverse primer based on the non-TIR RNBS-D motif (CFLYCALFPED). It is possible that the P-loop/GLPL

degenerate primer set is biased toward TIR-class or non-TIR-class sequences in a species-dependent manner. The fact that only one RGA was obtained using the as3 reverse primer may also reflect primer bias.

Although the isolation of non-RGA sequences among the PCR products generated by degenerate RGA primers seems inevitable, many RGA reports do not mention whether any non-RGA sequences were obtained with degenerate RGA primers. In the present study, two types of non-RGA sequences were common. Use of degenerate RGA primers generated seven PCR products that had high identity to transposable element-like sequences. Sequences with identity to transposable elements have been mentioned in a few RGA reports (Noir et al., 2001; Timmerman-Vaughan et al., 2000). Five of our mint degenerate primer PCR products had high identity to ABC transporter sequences. Joyeux et al. (1999) amplified an ABC-transporter-like sequence from *Brassica napus* L. using degenerate primers nearly identical to s2 and as2. Amplification of ABC transporter sequences with NBS degenerate primers is not surprising, since ABC transporters have an NBS motif similar to that of NBS-LRR disease resistance genes (Theodoulou, 2000). If the isolation of non-RGA sequences was detailed in more reports of degenerate RGA primer studies, potentially useful insight could be gained into commonalities and patterns of occurrence of such sequences in different species.

***Ve***

Since *Ve* is not an *R* gene of the NBS-LRR class, it would not be expected to be amplified by NBS-targeted degenerate primers such as those used to amplify

RGAs. We used degenerate primers targeting part of the tomato *Ve* LRR region to isolate an original *Ve*-like *M. longifolia* sequence of 445 bp from the verticillium wilt-resistant accession CMEN585. The 56% to 57% amino acid identity of this sequence to the corresponding regions of the tomato *Ve*1 and *Ve*2 sequences is higher than that for any other *Ve*-like sequence found in GenBank®. Design of mint-specific IPCR primers enabled us to extend the acquisition of mint *Ve* sequence to a 1077 bp segment, that corresponds to the distal B region and the C,D, and E regions of the tomato *Ve* genes as described by Kawchuk et al. (2001). To date, this is the first reported use of degenerate PCR primers to isolate a *Ve*-like sequence, and the first targeted isolation of such a sequence outside of the Solanaceae.

Specific primers based on the original mint *Ve* sequence detected *Ve* homologues in several wilt-resistant and wilt-susceptible *M. longifolia* accessions, and from wilt-susceptible cultivar Black Mitcham, indicating that this primer pair and others based upon the extended *Ve* sequence will be a valuable resource for use in a broader assessment of the diversity of *Ve* sequences among *Mentha* accessions and species.

Thus far, no verticillium wilt-related phenotypes have been associated with any of the *Ve*-like sequences reported in *O. sativa* or *A. thaliana*. The potential rice *Ve* homologues have been listed as such in GenBank®, based on inference, as a consequence of the annotation of the *O. sativa* genome, but monocot species are not known to be hosts for *Verticillium* fungi. To date, no complementation studies have been performed to elucidate any connection between rice *Ve* homologues and non-host resistance to verticillium wilt. For *A. thaliana*, differences in symptom

development such as stunting and early flowering have been observed, but no verticillium wilt-resistant ecotypes have been identified (Steventon et al., 2001; Veronese et al., 2003). Veronese et al. (2003) identified a locus (but not a gene) that was correlated with less severe symptom development in ecotype C-24, but they noted that disease development was difficult to parse from normal plant developmental and senescence features. Distinctly verticillium wilt-resistant and wilt-susceptible *M. longifolia* accessions have been described (Vining et al., 2005), making this a much more appropriate species than either *O. sativa* or *A. thaliana* for identification and characterization of *Ve* homologues and assessment of their possible roles in conferring verticillium wilt resistance. As yet, the available data do not permit association of the mint *Ve*-like sequences with resistant or susceptible phenotypes.

The isolation of candidate resistance genes such as RGAs provides a useful foundation for further study of the genetics and genomics of plant disease resistance in mint. This study also describes the development of novel degenerate primers and their use in the isolation of potential *Ve* homologues in mint. The present work is the first report of *R* gene-like sequences in *Mentha* and in the Lamiaceae, and the first to report the targeted isolation of *Ve* homologues outside of the Solanaceae. This work advances the development of *M. longifolia* as a model system for disease resistance in perennials and plants in general, and constitutes a step toward identification of associations between candidate *R* genes and phenotypically assessed resistances against pathogens, such as *V. dahliae*, for which clear resistance/susceptibility is not available in model plant species.

The isolation of a *Ve* homologue from *M. longifolia* would open opportunities for introduction of verticillium wilt resistance into wilt-susceptible commercial mint varieties, although the marker-assisted introgression of verticillium wilt resistance from *M. longifolia* into peppermint may be precluded by the sterility of the latter, and by differences in ploidy. Transformation systems mediated by *Agrobacterium tumefaciens* Smith & Townsend, are available for peppermint and other commercial mints (Niu et al., 1998). The *Ve*-like mint sequences reported here provide a potential tool for isolation of full-length, functional *Ve* alleles as candidates for introduction into wilt-susceptible commercial mint varieties.

Table 2.1. PCR conditions used for amplification of resistance gene analog (RGA) and verticillium wilt resistance-like (*Ve*) sequences from *Mentha longifolia*. Degenerate, specific and inverse PCR profiles had an initial 2 min, 94 °C denaturation step and a final 7 min, 72 °C extension step. The reverse-transcription PCR (RT-PCR) profile had an initial, first-cDNA-stand synthesis step at 37 °C for 45 min and a final extension at 68 °C for 10 min.

| PCR profile | PCR primer set[z] | Denaturation step | Annealing step | Extension step | Cycles (no.) |
|---|---|---|---|---|---|
| Degenerate | 1+2, 1+3, 4+5 | 94 °C 50 s | 56 °C 1 min 20 s, 45 °C 30 s | 72 °C 1 min | 34 |
| Reverse transcription | 1+2 | 95 °C 1 min | 56°C 1 min | 68 °C 2 min | 39 |
| Specific | 6+7, 8+9, 10+11 | 94 °C 1 min | 55 °C 1 min | 72 °C 1 min | 30 |
| Inverse | 12+13 | 94 °C 1 min | 58 °C 1 min | 72 °C 2 min 30 s | 30 |

[z] The listed numbers correspond to PCR primer sequences in Table 2.2.

Table 2.2. Oligonucleotide primers used to amplify resistance gene analog (RGA) and verticillium wilt resistance-like (*Ve*-like) sequences from *Mentha longifolia*. Primers s2, as2 and as3 are from Leister et al. (1996). All other primers were designed during the present study for amplification of *M. longifolia Ve*-like sequences.

| Primer name | Target sequence | Primer sequence (5' to 3') |
|---|---|---|
| 1. s2 (forward) | P-loop (GGVGKTT) | GGNGGNGTNGGNAANACNAC |
| 2. as2 (reverse) | Hydrophobic domain (GLPLAL) | NAANGCNAGNGGNAANCC |
| 3. as3 (reverse) | Hydrophobic domain (GLPLAL) | NAGNGCNAGNGGNAGNCC |
| 4. Ve5F19 (forward) | Tomato *Ve1*-specific | GGAGTACGTGGAACAGCCT |
| 5. Ve5R23 (reverse) | Tomato *Ve1*-specific | AGTTATTCGCACTGAGGTCTAAT |
| 6. Ve2398-2416F Actual sequence at primer site in CMEN585 | *Ve* LRR | ATYNTNGATATAGCYTCCAA (ATCATCGATATAGCTTCCAA) |
| 7. Ve2791-2810R Actual sequence at primer site in CMEN585 | *Ve* LRR | CTGTGGRATYTCYCC (ATTCGGGATCTCTCC ) |
| 8. Ve4-24F Actual sequence at primer site in CMEN585 | Mint *Ve* | GTCGATATAGCTTCCAACAAT (ATCGATATATCTTCCAACAAT) |
| 9. Ve403-386R | Mint *Ve* | TCGGTACATGCCCTGACA |
| 10. Ve111-133F | Mint *Ve* | TACCGCCGAGCTTGTCCGTCCTT |
| 11. Ve974-951R | Mint *Ve* | TTCCGCGTGGTCTTCTTCTTCTCC |
| 12. IPCR138-153F | Mint *Ve* | GAAGCATCCCGAAGTC |
| 13. IPCR325-304R | Mint *Ve* | CAGCTAGAGAAGTTTATTGATT |

Table 2.3. *Mentha longifolia* resistance gene analogs (RGAs) obtained with degenerate primer PCR from various *M. longifolia* accessions. PCR templates were genomic DNA (gDNA) or cDNA. RGA clone designations in parentheses indicate redundant (identical) gDNA or cDNA clones.

| ***M. longifolia*** **USDA identification no.** | **RGA GenBank® accession no.** | **RGA clone** | **PCR template** | **Primer set (Table 2.2)** | **Length (bp)** |
|---|---|---|---|---|---|
| CMEN 17.001 | AF469684 | 17B9 | cDNA | 1+2 | 492 |
| CMEN 17.001 | AF481104 | 171 | gDNA | 1+2 | 493 |
| CMEN 17.001 | AY029196 | 178 | gDNA | 1+2 | 493 |
| CMEN 17.001 | AF474173 | 1736 | gDNA | 1+2 | 495 |
| CMEN 17.001 | AF469683 | 1711 | gDNA | 1+2 | 496 |
| CMEN 17.001 | DQ174111 | 173 | gDNA | 1+2 | 515 |
| CMEN 17.001<br>(CMEN 17.001) | AF469685 | 1712<br>(1725) | gDNA<br>gDNA | 1+2 | 526 |
| CMEN 81.001 | DQ174431 | 81CRCG2 | cDNA | 1+2 | 481 |
| CMEN 81.001 | AF481109 | 813 | gDNA | 1+2 | 501 |
| CMEN 81.001 | AF481108 | 8126 | gDNA | 1+2 | 506 |
| CMEN 81.001 | DQ174432 | 81CRCG4 | cDNA | 1+2 | 511 |
| CMEN 81.001 | AF474174 | 8116 | gDNA | 1+2 | 522 |
| CMEN 81.001 | DQ174109 | 8120 | cDNA | 1+2 | 680 |
| CMEN 501.001 | DQ174433 | 501as38a | gDNA | 1+3 | 511 |
| CMEN 585.001 | DQ174435 | 5857 | gDNA | 1+2 | 289 |
| CMEN 585.001 | DQ174434 | 5853 | gDNA | 1+2 | 385 |
| CMEN 585.001 | DQ174108 | 58517 | cDNA | 1+2 | 386 |
| CMEN 585.001 | DQ174110 | 58519 | cDNA | 1+2 | 436 |
| CMEN 585.001<br>(CMEN 585.001)<br>(CMEN 17.001) | AF469686 | 5856<br>(58531)<br>(1744) | gDNA<br>gDNA<br>gDNA | 1+2 | 495 |
| CMEN 585.001 | AF469687 | 5858 | gDNA | 1+2 | 511 |
| CMEN 585.001 | AF481107 | 5854 | gDNA | 1+2 | 525 |
| CMEN 585.001 | AF481106 | 58530 | gDNA | 1+2 | 525 |
| CMEN 585.001<br>(CMEN 17.001) | AF481105 | 58513<br>(1719) | gDNA<br>(cDNA) | 1+2 | 548 |

Table 2.4. Closest matches (blast hits) for mint resistance gene analog (RGA) sequences obtained by blastx searches of the GenBank® nr database. Each listed sequence is represented on the neighbor-joining tree (Fig. 2.2). RGA family designations refer to those in Fig. 2.2.

| Mint RGA family | Top blast hits (GenBank accession no.) | Plant species | Confers resistance to | Amino acid identity (%) | E-value |
|---|---|---|---|---|---|
| A | disease resistance-like protein (CAC82603) | *Coffea arabica* L. | | 48 | 2e-34 |
| B | NBS-LRR-like (AA089149) | *Gossypium barbadense* L. | | 41 | 1e-26 |
| | RGA Pt19 (AAN08179) | *Citrus grandis* Osbeck *xPoncirus trifoliata* (L.) Raf | | 43 | 1e-32 |
| | putative RGA (AAM77267) | *Malus prunifolia* | | 44 | 2e-33 |
| C,D | Bs2 (AAF09256) | *Capsicum chacoense* | *Xanthomonas campestris* Pammel | 44 | 2e-29 |
| | Rx (CAB50786) | *Solanum tuberosum* L. | Potato virus X | 44 | 6e-29 |
| | Gpa2 (AAF04603) | *S. tuberosum* | *Globodera padilla* (Stone) Behrens | 52 | 2e-45 |
| E | disease resistance-like protein (CAC82598) | *C. arabica* L. | | 40 | 1e-22 |
| | NBS-LRR protein (AAZ07904) | *Ipomoea batatas* (L.) Lam. | | 37 | 2e-22 |
| F | NBS-LRR protein (AAZ07913) | *I. batatas* | | 53 | 6e-46 |
| | RGA (AAL30111) | *Solanum phureja* Juz & Bukasov *x Solanum stenotomum* Juz & Bukasov | | 49 | 8e-43 |
| | kinase/encodes NBS (CAC79996) | *Solanum pinnatisectum* Dunal | | 51 | 6e-47 |
| G | RPI (AAR29069) | *Solanum bulbocastanum* Dunal | *Phytophthora infestans* (Mont.) de Bary | 52 | 2e-45 |
| | RGA1 (AAR29070) | *S. bulbocastanum* | *P. infestans* | 53 | 5e-45 |

Figure 2.1. Multiple amino acid alignment of mint resistance gene analogs (RGAs) and most closely related R genes and RGAs from other plant species. *M. longifolia* RGAs are in green type; most closely related sequences identified by blastx and tblastx are in black type. Conserved nucleotide binding site (NBS) motifs as defined by Meyers et al. (1999) are boxed. The P-loop and GLPL motifs are the sites of the s2 and as2 primers, respectively.

Figure 2.2. Neighbor-joining tree generated using Clustal X. Mint resistance gene analogs (RGAs) are in boldface type; most closely related sequences identified by blastx and tblastx are in lightface type. Numbers (percents) on branches represent bootstrap values for 1000 iterations. Roman numerals represent major groupings of RGAs. Branches defining families of closely related sequences are labeled A-G.

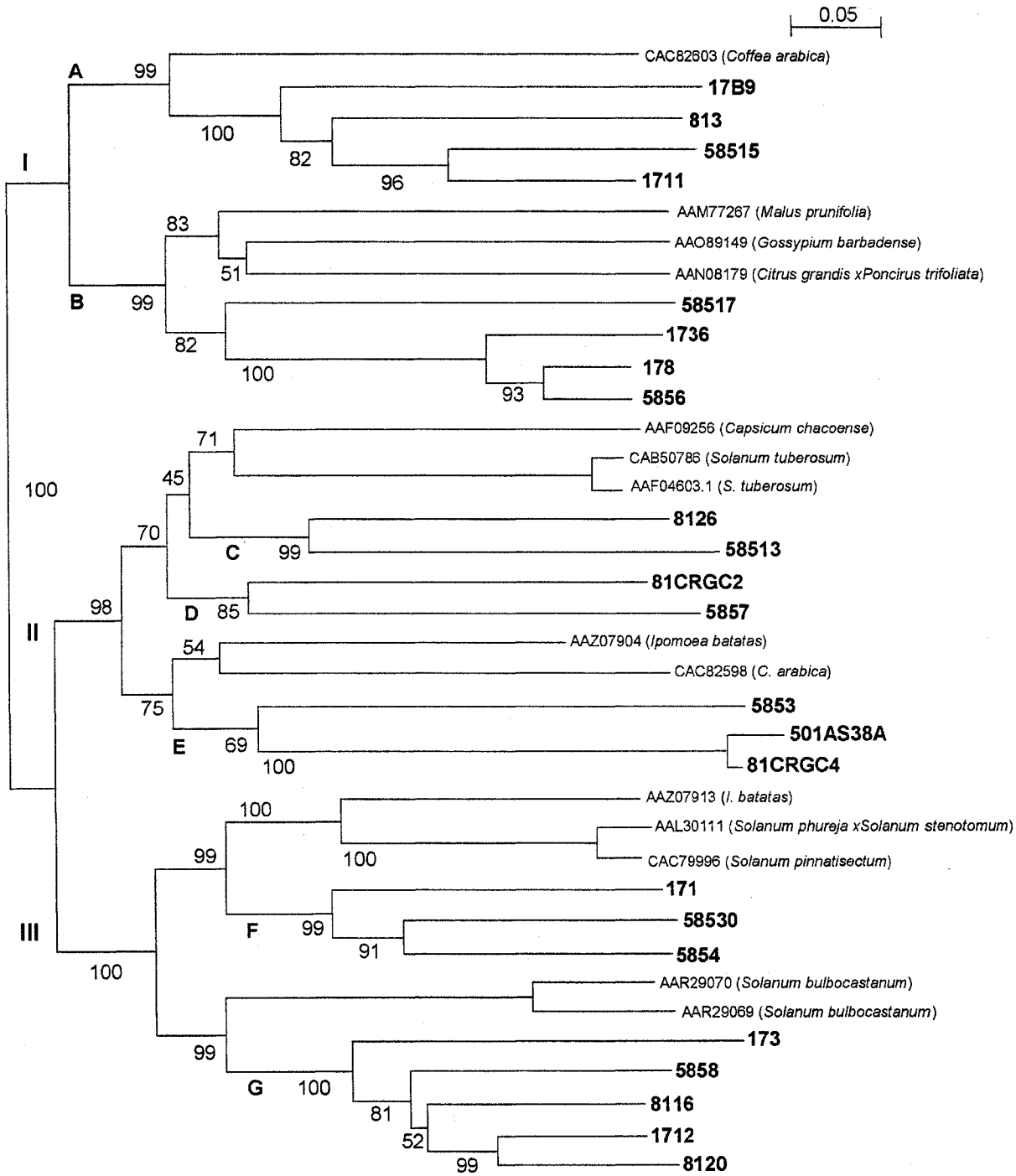
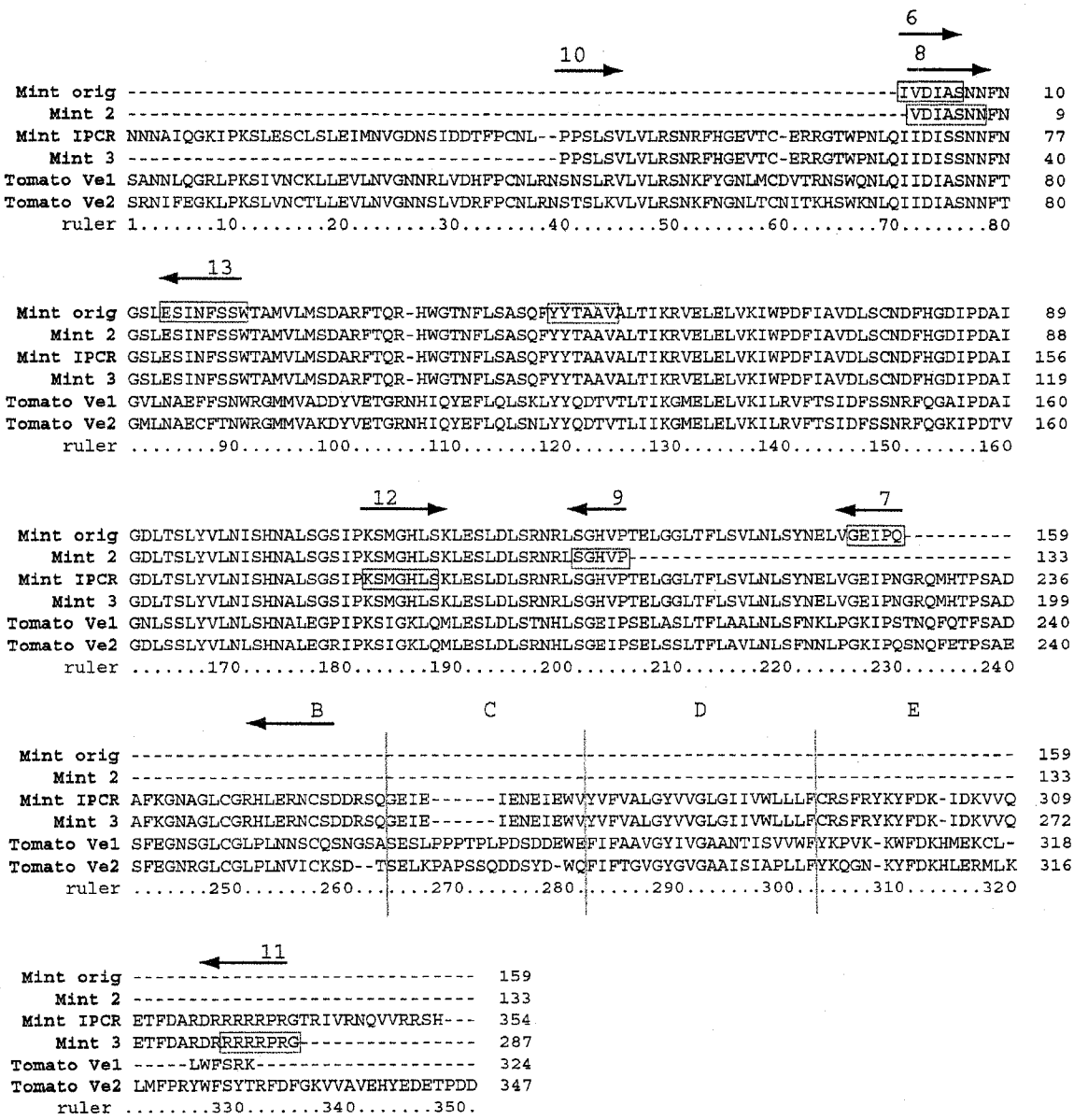


Figure 2.3. Multiple amino acid alignment of mint Ve sequences with partial tomato Ve sequences. The segments of tomato Ve1 and Ve2 shown here correspond to amino acids 730-1053 and 728-1074, respectively, of the entire predicted gene products. Numbered arrows and corresponding boxes indicate locations of primer sequences listed in Table 2.2. Letters indicate regions of tomato Ve defined by Kawchuk et al. (2001).



CHAPTER III.

# ISOLATION OF A POTENTIAL VERTICILLIUM WILT RESISTANCE GENE, *MVE1*, FROM *MENTHA LONGIFOLIA* (L.) HUDS, AND ANALYSIS OF *MVE1* ALLELES IN F1 AND F2 SEGREGATING POPULATIONS


## Abstract

We report the cloning and sequencing of the *mVe1* gene from the mint species *Mentha longifolia* (L.) Huds. *mVe1* is a potential homolog of the tomato (*Solanum lycopersicum* L.) verticillium wilt resistance genes *Ve1* and *Ve2*, and shares 51.4-51.7% and 50.8-51.0% predicted amino acid identity, respectively, with those genes. The *mVe1* gene has a coding region of 3051 bp (1017 amino acids), which ends with three consecutive stop codons. A major feature of the *mVe1* predicted protein is a leucine-rich repeat domain, which is a common feature of plant disease resistance proteins.

We isolated and compared seven *mVe1* alleles from two South African (SA) accessions (CMEN 584 and CMEN 585) and two European (E) accessions (CMEN 17 and CMEN 516) of *M. longifolia*. Two different alleles were isolated from each of three accessions, while only one allele was obtained from CMEN 17. These seven alleles shared 96.2-99.6% nucleotide identity. All polymorphisms were base substitutions.

Four progeny populations derived from resistant (R) x susceptible (S) crosses were genotyped with respect to *mVe1* alleles: SAF1 and SAF2 populations from the cross CMEN 585 (R) x CMEN 584 (S), and EF1 and EF2 populations from the cross CMEN 17 (R) x CMEN 516 (S). We screened subsets of each population for verticillium wilt resistance and looked for associations between *mVe1* genotype and wilt resistance or susceptibility. All four populations were found to be segregating with respect to identified *mVe1* alleles and wilt resistance vs. susceptibility. No association was found between *mVe1* genotype and wilt phenotype. Implications of these results are discussed.

## Introduction

Verticillium wilt is a fungal disease of plants that is caused by two deuteromycete species: *Verticillium albo-atrum* and *Verticillium dahliae*. These soil-borne plant pathogens are distributed throughout the world, and are of major economic importance because of the breadth of their host range and the amount of crop damage they cause. As many as 300 different plant hosts have been reported (University of Illinois Extension, 1997; Bhat and Sabbarao, 1999). Based upon shared patterns of host range specificity, both *Verticillium* species have been subdivided into races 1 and 2 (Bender and Shoemaker, 1984). Genotypes and varieties with varying levels of verticillium wilt resistance have been described in several plant species and a major locus (*Ve*) conferring qualitative resistance to this disease in tomato and potato has been cloned and

characterized (Kawchuk et al., 2001). However, sources of qualitative resistance have not been identified in most host species.

On the organismal level, early events in *Verticillium* spp. infection have been studied using cultured roots and seedlings from tomato and eggplant. Typically, infection occurs when the fungus penetrates root hairs and the root epidermis, both inter- and intracellularly (Al-Shukri, 1968; Tjamos and Smith, 1975). Hyphae then spread through the root cortex inter- and intracellularly, reach the stele, and invade the xylem (Tjamos and Smith, 1975). Once in the xylem, hyphae produce waves of conidia that spread upward in the transpiration stream and germinate in the vascular system (Brandt et al., 1984; Garas et al., 1986). In advanced infections hyphae clog the xylem, blocking water flow through the stem and ultimately causing plant death. In many affected crops, cultivars referred to as "resistant" or "tolerant" to verticillium wilt are still colonized by the fungus, but only to a moderate extent, and disease symptoms are mild or absent. Regardless of disease symptom expression, yields are reduced in *Verticillium* spp.-infected plants (McFadden et al., 2001). Relative resistance or susceptibility seems to hinge on the ability of the pathogen to proliferate in the vascular system of the plant host.

On the cellular/molecular level, most studies of plant-*Verticillium* spp. interactions have identified components of general, basal plant defense responses rather than pathogen-specific recognition mechanisms. Timing of these basal defense responses to *Verticillium* spp. infection is slower in susceptible plants than in resistant plants (Williams et al., 2002). Basal defense

responses have been demonstrated to include elemental sulfur accumulation in vascular tissue and associated parenchyma cells (Williams et al., 2002) and expression of pathogenesis-related (PR) protein genes (Hill et al., 1999; McFadden et al., 2001).

The only genes confirmed to play a role in *Verticillium* spp.-specific resistance have been identified in tomato. Resistance to *Verticillium* spp. race 1 in tomato was identified in a wild tomato accession and the single-gene nature of this resistance was subsequently shown in classical genetics experiments (Schaible et al. 1951). Tomato plants carrying the dominant *Ve* allele are resistant to race 1, but not to race 2, while homozygous recessive (*ve/ve*) plants are susceptible to both races.

When the *Ve* locus was mapped to tomato chromosome IX and subsequently positionally cloned and sequenced, it was found to consist of a pair of functional genes (*Ve1* and *Ve2*) (Kawchuk et al., 2001). These tomato *Ve* genes each contain a lengthy leucine-rich-repeat (LRR) domain, a common feature of plant disease resistance genes that is thought to play a role in recognition of pathogen-derived molecules, either extracellularly or intracellularly. The tomato *Ve* gene products are predicted to be cell surface receptor proteins. Both genes were used to transform susceptible tomato and potato varieties, and each gene conferred resistance (Kawchuk et al., 2001). Since the cloning and characterization of *Ve1* and *Ve2*, a few subsequent studies have focused on identification of *Ve* homologs in other Solanaceae species (Chai et al., 2003; Fei et al., 2004), on developing markers for marker-assisted selection in potato

(*Solanum tuberosum* L.), (Simko et al., 2004a, Simko et al., 2004b), and on identifying virulence-related genes in the pathogen (Rauyaree et al., 2005). In addition, the *EDS1* gene was subsequently shown to be required for *Ve*-mediated resistance in tomato (Hu et al., 2005). Much work remains to understand the molecular details of *Verticillium* spp. infection of plant hosts, including identification of other plant defense components contributing to wilt disease resistance, and elucidation of the timing of virulence expression in the fungus and induction/expression of resistance in host cells.

Among the many crops affected by verticillium wilt are the commercial mints. Two of the three commercially-grown mint species in the United States, peppermint (*Mentha ×piperita* L.) and Scotch spearmint (*M. ×gracilis* Sole), are susceptible to verticillium wilt, while the third, native spearmint (*M. spicata* L.) is relatively resistant (Berry and Thomas, 1961; Lacy and Horner, 1965; Sink and Grey, 1999). These species are perennial, clonally propagated plants that have been subjected to very little conventional breeding, and thus are, for all intents and purposes, wild plants. The mints are specialty crops in the United States. In 2005, 76,000 acres of peppermint were harvested, yielding 6,980,000 lbs. of oil worth $84 million, and 18,000 acres of spearmint produced 1,933,000 lbs. of oil worth $20 million (NASS, 2006).

*Mentha ×piperita* L.(peppermint) is a hybrid of *M. spicata* and *M. aquatica* (Tucker et al., 1980). Peppermint plants bearing the names Mitcham Peppermint or Black Peppermint, brought from England to New England by colonists, escaped from gardens and became naturalized around streams; these plants, re-

collected from the wild and put into cultivation starting in the 1790s, became the foundation of the U.S. mint industry (Landing, 1969). In 1883, stolons of peppermint cultivar Black Mitcham, presumed to be the same plant previously grown by colonists, were imported to Michigan from Mitcham, England by A.M. Todd for purposes of commercial production (Landing, 1969). Black Mitcham is still the primary peppermint cultivar in the U.S.

'Native Spearmint,' like peppermint, is believed to be an early colonial import from Europe that escaped from gardens and became naturalized (Landing, 1969). 'Scotch spearmint,' a hybrid of *M. arvensis* and *M. spicata* (Tucker et al., 1991), was discovered in a Wisconsin garden, purported to be an import from Scotland, and introduced into commercial production in 1910 (Landing, 1969).

The incidence of verticillium wilt has shaped the history of the United States mint industry, particularly because wilt-susceptible 'Black Mitcham' peppermint has always been the predominant cultivated crop. Until the 1940s, most mint was produced in Indiana and Michigan. Verticillium wilt was first observed in Michigan peppermint fields during the 1924 and 1925 growing seasons (Nelson, 1926). By the 1950s, the economic damage caused by verticillium wilt had triggered a serious decline in acreage devoted to peppermint production (Green, 1951). Peppermint cultivation shifted to the northwestern U.S., and today, 89% of U.S. mint oil is produced in Washington, Oregon, and Idaho (NASS, 2006). Verticillium wilt continues to be the most limiting disease of

commercial mints, particularly peppermint, and is suspected to have spread to northwestern mint fields via infected stolons (Douhan and Johnson, 2001).

'Black Mitcham' peppermint is a male-sterile hexaploid (2n=6x=96) (Chambers and Hummer, 1994). As such, it is not amenable to conventional plant breeding. In an attempt to produce wilt-resistant peppermint cultivars, workers from the A.M. Todd Company irradiated 'Black Mitcham' stolons in the late 1960s, and screened surviving (presumably mutant) stolons for wilt resistance, as well as for oil yield and quality (Murray and Todd, 1975). The result of this effort was the release of two peppermint cultivars, Todd's Mitcham (Murray and Todd, 1972) and Murray Mitcham (Todd et al., 1977), which, along with Black Mitcham, are still in commercial production (Lundy, pers. comm.[1]).

More recently, somatic hybrids of 'Black Mitcham' and 'Native Spearmint' were generated via protoplast fusion, with the intention of combining the wilt resistance of the latter with the desirable oil quality and yield of the former (Krasnyanski et al., 1998). Sink and Grey (1999) screened 743 of these clonal plants for verticillium wilt resistance and found nine relatively resistant clones. The Mint Industry Research Council, a collective of mint oil producers, end-users and researchers, sponsored a second induced mutation study, conducted by the Plant Technologies company in Albany, OR, the results of which were evaluated by Johnson and Cummings (2000), who found one wilt-resistant irradiated mutant.

[1] Rocky Lundy, Executive Director, Mint Industry Research Council, P.O. Box 971, Stevenson, WA 98648-0971. Email: mirc@gorge.net.

While considerable effort has been made to evaluate mints for field resistance to verticillium wilt, no previous work has focused on the genetics of wilt resistance in mint. We are developing the wild, diploid mint species *M. longifolia* as a model for the study of mint genetics and plant-microbe interactions. Toward this end, we have identified *M. longifolia* accessions from the United States Department of Agriculture (USDA) with varying resistance/susceptibility to verticillium wilt (Vining et al., 2005). As a first step toward identifying molecular determinants of verticillium wilt resistance in mint, we have also cloned resistance gene analog (RGA) sequences from some of the USDA *M. longifolia* accessions (Vining et al., 2007). Furthermore, we used degenerate primers based on the tomato *Ve1* gene to isolate an initial 445 bp segment of a *Ve*-like sequence from mint, then used inverse PCR to walk outward, thereby obtaining a 1413-bp *M. longifolia* sequence that had 56-57% predicted amino acid identity to the equivalent regions of tomato *Ve1* and *Ve2* (Vining et al., 2007). Here, we report the putatively complete coding sequence of an *M. longifolia Ve* homolog, *mVe1*, evaluate verticillium wilt susceptibility/resistance in *M. longifolia* F1 and F2 populations derived from crosses of wilt-resistant and wilt-susceptible *M. longifolia* accessions, and look for associations between *mVe1* genotype and wilt disease resistance or susceptibility phenotype.

## Materials and methods

### Plant materials

*Mentha longifolia* accessions were obtained as stolons or rooted cuttings from the National Clonal Germplasm Repository (NCGR) in Corvallis, Oregon. Based on the results of prior germplasm screenings (Vining et al., 2005), verticillium wilt-resistant, diploid accessions CMEN 585 (South Africa) and CMEN 17 (unknown European country) and wilt-susceptible, diploid accessions CMEN 584 (South Africa) and CMEN 516 (Italy), were chosen as parents for use in resistant x susceptible crosses.

The cross CMEN 585 × CMEN 584 was performed by rubbing fingers on CMEN 584 inflorescences to gather pollen, and then applying the pollen directly to CMEN 585 stigmas, or by rubbing together inflorescences of the two parent plants. Self-pollination of CMEN 585 was not an issue, because CMEN 585 was never demonstrated to have functional pollen: numerous attempts to self-pollinate CMEN 585 and to use CMEN 585 as a pollen parent were unsuccessful. In addition, during the summer months, CMEN 585 anthers typically failed to develop and pollen was rarely observed. An F1 population from the cross CMEN 17 × CMEN 516 was available at the onset of this investigation.

The F1 population produced from the South African (SA) cross CMEN 585 × CMEN 584 will hereafter be referred to as SAF1, and the F1 population from the European (E) CMEN 17 × CMEN 516 cross will be called EF1. Respective F2 generation populations (SAF2 and EF2) were then obtained by allowing one

selected plant from each F1 population to self-pollinate in isolation. In addition, an S1 population ('17 self' or 17⊗) was developed from accession CMEN 17 via self-pollination in isolation.

***mVe1* sequencing**

Three mint-specific primers based on the previously isolated m*Ve*1 sequence from CMEN 585 (Vining et al., 2007) and four generic, degenerate primers (Table 3.1) were used in Thermal Asymmetric InterLaced (TAIL)-PCR (Liu et al., 1995; Liu and Whittier, 1995) with template DNA from *M. longifolia* accession CMEN 585. Thermal cycling parameters are listed in Table 3.2. All reactions were performed using the Accuprime™ Taq DNA Polymerase System (Invitrogen, Carlsbad, CA) in 25 µl total reaction volumes. Primary TAIL-PCRs contained 3 µM degenerate primer and 0.4 µM specific primer. Products from primary TAIL-PCR were diluted 1:50 with nuclease-free water (Ambion, Inc., Foster City, CA) and used as templates in secondary PCR. Secondary and tertiary TAIL-PCRs contained 1.8 µM degenerate primer and 0.4 µM specific primer. Products from secondary PCR were used directly in tertiary PCR, without dilution. Secondary and tertiary TAIL-PCR products obtained from the combination of degenerate primers AD1 and AD6 with the *mVe*-specific primers were cloned directly, using the TOPO TA Cloning® Kit for Sequencing (Invitrogen). Resulting colonies were PCR-screened to determine insert size, and recombinant plasmids were isolated using the Wizard® SV Plasmid Purification kit (Promega Corporation, Madison, WI). Sequencing reactions were performed

using the UNH Hubbard Center for Genome Studies protocol (http://hcgs.unh.edu/protocol/sequence/), with DYEnamic™ Terminator ET cycle sequencing premix (Amersham Biosciences, Piscataway, NJ). Reaction products were analyzed using an ABI PRISM® 377 DNA sequencer (Applied Biosystems, Foster City, CA).

Once the preliminary draft of a putatively complete *mVe1* coding region was determined, PCR primers were designed in order to amplify PCR products containing this entire coding region from *M. longifolia* accessions CMEN 585, CMEN 584, CMEN 17, and CMEN 516. These products were amplified and cloned directly, as described for TAIL-PCR products above. To confirm accuracy of the *mVe* sequences, overlapping sequence reads were obtained from the clones by using a series of *mVe*–specific sequencing primers (Table 3.3), plus standard vector (M13F, M13R) primers.

Sequence alignments and ORF translations were performed using Lasergene Megalign software (DNASTAR, Madison, WI) and Clustal X (Thompson et al., 1997). GenBank® BLAST searches were done on the National Center for Biotechnology web site (http://www.ncbi.nlm.nih.gov). Alignments of *mVe1* predicted amino acid sequence with the tomato *Ve* predicted amino acid sequences were used to delineate protein domains. Further structural analyses of the putative mVe1 protein were executed on the Swiss EMBnet node server (http://www.ch.embnet.org) and Expasy Molecular Biology Server (http://www.expasy.org).

## Genotyping of F1 and F2 populations

Restriction digests of *mVe* PCR products were employed to determine allele compositions of parent plants, and of plants from F1 and F2 populations. Primers MA148 and MA151, which amplified the entire *mVe1* coding region as described above, were used to generate PCR products from all plants that had been screened for verticillium wilt resistance. To confirm PCR template quality, a pair of primers targeting a mint chalcone synthase gene (Lange et al., 2000) was employed. The chalcone synthase primers (ML414_79F, sequence 5'GCATCACCAACAGCGAACA3', and ML414_528R, sequence 5'CGGAGCAGACGATGAGGAC3') were based on peppermint cDNA clone ML414 (Genbank® accession number AW255394).

For restriction digests, 5 µl of PCR products were digested in 20 µl reactions with 10 units of either *Nco*I and *Xba*I (SAF1, SAF2), or 10 units of *Age*I (EF1, EF2) (New England Biolabs, Ipswich, MA). All digests were incubated overnight at 37 °C and digest products were electrophoretically separated on 2% agarose gels.

## Verticillium wilt resistance screening

Verticillium wilt resistance screenings of F1 and F2 generation and parental plants were performed as described in Vining et al. (2005). Briefly, rooted cuttings were inoculated by root-dipping in 20 ml of a fungal spore suspension (concentration ~$10^7$ spores/ml) for 5 min, then replanted in sterilized soil. Following inoculation, plants were maintained in a growth chamber under a

cool-temperature, short-day light regimen for two weeks before being moved to a greenhouse bench under ambient light and temperature conditions. Disease symptoms were evaluated at eight weeks post-inoculation using a rating scale on which zero designated a healthy plant, increasing scores indicated increasing symptom severity, and four described a dead plant. At least six rooted cuttings (clones) were evaluated per plant, and mean symptom ratings and standard deviations were calculated for each set of clones.

### Genotype-phenotype association testing

The hypothesis of association between verticillium wilt resistance score and allele composition was tested using two statistical methods: one-way, parametric analysis of variance (ANOVA), and nonparametric, Kruskal-Wallis single-factor analysis of variance by ranks, both performed with Systat© version 10 (SPSS Inc., Chicago, IL).

## Results

### Population development

The South African (SA) cross CMEN 585 x CMEN 584 produced an F1 population (SAF1) of 55 plants, and self-pollination of one wilt-resistant F1 plant (SAF1-1) produced an F2 generation population (SAF2) of 120 plants. The European (E) cross CMEN 17 x CMEN 516 produced an $F_1$ population (EF1) of 16 plants, and self-pollination of one wilt-resistant $F_1$ plant (EF1-8) produced an

$F_2$ generation population (EF2) of 18 plants. Self-pollination of CMEN 17 produced an $S_1$ population (17⊗) of 31 plants.

Due to logistical considerations, including resource-based limitations on the number of plants that could be screened, and death of some plants during the term of this study, various population subsets were used in most of the population genotyping and wilt resistance screening assays. An obvious trait segregating in the SAF1 population was foliar scent, in that approximately half the individuals had a "spearmint" scent, while the other half had a "non-spearmint" scent. Therefore, plants representing each scent category were included in the SA population subsets chosen for wilt resistance screening.

***mVe1***

Using TAIL-PCR to extend a previously-obtained *Ve*-like fragment, we completed a draft sequence of a putative homolog, *mVe1*, of the tomato *Ve* genes from verticillium wilt-resistant *M. longifolia* accession CMEN 585. Initial sequence assembly of the degenerate, inverse and TAIL-PCR fragments produced a sequence of 3328 nucleotides (nt), including 210 nt upstream of the putative start codon and 58 nt downstream of the putative stop codon (Appendix C). The *mVe1* sequence has a predicted coding region of 3051 nt, with no introns. The predicted stop codon (TAA) is followed immediately (in frame) and consecutively by two more stop codons (TGATGA) (Fig. 3.1).

The predicted mVe1 protein, 1017 amino acids in length, is rich in leucine residues (15.2%). Five distinct sequence domains were identified in this

predicted protein corresponding to the domains A through E of the tomato Ve proteins as defined by Kawchuk et al. (2001) (Fig. 3.2). The amino terminus of 24 amino acids (domain A) precedes an LRR region (amino acids 25-934, domain B) with 36 imperfect copies of the consensus sequence [XXIXNLXXLXXLXLSXNXLSGXIP]. Immediately adjacent to the LRR domain, a negatively-charged stretch (amino acids 935-951, domain C) precedes a putative transmembrane domain (amino acids 952-973, domain D). The carboxy terminus (amino acids 974-1017, domain E) is positively charged, and contains a stretch of seven consecutive arginine residues.

The predicted *mVe1* gene product has 51.4% and 50.8% amino acid identity with tomato *Ve1* and *Ve2* gene products, respectively. Blastx searches of the nr database using translated *mVe1* sequences as search queries have tomato Ve2 as the top Blast hit (score=854, e-value=0.0). Tomato Ve1 is within the top 10 Blast hits, as are the SlVe1 and SlVe2 sequences from *Solanum lycopersicoides* (Chai et al., 2003). Sequence identities among the translated *mVe1* sequences and the *Solanum* sequences are 48-50%. Other *mVe1* tblastx hits are leucine-rich-repeat proteins and predicted proteins from a variety of plant taxa (≤41% identities).

When each of the tomato *Ve*s was used as a search query in a blastx search of the nr database the top hits were sequences from other Solanaceae species (≥80% identities). Outside of the Solanaceae, the closest blastx hits for the tomato *Ve* translations were from *Medicago truncatula* Gaertner (40-45% identities) and *Oryza sativa* (≤40% identities).

With the objective of isolating and assessing sequence variation among *mVe1* alleles, PCR primers (MA148 and MA151, Table 3.1) based on the 3328 nt sequence were used to amplify products containing putatively complete coding regions from the four the *M. longifolia* accessions used as crossing parents (CMEN 585, CMEN 584, CMEN 17, CMEN 516). These *mVe1* PCR products were cloned and sequenced, identifying two alleles each from CMEN 585 (alleles I and II) and CMEN 584 (alleles III and IV), one allele from CMEN 17 (allele V), and two from CMEN 516 (VII and VIII) (Fig. 3.3). The seven distinct *mVe1* alleles have 96.2-99.6% predicted amino acid identity with each other. All encode predicted proteins of the same length (1017 amino acids), and all coding regions end in three consecutive stop codons. All differences among the *mVe1* alleles are base substitutions; there are no indel polymorphisms.

Comparisons of *mVe1* alleles within and among plant accessions revealed strong similarities among accessions from common geographic origins, as well as striking differences between alleles of South African plants and those from European plants (Fig. 3.3, Table 3.4). The two most similar alleles were from South African accessions CMEN 585 and CMEN 584 (alleles I and III, respectively), which differed at only two nucleotides, and were identical in amino acid sequence. Pairs of alleles from South African plants averaged 9 nucleotide differences, and alleles from European accessions differed by an average of 4 nucleotides. In contrast, there was an average difference of 73 nucleotides in pairwise comparisons of South African with European alleles.

***mVe1* genotyping**

The same PCR primers used to amplify and clone *mVe1* alleles from parent accessions were used to amplify alleles of plants in the SAF1, SAF2, EF1 and EF2 populations derived from these accessions. Based upon the sequence polymorphisms described above, restriction enzymes that differentially cut particular alleles were chosen and employed for allele genotyping in the four populations.

Restriction digests of *mVe1* PCR products from SA parent accessions produced fragment sizes expected on the basis of the known restriction site positions (Fig. 3.4). In the SAF1 population, sixteen plants chosen for verticillium witl resistance screening (see below) were genotyped with respect to *mVe1* alleles. Given the inferred genotypes of the parent plants, there were four possible F1 allele compositions (each of which could be inferred from its respective banding pattern). All four of these expected allele compositions were represented in the SAF1 population sample. Eight SAF1 plants had allele composition I / IV; four plants had allele composition II / IV; three plants had allele composition II / III; a single plant (SAF1-39) had allele composition I / III. The SAF1 plant (SAF1-1) that was self-pollinated to produce the SAF2 generation had allele composition I / IV. In the SAF2 population, 74 plants were genotyped. As inferred from gel banding patterns, twenty-seven were homozygous for the CMEN 585-derived allele (I / I), thirty-seven were heterozygous (I / IV), and ten were homozygous for the CMEN 584-derived allele (IV / IV).

Death of several plants in the EF1 and EF2 populations prohibited genotyping of every individual from the initially produced populations. A further complication in genotyping of surviving plants was the high sequence identity/low number of polymorphisms among the European alleles. CMEN 17 allele V and CMEN 516 allele VIII were distinguishable at the DNA sequence level, but could not be differentiated on the basis of restriction digests. Initially, evidence of only one kind of *mVe1* allele (allele V) in CMEN 17 suggested that this accession might be homozygous, but the EF1 and EF2 population genotyping results refuted this hypothesis, as explained below.

Restriction digest banding patterns obtained or expected in the European parents, EF1 and EF2 populations are depicted in Figure 3.5, which provides a necessary basis for the following description and explanation of results. The banding pattern obtained for CMEN17 conformed to the known presence of a single restriction site in allele V, the only allele identified by cloning in CMEN 17. The pattern for CMEN 516 conformed to the known presence of one restriction site in allele VII (resulting in two bands of 1362 bp and 1769 bp, equal in size to the two bands characteristic of allele V) and two restriction sites in allele VII (resulting in bands of 1362, 1198, and 571 bp). Because alleles VII and VIII each produce a band of 1362 bp, CMEN 516 has only four gels bands, as shown.

A pivotal outcome was the absence of any *mVe1* PCR product band in 8 plants of the SAF2 population (even though the same templates produced a band with the control chalcone synthase primer pair). This outcome prompted the inference that a null (non-amplifiable) *mVe1* allele was segregating in this

population. Given that only one *mVe1* allele had been identifiable via PCR amplification and cloning in CMEN17, it was hypothesized that CMEN 17 was the source of the null allele, which was designated allele VI(N). This hypothesis was tested by performing PCR amplifications of *mVe1* on the 17S1 population. If CMEN 17 carried the null allele, and had genotpye V / VI(N), then 25% of the 17S1 plants would be expected to have genotype VI(N) / VI(N), and produce no mVe1 PCR product. In support of the null allele hypothesis, six out of nineteen 17S1 plants produced no *mVe1* PCR product (although the templates were good as confirmed by chalcone synthase amplifications).

Proceeding on the assumption that CMEN17 had the genotype V / VI(N), four possible genotypes, and three possible banding patterns, would be expected in the EF1 (Fig. 3.5). Of the nine EF1 plants, 6 had banding pattern A, and 3 had banding pattern B. Expected pattern C did not occur, but its absence could be due to sampling "error" in this small population sample.

**Verticillium wilt resistance screening**

All 16 plants from the EF1 population were screened for wilt resistance, while a subset of 16 SAF1 plants was screened. Thirty-three SAF2 plants and 18 EF2 plants were screened for wilt resistance. The original crossing parents were re-evaluated alongside progeny plants in the SAF1, SAF2 and EF2 trials.

Data from sets of clones with symptom ratings having standard deviations ≥1.2 were considered uninformative and were discarded; therefore, the final data

set consisted of 16 SAF1 plants, 24 SAF2 plants, 16 EF1 plants and 16 EF2 plants.

In the final SAF1 data set, six plants were resistant (mean ratings ≤0.5), six were moderately susceptible (mean ratings from 1.0 to 2.1), and four were highly susceptible (mean ratings >3.0) (Fig. 3.6A). In the SAF2 population of 24 plants, 13 plants were resistant (mean ratings ≤0.6), eight plants were moderately susceptible (mean ratings from 1.4 to 2.0), and six were susceptible (mean ratings from 2.3 to 3.4) (Fig. 3.6B).

In the EF1 population, plants were either highly resistant (six plants with mean ratings < 0.5), or highly susceptible (ten plants with mean ratings ≥3.5) (Fig. 3.7A). In the EF2 population of fifteen plants, ten plants were categorized as resistant (mean ratings ≤0.5), while the remaining five plants had various levels of disease symptoms (ratings ranging from 1.5 to 3.8) (Fig. 3.7B),.

Of the SAF2 plants subjected to verticillium wilt resistance screening, seven were homozygous for the CMEN 585-derived allele (I / I); three were homozygous for the CMEN 584-derived allele (IV / IV); and 14 were heterozygous (I / IV). Analyses of variance did not show a significant effect of plant genotype on disease symptom rating in SAF2 population (p=0.137 for parametric ANOVA, p=0.100 for Kruskal-Wallis nonparametric ANOVA) or the EF2 population (p=0.085 for parametric ANOVA, p=0.099 for Kruskal-Wallis non-parametric ANOVA).

## Discussion

Verticillium wilt has been a particularly challenging problem in plant pathology, because of the broad host range of the pathogen, and because of the dearth of resistant genotypes in many plant species. Verticillium wilt has plagued the U.S. mint industry for more than 80 years, in large part due to the genetic homogeneity of the susceptible cultivated peppermint. We have developed *Mentha longifolia* as a model species for the study of mint genetics, with particular attention to the problem of verticillium wilt resistance. We found in a previous study that some *M. longifolia* USDA accessions are highly resistant to verticillium wilt, while others are highly susceptible, under greenhouse conditions (Vining et al., 2005).

In the present study, we performed crosses between pairs of resistant and susceptible *M. longifolia* accessions from South Africa and from Europe, and generated segregating F1 and F2 generation populations as resources for genetic studies of verticillium wilt resistance. We have isolated putative homologs of the tomato verticillium resistance (*Ve*) genes from the four parental accessions used in these crosses, screened F1 and F2 populations derived from wilt-resistant x wilt-susceptible crosses, and looked for genotype-phenotype associations in F1 and F2 populations.

Starting with a gene segment obtained via degenerate PCR and inverse PCR (Vining et al., 2007), we used TAIL-PCR to acquire the putatively complete coding sequence of a potential *Ve* homolog from *M. longifolia* CMEN 585, and

then cloned *mVe1* sequences from this and three other *M. longifolia* accessions. All of the *mVe1* sequences obtained from these accessions are complete ORFs of equal length and could encode functional proteins. Like the tomato *Ve* genes, there are no predicted introns in *mVe1*. The *mVe1* gene is predicted to encode 36 leucine-rich repeats (LRRs), in contrast to the 38 LRR units of tomato *Ve1* and *Ve2* genes. The amino acid alignment of the predicted mVe1 and tomato Ve1 and Ve2 proteins showed two gaps in the mVe1 sequences relative to the tomato sequences, which corresponded to LRR units 8 and 17 in both Ve1 and Ve2. These gaps could represent a loss of individual repeats in the mint genes, or a gain of repeats in the tomato genes. Interestingly, when Ve homolog sequences from other species of Solanaceae were included in alignments, the same gaps existed in *M. longifolia* sequences relative to all Solanaceae sequences. Therefore, whatever gains or losses of LRR units occurred, they likely preceded the divergence of Solanaceae species.

Leucine-rich-repeat domains are common to the majority of known plant disease resistance genes. Most variation among homologs of disease resistance genes occurs in LRRs; there is evidence of diversifying selection at loci containing multiple orthologous LRR disease resistance genes (McDowell et al., 1998; Meyers et al., 1998; Dodds et al., 2006; Seah et al., 2007). Homologs may have many amino acid differences, yet confer equivalent resistance to a pathogen, as is the case at the tomato *Cf-9* locus, where two homologs with 61 amino acid differences each confer resistance to the same isolate of *Cladosporium fulvum* (van der Hoorn et al., 2001). In other cases, closely related

homologs that differ by only a few LRR amino acids may confer resistance to different isolates of a pathogen, as is the case for the *L* locus in flax, in which homologs encoding proteins with six LRR amino acid differences confer resistance to different isolates of *Melampsora lini* (Dodds et al. 2001).

The tomato *Ve* locus contains two presumably paralogous genes, both of which confer resistance to *Verticillium* spp. race 1. Homologs of the tomato *Ve* genes, *StVe1* and *SlVe1* from potato species, have not been tested for functionality. It is not known which, if any, amino acids in the LRR domain of *Ve* are key determinants of wilt resistance. The current picture of the mint-*Verticillium dahliae* interaction is even more murky. At the present time, there is no genetic map for any mint species; so there is no basis as yet to map either the *mVe1* locus or qualitative/quantitative trait loci involved in resistance itself. We have identified a second *mVe* homolog, *mVe2*, the preliminary analysis of which does not show any linkage to *mVe1* (data not shown). Therefore, the *mVe1* locus thus far is defined by a single gene.

The cloning of the *mVe1* sequence from *M. longifolia* accession CMEN 585 facilitated the cloning of *mVe1* sequences from other *M. longifolia* accessions. Similarly, the identification of the tomato *Ve* genes provided a foundation for the homology-based cloning of potential *Ve* homologs from other members of the Solanaceae: *SlVe1* from diploid potato *Solanum lycopersicoides* (Chai et al., 2003), *StVe* from diploid potato *Solanum torvum* Swartz (Fei et al., 2004), and 11 *StVe1* alleles from the cultivated, tetraploid potato *Solanum tuberosum* L. (Simko et al. 2004a, Simko et al. 2004b). The 11 *StVe1* alleles of

*S. tuberosum* share 76.2-99.6% predicted amino acid identity with each other, and 82.9-90.8% and 74.2-90.8% to tomato *Ve1* and *Ve2*, respectively. In mint, among the *mVe1* alleles, overall amino acid similarity ranges from 96.2-99.6%. The *mVe1* alleles from European accessions are more similar to each other than are the alleles from South African accessions. This is not surprising, since the two South African accessions are considered different subspecies: CMEN 585 is *M. longifolia* subsp. *capensis*, while CMEN 584 is *M. longifolia* subsp. *polyadenia* (Tucker and Naczi, 2005).

The detection of a 'null allele' in CMEN 17 inhibited genotyping in the EF1 and EF2 populations. This null allele could be a deletion of part or all of an allele. Alternatively, the *mVe1* primer sites could be missing or altered. Further investigations into the nature of this allele may include tests of alternative *mVe1* primers and restriction digests of any partial *mVe1* products to further distinguish allele VI from CMEN 17; these investigations are necessarily open-ended, and extend beyond the scope of this dissertation.

In addition to cloning *mVe1* alleles from *M. longifolia*, we have begun cloning *mVe1* alleles from other species in the *Mentha* genus, including the cultivated spearmint, *M. spicata*, and the *Mentha* ×*piperita* peppermint cultivar Black Mitcham (data not shown). It is probable that *mVe1* alleles are ubiquitous in *Mentha*. It remains to be seen whether any of these alleles encode proteins that confer verticillium wilt resistance in mint.

Plant defense responses to pathogen infections are complex and include pathogen-specific, early-recognition responses as well as all-purpose defenses

that may be invoked by general environmental stresses in addition to biotrophic stresses. The tomato *Ve* protein, as a predicted integral membrane receptor molecule, is implicated as one of the early actors in the plant defense response to *Verticillium* spp. invasion, and, as such, may be essential to wilt disease resistance. However, Ve protein localization in cells has not been confirmed. In tomato, *Ve*-mediated resistance to verticillium wilt is dependent on EDS1, a gene product that is also required for plant disease resistance conferred by certain proteins of the Toll-Interleukin-Receptor/Nucleotide-Binding-Site/Leucine-Rich-Repeat (TIR-NBS-LRR) class (Hu et al., 2005). It is not known what pathogen elicitor is recognized by Ve, or by what direct or indirect mechanism. Interactions of Ve with EDS and with other components of the plant defense system have not been determined. Other components of the Ve-associated defense mechanism remain to be identified.

The *Verticillium dahliae* screens of F1 and F2 *M. longifolia* populations showed that all were segregating for resistance/susceptibility to verticillium wilt. The EF1 progeny of a cross of European *M. longifolia* accessions were segregating approximately 1:1 for wilt resistance/susceptibility, while the EF2 population presented more continuous data. The SAF1 and SAF2 populations derived from a cross of South African *M. longifolia* accessions both exhibited a continuous distribution of disease symptom ratings.

In each of the *M. longifolia* F1 and F2 populations tested in the present study, there was no association between *mVe1* allele composition and resistance or susceptibility to verticillium wilt. At the same time, the present analysis of

*mVe1* sequences in resistant and susceptible *M. longifolia* accessions has uncovered only putatively complete ORFs and no apparent pseudogenes, suggesting that the identified *mVe1* sequences belong to functional genes. Therefore, despite the observed lack of association between alleles of the *mVe1* gene and a verticillium wilt phenotype in the studied population, it remains an open question as to whether a functional mVe1 protein is required for verticillium wilt resistance in mint. Plant resistance to any pathogen is dependent upon the expression of multiple genes, a defect in any one of which could confer disease susceptibility.

The *M. longifolia* accessions used as crossing parents in this study, particularly the South African accessions, were highly phenotypically diverse (Vining et al., 2005). It must be presumed that this diversity extends to alleles of genes involved with all aspects of plant defense, and that alleles at many resistance-related loci may be segregating in the studied populations. The continuous distribution of disease symptom ratings in three of the four *M. longifolia* populations points to a polygenic model, although the studied populations are too small to provide a test of this hypothesis. Further population development, including the derivation of inbred lines for future crosses and genetic studies, will likely aid in the identification of genes of qualitative effect.

In sum, it is possible that the *mVe1* gene does play a role in the resistance of *M. longifolia* to verticillium wilt, but that the verticillium susceptibility observed in parents CMEN 584 and CMEN 516 and in segregating progenies derived from these parents is not due to mutation at the *mVe1* locus. Instead, it is possible that

polymorphism in another gene or genes makes the difference between wilt resistance and susceptibility in these plants. Details of the molecular interactions between mint and *Verticillium* fungi remain to be elucidated. Future work with the *M. longifolia-Verticillium dahliae* interaction will include analysis of *mVe1* expression and a search for a mint *EDS1* homolog and for other candidate genes of possible relevance to verticillium wilt resistance in mint.

Table 3.1. Primers used in TAIL-PCR and *mVe1* coding region cloning. Primers AD1 and AD6 are from Liu et al. (1995); primers AD2 and AD3 are from Liu and Whittier (1995). All other primers were designed as part of the present study.

| **Primer name** | **Purpose** | **Sequence 5' ->3'** |
|---|---|---|
| AD1 | TAIL-PCR degenerate | NGTCGASWGANAWGAA |
| AD2 | TAIL-PCR degenerate | TGWGNAGSANCASAGA |
| AD3 | TAIL-PCR degenerate | AGWGNAGWANCAWAGG |
| AD6 | TAIL-PCR degenerate | WGTGNAGWANCANAGA |
| 396R (MA129) | TAIL-1 PCR *mVe1*-specific | TCTCCAACGACATGCAACTCTCA |
| 362R (MA131) | TAIL-2 PCR *mVe1*-specific | GATTTTCCCTTGTATAGCATTGTTGTTA |
| 290R (MA128) | TAIL-3 PCR *mVe1*-specific | GATGTGTCCGCTGATGTTGTTTC |
| 188F (MA151) | *mVe1* coding region cloning | CATCATCACTCATCTCCTTCACAA |
| 2789R (MA148) | *mVe1* coding region cloning | TTGCGCAGAAACCTACA |

Table 3.2. Thermal cycling parameters used in TAIL-PCR, based on those of Liu et al. (1995). An Eppendorf Mastercycler ep was used for all reactions. Thermocycler lid temperature was 105°C. Temperature ramp was set at the default of 3°C/sec unless otherwise indicated.

| Reaction | No. of cycles | Thermal Settings |
|---|---|---|
| Primary | 1 | 93°C, 1 min; 95°C, 1 min |
| | 5 | 94°C, 30 sec; 62°C, 1 min; 72°C, 2.5 min |
| | 1 | 94°C, 30 sec; 25°C, 3 min; ramp 0.3°C/sec to 72°C; 72°C, 2.5 min |
| | 15 | 94°C, 30 sec; 68°C, 1 min; 72°C, 2.5 min;<br>94°C, 30 sec; 68°C, 1 min; 72°C, 2.5 min;<br>94°C, 30 sec; 44°C, 1 min; 72°C, 2.5 min; |
| | 1 | 72°C, 5 min, hold 4°C |
| Secondary | 12 | 94°C, 30 sec; 64°C, 1 min; 72°C, 2.5 min<br>94°C, 30 sec; 64°C, 1 min; 72°C, 2.5 min<br>94°C, 30 sec; 44°C, 1 min; 72°C, 2.5 min |
| | 1 | 72°C, 5 min, hold 4°C |
| Tertiary | 40 | 94°C, 30 sec; 44°C, 1 min; 72°C, 2.5 min |
| | 1 | 72°C, 5 min, hold 4°C |

Table 3.3. Primers used in confirmatory sequencing of *mVe1*. Primer sites refer to locations within the *mVe1* sequence shown in Figure 1B.

| **Primer** | **Site in *mVe1* (bp)** | **Sequence 5' ->3'** |
|---|---|---|
| MA028 | 2455-2475 | ATCGATATATCTTCCAACAAT |
| MA030 | 2837-2854 | TCGGTACATGCCCTGACA |
| MA033 | 2971-2987 | AACGCGGGATTATGTGG |
| MA123 | 3302-3318 | TTGCGCAGAAACCTACA |
| MA133 | 1965-1991 | GGGCGACCTGCACCTCTTTATTTCTCC |
| MA138 | 3087-3110 | ACAATGATTCCTGAGCCCACAACA |
| MA142 | 1048-1071 | CAAAAACCCACTCTACAAAATCTC |
| MA144 | 2127-2140 | GACAAGGGGCTATA |
| MA150 | 775-794 | GAAACATCGACACCATCAAG |
| MA152 | 424-437 | TGGCCTGCACCGTC |
| MA157 | 253-276 | ACAGCAACAACTTTTACTACTCTT |
| MA227 | 350-374 | CTACAACGTCTTCAACCGCACTCA |
| MA228 | 679-702 | GCAAAGGCCCAGAAACACACAAT |
| MA229 | 661-685 | GGAGTTTGAGCTTGCGTTATTGTA |
| MA230 | 1562-1585 | TCAGGGAAGGCATGCAAGTCACA |
| MA231 | 1136-1160 | TTCCACGTTGTTTCGAGGTCTATC |
| MA232 | 1110-1133 | AATGAGCCGGTGAAGAAGTTAGC |

Table 3.4. Pairwise comparisons of *mVe1* nucleotide sequences. Numbers of nucleotide differences are given, followed by number of amino acid differences in predicted protein sequences. Alleles I and II are from *M. longifolia* accession CMEN 585 (South Africa); alleles III and IV are from CMEN 584 (South Africa); allele V is from CMEN 17 (unknown European country); alleles VII and VIII are from CMEN 516 (Italy).

| ALLELES | I | II | III | IV | V | VII | VIII |
|---|---|---|---|---|---|---|---|
| I | - | 13 (7) | 2 (0) | 6 (1) | 74 (33) | 73 (32) | 75 (33) |
| II | | - | 13 (7) | 13 (8) | 71 (34) | 70 (33) | 72 (34) |
| III | | | - | 6 (1) | 75 (33) | 73 (32) | 75 (33) |
| IV | | | | - | 74 (34) | 73 (33) | 75 (34) |
| V | | | | | - | 3 (3) | 7 (4) |
| VII | | | | | | - | 6 (3) |
| VIII | | | | | | | - |

Figure 3.1. Schematic drawing of *Mentha longifolia mVe1* sequence (not to scale). The predicted coding region is shown in gray; 5' and 3' UTRs are white. A. Initial assembly of *mVe1* from degenerate, inverse and TAIL-PCR fragments. The 3' 1413 bp were obtained with degenerate PCR and inverse PCR (Vining et al., 2007). The 5' 1915 bp were obtained during the present study with TAIL-PCR. Relative locations of nested specific primers used in TAIL-PCR (sequences in Table 3.1) are indicated. B. Locations of *mVe1* primers used to clone coding regions from *M. longifolia* accessions. Clones included 40 bp of 3' UTR.

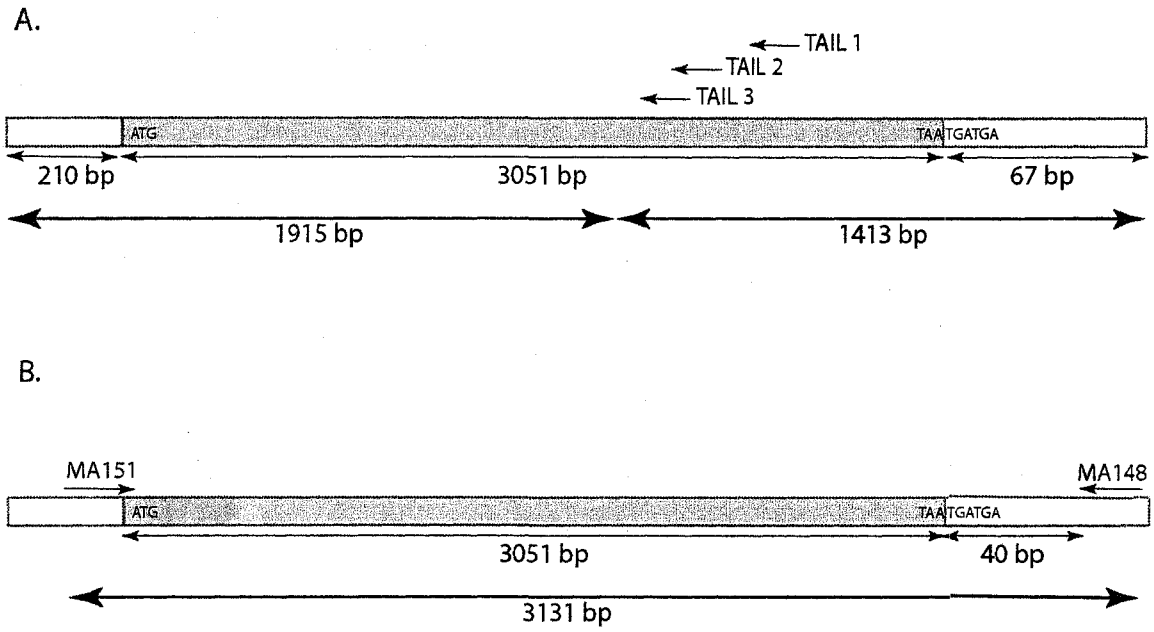


Figure 3.2. Primary amino acid structure of *M. longifolia* mVe1 deduced from genomic DNA sequence. The polypeptide has been divided into predicted domains A-E as described in the text, based on alignments with the tomato Ve putative protein sequences.Domain A: amino terminus. Domain B: leucine-rich repeat (LRR) domain, separated into individual, imperfect copies of the concensus motif: [XXIXNLXXLXXLXLSXNXLSGXIP]. Leucine residues are in boldface type. Domain C: extracytoplasmic domain; Domain D: transmembrane domain; Domain E: cytoplasmic domain.

```
A                MANLFLSVLMISIITATTFTTLSY                24
B           SQQCLHHQKTSLLQLKNELKFDSSNSTKLVQWNR           58
          KNNDCCNWYGVGCDGAGHVTSLQLDHEAISGGIDD            93
                  SSSLFRLEFLEKLNLAYNVFNRTQIP            119
                   RGIQNLTYLTHLNLSNAGFTGQVP             143
                    LQLSFLTRLVSLDISKFRRGIEP             166
             LKLERPNLETLLQNLSGLRELCLDGVDVSSQKS          199
                 EWGLIISSCLPNIRSLSLRYCSVSGPLH           227
                     ESLSKLQSLSILILDGNHLSSVVP           251
                     NFFANFSSLTTLSLKNCSLEGSFP           275
                     EMIFQKPTLQNLDLSQNMLLGGSIP          300
                      PFTQNGSLRSMILSQTNFSGSIP           323
                     SSISNLKSLSHIDLSYNRFTGPIP           347
                     STLGNLSELTYVRLWANFFTGSLP           371
                    STLFRGLSNLDSLELGCNSFTGYVP           396
                     QSLFDLPSLRVIKLEDNKFIGQVE           420
                  EFPNGINVSSHIVTLDMSMNLLEGHVP           447
                     ISLFQIQSLENLVLSHNSFSGTFQ           471
                     MKNVGSPNLEVLDLSYNNLSVDAN           495
                   VDPTWHGFPKLRELSLASCDLHAFPE           521
                       FLKHSAMIKLDLSNNRIDGQIP           543
                  RWIWGTELYFMNLSCNLLTDVQK               566
                       PYHIPASLQLLDLHSNRFKGDLHL         590
                  FISPIGDLTPSLYWLSLANNSFSGSIP           617
                     TSLCNATQLGVIDLSLNQLSGDIA           641
                    PCLLENTGHIQVLNLGRNNISGHIP           666
                     DNFPSQCGLQNLDLNNNAIQGKIP           690
                     KSLESCMSLEIMNVGDNSIDDTFP           714
                       CMLPPSLSVLVLRSNRFHGEVT           736
                    CERRGTWPNLQIIDISSNNFNGSLE           761
                  SINFSSWTAMVLMSDARFTQRHWG              785
            TNFLSASQFYYTAAVALTIKRVELELVKI               814
                         WPDFIAVDLSCNDFHGDIP            833
                     DAIGDLTSLYVLNISHNALGGSIP           857
                     KSLGQLSKLESLDLSRNRLSGHVP           881
                     TELGGLTFLSVLNLSYNELVGEIP           905
           NGRQMHTFSADAFKGNAGLCGRHLERNCSD               935
C                DRSQGEIEIENEIEWV                       951
D              YVFVALGYVVGLGIIVWLLLF                    972
E           CRSFRYKYFDKIDKVVQETFDARDRRRRRRRGTRIVRNQVVRRSH  1017
```

Figure 3.3. Clustal X alignment of *mVe1* deduced amino acid sequences from four *M. longifolia* accessions. Amino acids are colored as follows: M,L,V,I = green; G,S,T,P = brown; FWY = blue; R,H,K,= red; A,Q,N,C,E,D = white. Symbols above columns denote levels of conservation: an asterisk (*) for single, fully conserved residue; a colon (:) or a period (.) for conservation within one of the following "strong" or "weak" groups, respectively, as calculated by Clustal X (which in turn is based on the Gonnet Pam50 matrix, where a "strong" score is >0.5 and a "weak" score is <0.5). strong: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW; weak: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM, HFY.

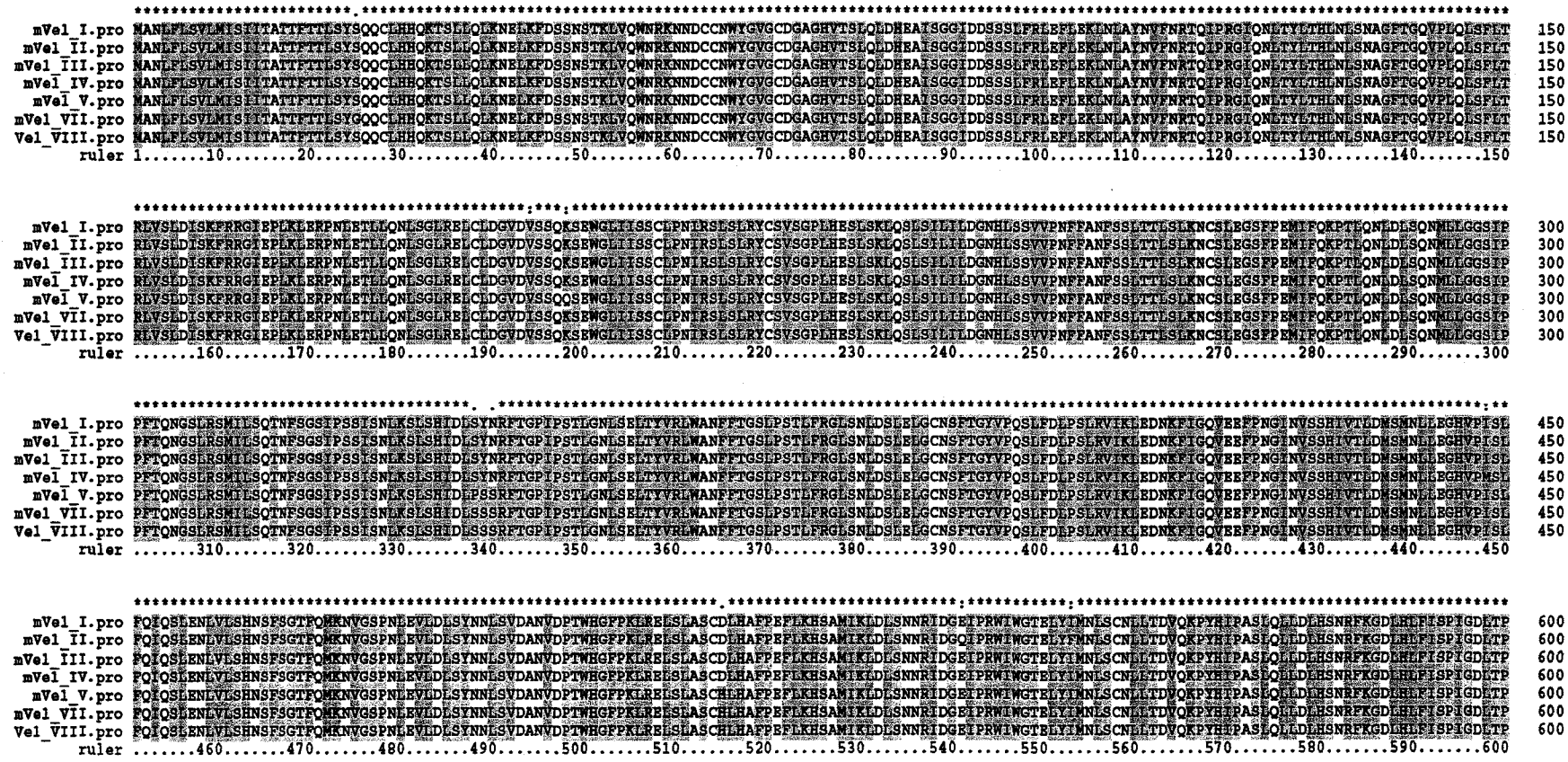


Figure 3.3. Continued.

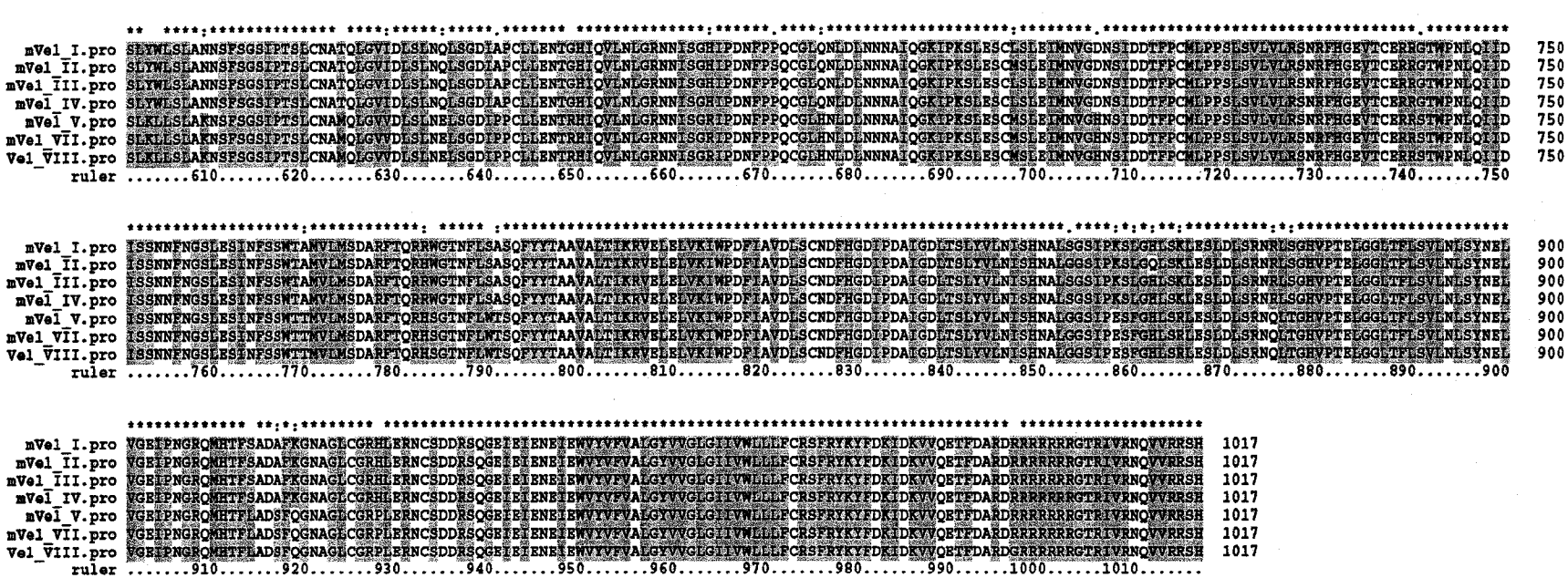


Figure 3.4. Restriction digest patterns used to distinguish *mVe1* alleles. Lane 1: 1Kb Plus ladder (Invitrogen), lane 2: CMEN 585, undigested; lane 3: CMEN 584, undigested; lane 4: CMEN 585, *Nco*I digested; lane 5: CMEN 584, *Nco*I digested; lane 6: CMEN 585, *Xba*I digested; lane 7: CMEN 584, *Xba*I digested.

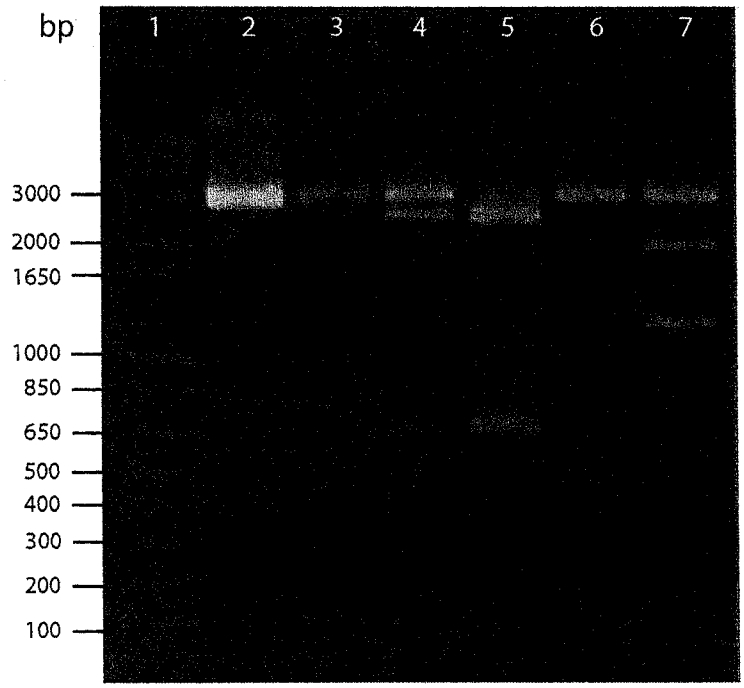
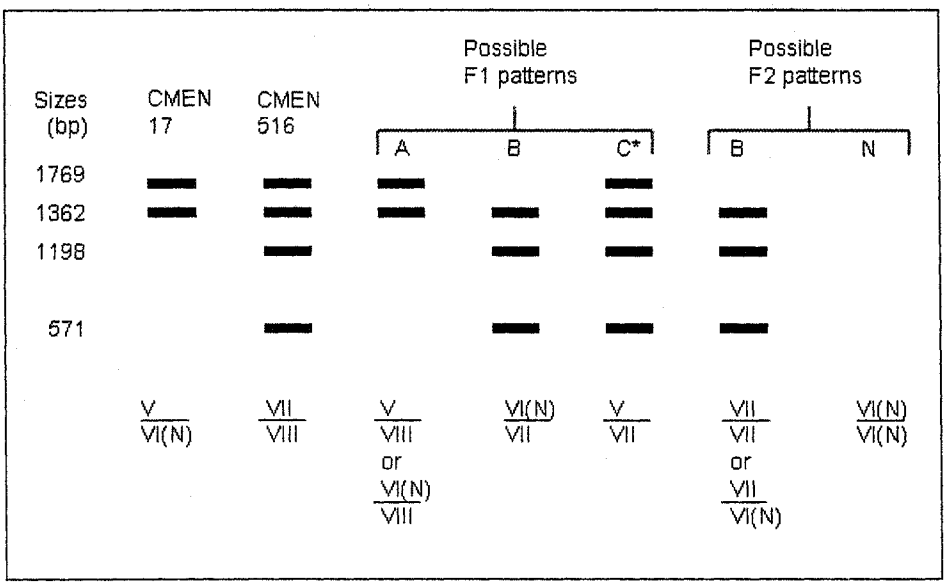


Figure 3.5. Genotypes of *mVe1* observed and expected in EF1 and EF2 populations.




**C* this expected pattern did not occur.**
**N = null allele (no PCR product)**

Figure 3.6. Disease symptom ratings and genotypes of plants in F1 (A.) and F2 (B.) populations derived from South African (SA) germplasm. Each graph bar represents an average rating of 6 clones, with standard deviations. Alleles of each plant, as identified with restriction digests, are represented with plusses (+). Parents of the SAF1 population are shown in boldface type. The SAF1 parent plant of the SAF2 population is boxed.

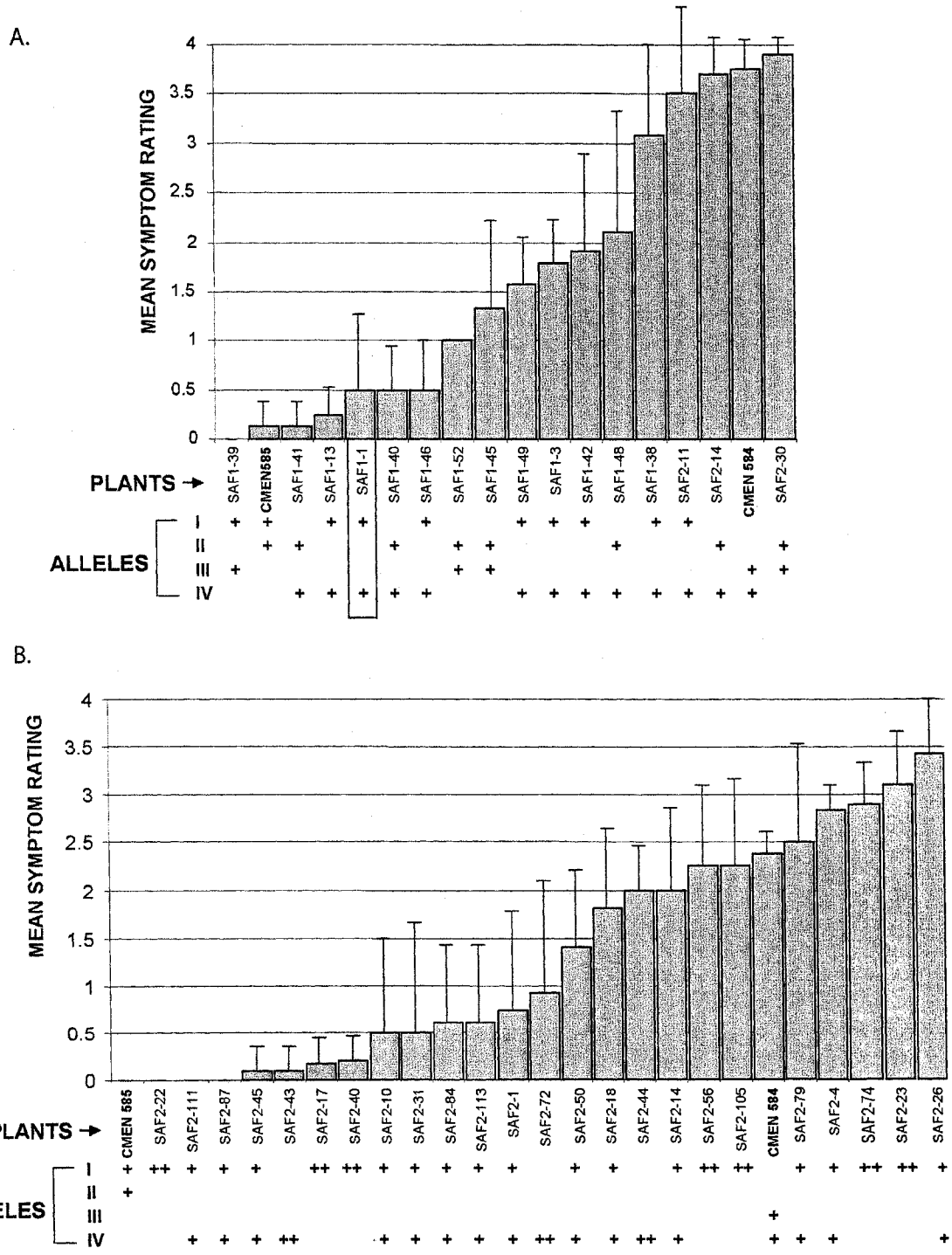
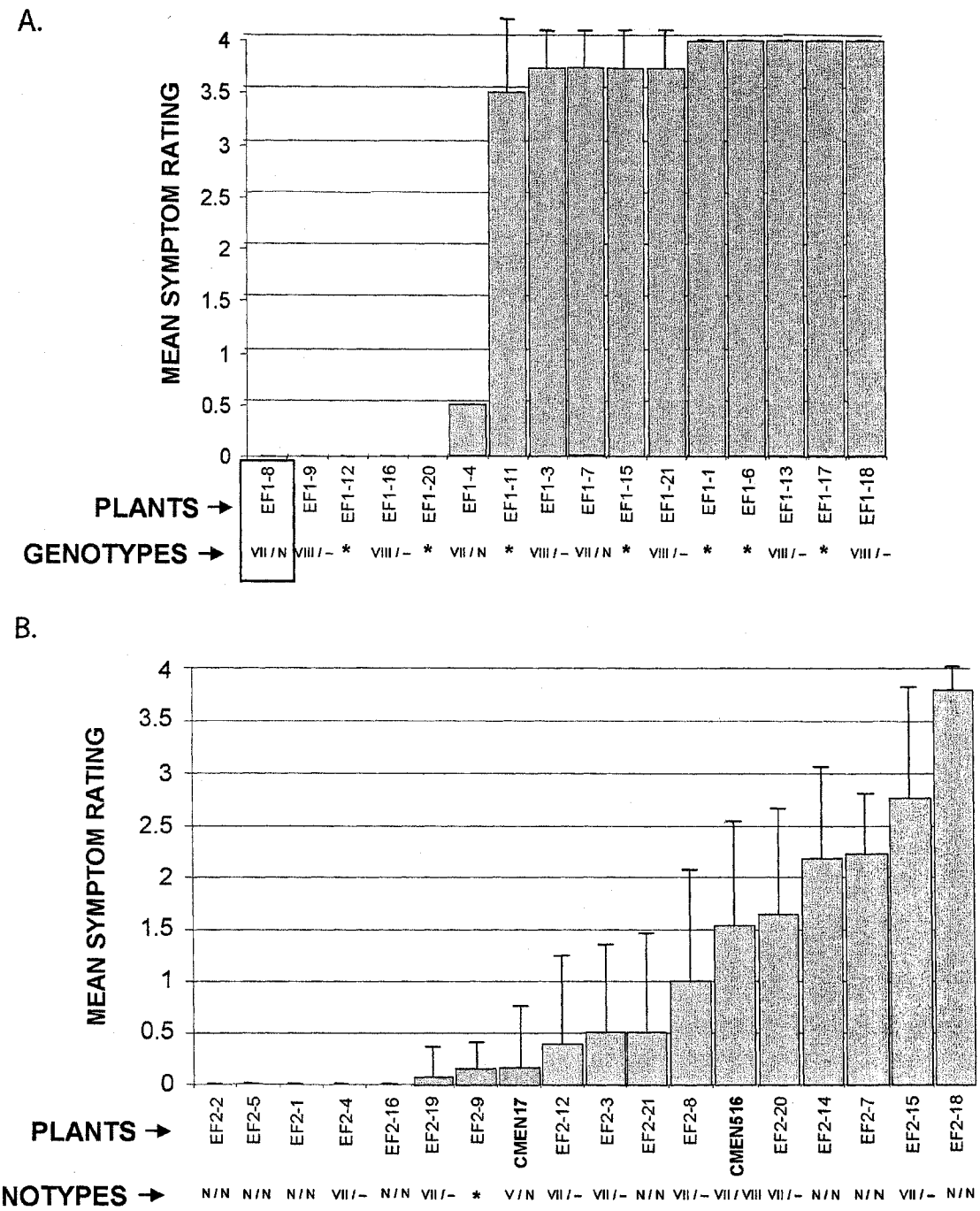


Figure 3.7. Disease symptom ratings and genotypes of plants in F1 (A.) and F2 (B.) populations derived from European (E) germplasm. Each graph bar represents an average rating of screened clones, with standard deviations. The allele composition of each plant is listed below the plant name; asterisks (*) represent plants that were not genotyped. Parents of the F1 generation are shown in boldface type. The EF1 parent plant of the EF2 generation is boxed.



# CONCLUSION

This project was initiated with the specific goal of identifying genes related to verticillium wilt resistance that could be used by the mint industry to genetically improve peppermint. Because of the complications of polyploidy, sterility, and genetic homogeneity in commercial peppermint, a wild diploid mint species, *Mentha longifolia*, was chosen as a focal point for genetic investigations. As the work progressed, the merits of this species as a generalized model for study of plant disease resistance and other genetic traits became apparent. Given this context, the results that have been presented in this dissertation are of three distinct types: 1) characterization and development of valuable germplasm resources; 2) refinement of protocols for reproducible verticillium wilt resistance screening; and 3) cloning of resistance gene candidates of two types. These categories reflect the history of the project, with the applied aspect of gene discovery and the broader, longer-term development of *M. longifolia* as a model plant for the study of mint genetics and plant-microbe interactions. The accomplishments in each of these areas are briefly summarized below, accompanied by recommendations for directions of future studies.

The work presented in Chapter I shows the evaluation of the mint species *M. longifolia* as a model plant for the study of mint genetics and plant pathology. The survey of USDA accessions uncovered diversity in plant morphology, oil type and verticillium wilt resistance. Because of the verticillium wilt observations, two

accessions from South Africa (CMEN 585, CMEN 584) and two accessions from Europe (CMEN 17, CMEN 516), each pair representing extremes of wilt resistance and susceptibility, became the focus of crossing efforts, and two F1 populations were produced. Screening of these F1 populations enabled identification of wilt-resistant individuals that were then allowed to self-pollinate to produce F2 populations. In verticillium wilt resistance trials of F2 populations, continuous variation in wilt phenotype was observed, indicating a resistance involving multiple genes. The F1 population (SAF1) produced from crossing South African accessions consists of 55 plants, and the respective SAF2 population has 120 plants. Additional wilt resistance screening trials could be conducted with these populations in order to increase the power of future genotype-phenotype association studies. The F1 and F2 populations derived from European accessions (EF1, EF2) were small, consisting of only ~20 plants each, and therefore, these populations would not be practical subjects for further study unless they were expanded.

Success in wilt resistance screening was dependent on methodological improvement. The employment of a growth chamber for strict environmental control during rooting of cuttings prior to inoculation, and then for four weeks post-inoculation, was critical. Strict short-day light conditions ensured that floral initiation was prevented, and allowed reproducible screening results to be obtained year-round with long-day-flowering mints. Post-inoculation cool temperatures ensured fungal spore survival and germination was not inhibited. At four weeks post-inoculation, when plants were moved to greenhouse benches,

warmer, more stressful conditions favored the development and observation of disease symptoms.

At the outset of the mint project, many plant genes conferring resistance to a variety of diseases had been isolated and sequenced from many plant species, but a verticillium wilt resistance gene was not among them. Therefore, a broad-based approach was taken to isolate generic resistance-gene-like sequences from *M. longifolia* with the intention of identifying and investigating candidate wilt resistance genes. The resistance gene analogs (RGAs) presented in Chapter II were the results of this effort. These RGAs have the potential to be useful genetic markers, and some of them may one day prove to be parts of functional disease resistance genes. However, when the tomato verticillium wilt resistance (*Ve*) genes were cloned and sequenced, the concentration of the mint project shifted to identification of a mint *Ve* homolog. This focus ultimately resulted in the cloning of the *mVe1* gene, as described in Chapters II and III. In addition, a second gene, *mVe2*, has recently been cloned and characterization is underway. The *mVe* gene sequences have been shared with the Mint Industry Research Council, and plasmid clones of both genes have been given to collaborators Rodney Croteau and Mark Wildung at Washington State University. Although no association between *mVe1* genotype and wilt phenotype was uncovered in studies of the *M. longifolia* F1 and F2 populations, a role for *mVe1* in conferring verticillium wilt resistance cannot be ruled out. Much remains to be learned about *mVe1*. For example, the *mVe1* promoter remains to be sequenced, and

expression studies have yet to be performed. Future studies must also examine *mVe2* and other genes that may function in wilt resistance.

While verticillium wilt was the only disease addressed in this dissertation, *M. longifolia* is a suitable subject for other plant pathology studies. For example, informal observations of greenhouse outbreaks of powdery mildew indicate that CMEN 682 is a particularly susceptible accession, while most of the other accessions appear to be resistant. The *M. longifolia* germplasm might also be evaluated for resistance to the fungal disease mint rust (caused by *Puccinia menthae*), or for resistance to a number of viruses that are known to infect mints.

In addition to plant pathology studies, the *M. longifolia* SAF1 and SAF2 populations have several features that make could them especially useful for other types of genetic investigations. In Chapter I, parental accessions CMEN 585 and CMEN 584 were reported to be at once very similar to each other and different from all the other *M. longifolia* accessions, with lanceolate leaves and a tall upright growth habit. Despite these similarities, CMEN 585 and CMEN 584 were also found to be very different from each other with respect to verticillium wilt resistance and to oil type. CMEN 585 was highly wilt-resistant, while CMEN 584 was highly wilt-susceptible. CMEN 585 had a more "pepperminty" scent, while CMEN 584 was the only carvone-rich ("spearminty") *M. longifolia* accession. Results of wilt screenings in Chapter III show that the SAF1 and SAF2 populations are segregating for wilt resistance/susceptibility. Casual observations indicate that these populations are also segregating for oil type, but there is no obvious correlation between oil type and wilt phenotype.

One genomics tool that is lacking for mint is a genetic map. The production of an *M. longifolia* map would aid candidate gene-phenotype association studies. RGAs and other resistance gene candidate markers would be important components of such a map.

A genomic library was constructed from the verticillium wilt-resistant South African USDA accession CMEN 585 (Appendix B). This library, consisting of an estimated 114,000 clones, could provide ~8.5x coverage of the *M. longifolia* genome. Only 18,432 clones were picked and spotted on filters. An existing repository of frozen clones could be picked and spotted to expand the library's potential usefulness as a resource for gene discovery.

The endeavor to accumulate *M. longifolia* genetic resources at UNH has been productive and includes phenotypic documentation, breeding populations, and a collection of DNA sequences representing genes and gene fragments. These resources are ultimately intended to benefit commercial mint production, and constitute a model system for study of plant-pathogen interactions. Commercial peppermint—an extremely homogeneous, polyploid crop—is now served by an especially diverse diploid genetic base and a growing foundation of knowledge and genomic resources.

Abou-jawdah, Y., Sobh, H. , Salameh, A. 2002. Antimycotic activities of selected plant flora, growing wild in Lebanon, against phytopathogenic fungi. J. Agric. Food Chem. 50(11):3208-3213.

Akita, M.and J.P.T. Valkonen. 2002. A novel gene family in moss (*Physcomitrella patens*) shows sequence homology and a phylogenetic relationship with the TIR-NBS class of plant disease resistance genes. J. Mol. Evol. 55(5):595-605.

Al-Shukri, M.M. 1968. Mode of penetration of *Fusarium* and *Verticillium* species into intact radicle of cotton in culture. Biologia Bratislava 23(10):819-824.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and L. D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215(3):403-410.

Bender, C.G., Shoemaker, P.B. 1984. Prevalence of verticillium wilt of tomato and virulence of *Verticillium dahliae* race 1 and race 2 isolates in western North Carolina. Plant Dis. 68(4):305-309.

Bennett, M.D., Leitch, I.J. 2005. Nuclear DNA amounts in angiosperms – progress, problems and prospects. Ann. Bot.-London. 95(1):45-90.

Berry, S.Z.and C.A. Thomas. 1961. Influence of soil temperature, isolates, and method of inoculation on resistance of mint to verticillium wilt. Phytopathology. 51(3):169-174.

Bhat, R.G.and K.V. Subbarao. 1999. Host range specificity in *Verticillium dahliae*. Phytopathology. 89(12):1218-1225.

Brandt, W.H., M.L. Lacy, and C.E. Horner. 1984. Distribution of *Verticillium* in stems of resistant and susceptible species of mint. Phytopathology. 74(5):587-591.

Brotman, Y., L. Silberstein, I. Kovalski, C. Perin, C. Dogimont, M. Pitrat, J. Klingler, G.A. Thompson, and R. Perl-Treves. 2002. Resistance gene homologues in melon are linked to genetic loci conferring disease and pest resistance. Theor. Appl. Genet. 104(6-7):1055-1063.

Bunsawat, J., N.E. Elliott, K.L. Hertweck, E. Sproles and L.A. Alice. 2004. Phylogenetics of Mentha (Lamiaceae): evidence from chloroplast DNA sequences. Syst. Bot. 29(4):959-964.

Burke, C., K. Klettke and R. Croteau. 2004. Heteromeric geranyl diphosphate synthase from mint: construction of a functional fusion protein and inhibition by bisphosphonate substrate analogs. Arch. Biochem. Biophys. 422(1):52-60.

Cannon, S.B., H. Zhu, A.M. Baumgarten, R. Spangler, G. May, D.R. Cook, and N.D. Young. 2002. Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. J. Mol. Evol. 54(4):548-562.

Chai, Y., L. Zhao, Z. Liao, X. Sun, K. Zuo, L. Zhang, S. Wang, and K. Tang. 2003. Molecular cloning of a potential *Verticillium dahliae* resistance gene *SlVe1* with multi-site polyadenylation from *Solanum licopersicoides*. DNA Sequence. 14(5):375-384.

Chambers, H.L.and K.E. Hummer. 1994. Chromosome counts in the *Mentha* collection at the USDA-ARS National Clonal Germplasm Repository. Taxon. 43(3):423-432.

Chin, D.B., R. Arroyo-Garcia, O.E. Ochoa, R.V. Kesseli, D.O. Lavelle, and R.W. Michelmore. 2001. Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (*Lactuca sativa*). Genetics. 157(4):831-849.

Cordero, J.C.and D.Z. Skinner. 2002. Isolation from alfalfa of resistance gene analogues containing nucleotide binding sites. Theor. Appl. Genet. 104(8):1283-1289.

Croteau, R., Gershenzon, J. 1994. Genetic control of monoterpene biosynthesis in mints (Mentha: Lamiaceae) p. 193-229. In: G.K. B.E. Ellis, H.A. Stafford (ed.). Genetic Engineering of Plant Secondary Metabolism. Plenum Press, New York, NY.

Diaz, V.and E. Ferrer. 2003. Genetic variation of populations of *Pinus oocarpa* revealed by resistance gene analog polymorphism (RGAP). Genome. 46(3):404-410.

Diwan, N., R. Fluhr, Y. Eshed, D. Zamir, and S.D. Tanksley. 1999. Mapping of *Ve* in tomato, a gene conferring resistance to the broad spectrum pathogen, *Verticillium dahliae* race1. Theor. Appl. Genet.. 98(2):315-319.

Dodds, P.N., G.J. Lawrence, A.-M. Catanzariti, T. Teh, C.-I.A. Wang, M.A. Syliffe, B. Kobe, and J.G. Ellis. 2006. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proc. Natl. Acad. Sci. USA 103(23):8888-8893.

Dodds, P.N., G.J. Lawrence, and J.G. Ellis. 2001. Six amino acid ahanges confined to the leucine-rich repeat B-strand/B-turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. Plant Cell. 13(1):163-178.

Douhan, L.I.and D.A. Johnson. 2001. Vegetative compatibility and pathogenicity of *Verticillium dahliae* from spearmint and peppermint. Plant Dis. 85(3):297-301.

Fei, J., Y. Chai, J. Wang, J. Lin, X. Sun, C. Sun, K. Zuo, and K. Tang. 2004. cDNA cloning and characterization of the *Ve* homologue gene *StVe* from *Solanum torvum* Swartz. DNA Sequence. 15(2):88-95.

Garas, N.A., S. Wilhem, and J.E. Sagan. 1986. Relationship of cultivar resistance to distribution of *Verticillium dahliae* in inoculated cotton plants and to growth of single conidia on excised stem segments. Phytopathology. 76(10):1005–1010.

Ghoulami, S., A. Il Idrissi and F. Fkih-Tetouani. 2001. Phytochemical study of *Mentha longifolia* of Morocco. Fitoterapia. 72(5):596-598.

Gobert, V., S. Moja, M. Colson, and T. P. 2002. Hybridization in the section *Mentha* (Lamiaceae) inferred from AFLP markers. Am. J. Bot. 89(12):2017-2023.

Graham, M.A., L.F. Marek, and R.C. Shoemaker. 2002. Organization, expression and evolution of a disease resistance gene cluster in soybean. Genetics. 162(4):1961-1977.

Green, R.J. 1951. Studies on the host range of the Verticillium that causes wilt of *Mentha piperita* L. Science. 113(2930):207-208.

Green, R.J., Simon. J.E. 1996. Evaluations of MIRC peppermint and spearmint lines. Mint Industry Research Council Annual Proceedings.

Hefendehl, F.W. 1977. Monoterpene composition of a carvone containing polyploidy strain of *Mentha longifolia* (L.) Huds. Herba Hung. 16(1):39-43.

Hefendehl, F.W.and M.J. Murray. 1976. Genetic aspects of the biosynthesis of natural odors. Lloydia. 39(1):39-52.

Hendriks, H., F.H.L. van Os, and W.J. Feenstra. 1976. Crossing experiments between some chemotypes of *Mentha longifolia* and *Mentha suaveolens*. Planta Med. 30(2):154-162.

Hill, M.K., K.J. Lyon, and B.R. Lyon. 1999. Identification of disease response genes expressed in *Gossypium hirsutum* upon infection with the wilt pathogen *Verticillium dahliae*. Plant Mol. Biol. 40(2):289-296.

Hu, G., A.K.A. deHart, Y. Li, C. Ustach, V. Handley, R. Navarre, C.-F. Hwang, B.J. Aegerter, V.M. Williamson, and B. Baker. 2005. *EDS1* in tomato is required for resistance mediated by TIR-class *R* genes and the receptor-like *R* gene *Ve*. Plant J. 42(3):376-391.

Hubbard Center for Genome Studies, University of New Hampshire. 2002. Sequencing reactions for the ABI377. Accessed 4 Feb. 2004. <http://hcgs.unh.edu/protocol/sequence>

Irigoyen, M.L., Y. Loarce, A. Fominaya, and E. Ferrer. 2004. Isolation and mapping of resistance gene analogs from the *Avena strigosa* genome. Theor. Appl. Genet. 109(4):713-724.

Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. Bull. Soc. Vaudoise Sci. Nat. 44(5):223-270.

Jebanathirajah, J.A., S. Peri, and A. Pandey. 2002. Toll and interleukin-1 receptor (TIR) domain-containing proteins in plants: a genomic perspective. Trends Plant Sci. 7(9):388-391.

Johnson, D.A.and T.F. Cummings. 2000. Evaluation of mint mutants, hybrids, and fertile clones for resistance to *Verticillium dahliae*. Plant Dis. 84(3):235-238.

Joyeux, A., M.G. Fortin, R. Mayerhofer, and A.G. Good. 1999. Genetic mapping of plant disease resistance gene homologues using a minimal *Brassica napus* L. population. Genome. 42(4):735-743.

Kaiser, H. and W. Lang. 1951. Ueber die bestimmung des ätherischen oels in drogen. Deutsche Apotheker-Z./Süddeutsche Apotheker-Z. 91(4):163-166.

Kanazin, V., L.F. Marek, and R.C. Shoemaker. 1996. Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA. 93(21):11746-11750.

Kawchuk, L.M., J. Hachey, D.R. Lynch, F. Kulcsar, G. van Rooijen, D.R. Waterer, A. Robertson, E. Kokko, R. Byers, R.J. Howard, R. Fischer, and D. Prufer. 2001. Tomato *Ve* disease resistance genes encode cell surface-like receptors. Proc. Natl. Acad. Sci. USA98(11):6511-6515.

Kokkini, S., R. Karousou, and T. Lanaras. 1995. Essential oils of spearmint (carvone-rich) plants from the island of Crete (Greece). Biochem. Syst. Ecol. 23(4):425-430.

Kokkini, S.and V.P. Papageorgiou. 1988. Constituents of essential oils from *Mentha longifolia* growing wild in Greece. Planta Med. 1(1):59-60.

Krasnyanski, S., T. Ball, and K.C. Sink. 1998. Somatic hybridization in mint:identification and confirmation of Mentha piperita (+) M. spicata hybrid plants. Theor. Appl. Genet.. 96(5):683-687.

Kruijt, M., B.F. Brandwagt, and P.J.G.M. de Wit. 2004. Rearrangements in the *Cf-9* disease resistance gene cluster of wild tomato have resulted in three genes that mediate *Avr9* responsiveness Genetics. 168(3):1655-1663.

Lacy, M.L.and C.E. Horner. 1965. Verticillium wilt of mint: interactions of inoculum density and host resistance. Phytopathology. 55(11):1176-1178.

Landing, J.E. 1969. American Essence: a History of the Peppermint and Spearmint Industry in the United States. Kalamazoo Public Museum.

Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D. Croteau, R. 2000. Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. Proc. Natl. Acad. Sci. USA. 97(6):2934-2939.

Leister, D., A. Ballvora, F. Salamini, and C. Gebhardt. 1996. A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nat. Genet. 14(4):421-429.

Lescot, M., S. Rombauts, J. Zhang, S. Aubourg, C. Mathé, S. Jansson, P. Rouzé, and W. Boerjan. 2004. Annotation of a 95-kb *Populus deltoides* genomic sequence reveals a disease resistance gene cluster and novel class I and class II transposable elements. Theor. Appl. Genet.. 109(1):10-22.

Liu, J.-J.and A.K.M. Ekramoddoullah. 2003. Isolation, genetic variation and expression of TIR-NBS-LRR resistance gene analogs from western white pine (*Pinus monticola* Dougl. ex. D. Don.). Mol. Genet. Genomics. 270(5):432-441.

Liu, Y.-G., N. Mitsukawa, T. Oosumi, and R.F. Whittier. 1995. Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. 8(3):457-463.

Liu, Y.G.and R.F. Whittier. 1995. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. Genomics. 25(3):674-681.

Lopez, C.E., I.F. Acosta, C. Jara, F. Pedraza, E. Gaitan-Solis, G. Gallego, S. Beebe, and J. Tohme. 2003. Identifying resistance gene analogs associated with resistances to different pathogens in common bean. Phytopathology. 93(1):88-95.

Lorang, J.M., R.P. Tuori, P. Martinez, T.L. Sawyer, R.S. Redman, J.A. Rollins, T.J. Wolpert, K.B. Johnson, R.J. Rodriguez, M.B. Dickman, and L.M. Ciuffetti. 2001. Green fluorescent protein is lighting up fungal biology. Appl. Environ. Microbiol. 67(5):1987-1994.

Mahmoud, S.S.and R.B. Croteau. 2001. Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. Proc. Natl. Acad. Sci. USA. 98(15):8915-8920.

Mahmoud, S.S., M. Williams, and R.B. Croteau. 2004. Cosuppression of limonene-3-hydroxylase in peppermint promotes accumulation of limonene in the essential oil. Phytochemistry. 65(5):547-554.

Maleki, L., J.D. Faris, R.L. Bowden, B.S. Gill, and J.P. Fellers. 2003. Physical and genetic mapping of wheat kinase analogs and NBS-LRR resistance gene analogs. Crop Sci. 43(2):660-670.

Martinez Zamora, M.G., A.P. Castagnaro, and J.C. Diaz Ricci. 2004. Isolation and diversity analysis of resistance gene analogues (RGAs) from cultivated and wild strawberries. Mol. Genet. Genomics. 272(4):480-487.

McDowell, J.M., M. Dhandaydham, T.A. Long, M.G.M. Aarts, S. Goff, E.B. Holub, and J.L. Dangl. 1998. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of *Arabidopsis*. Plant Cell. 10(11):1861-1874.

McFadden, H.G., R. Chapple, R. deFeyter, and E. Dennis. 2001. Expression of pathogenesis-related genes in cotton stems in response to infection by *Verticillium dahliae*. Physiol. Mol. Plant P. 58(3):119-131.

Meyers, B.C., A.W. Dickerman, R.W. Michelmore, S. Sivaramakrishnan, B.W. Sobral, and N.D. Young. 1999. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J. 20(3):317-332.

Meyers, B.C., S. Kaushik, and R.S. Nandety. 2005. Evolving disease resistance genes. Curr. Opin. Plant Biol. 8(2):129-134.

Meyers, B.C., K.A. Shen, G. Pejman Rohani, Brandon S., and R.W. Michelmore. 1998. Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell. 10(11):1833-1846.

Mimica-Dukic, N., B. Bozin, M. Sokovic, B. Mihajlovic, and M. Matavulj. 2003. Antimicrobial and antioxidant activities of three *Mentha* species essential oils. Planta Med. 69(5):413-419.

Monosi, B., R.J. Wisser, L. Pennill, and S.H. Hulbert. 2004. Full-genome analysis of resistance gene homologues in rice. Theor. Appl. Genet.1091434-1447.

Murray, M.J.and W.A. Todd. 1972. Registration of Todds Mitcham peppermint. Crop Sci. 12(1):128.

Murray, M.J.and W.A. Todd. 1975. Role of mutation breeding in genetic control of plant diseases, pp. 172-176. In: G.W. Bruehl (ed.). Biology and Control of Soilborne Plant Pathogens. American Phytopathological Society, St. Paul, MN.

National Agricultural Statistics Service, United States Department of Agriculture. 2006. Agricultural Statistics 2006. Report number 01-016-036158-3. Accessed July 19, 2006. <http://www.nass.usda.gov/Publications/Ag_Statistics/agr06/agstats2006.pdf.>

Nelson, R. 1926. in: Plant Disease Reporter, Supplement 50. United States Department of Agriculture, Office of Mycology and Disease Survey. p. 474.

Niu, X., P.M. Hasegawa, R.A. Bressan, and S.C. Weller. 1998. Transgenic peppermint plants obtained by co-cultivation with *Agrobacterium tumefaciens*. Plant Cell Rep. 17(3):165-171.

Noir, S., M.-C. Combes, F. Anthony, and P. Lashermes. 2001. Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (*Coffea* L.). Mol. Genet. Genomics. 265(4):654-662.

Peñuela, S., D. Danesh, and N.D. Young. 2002. Targeted isolation, sequence analysis, and physical mapping of nonTIR NBS-LRR genes in soybean. Theor. Appl. Genet. 104(2-3):261-272.

International Rice Genome Sequencing Project. 2005. The map-based sequence of the rice genome. Nature. 436(7052):793-800.

Rauyaree, P., M.D. Ospina-Giraldo, S. Kang, R.G. Bhat, K.V. Subbarao, S.J. Grant, and K.F. Dobinson. 2005. Mutations in *VMK1*, a mitogen-activated protein kinase gene, affect microsclerotia formation and pathogenicity in *Verticillium dahliae*. Curr. Genet. 48(2):109-116.

Saraste, M., P.R. Sibbald, and A. Wittinghofer. 1990. The P-loop-a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. 15(11):430-434.

Schiable, L., O.S. Cannon, and V. Waddoups. 1951. Inheritance of resistance to verticillium wilt in a tomato cross. Phytopathology. 41(10):986-990.

Seah, S., A.C. Telleen, and V.M. Williamson. 2007. Introgressed and endogenous *Mi-1* gene clusters in tomato differ by complex rearrangements in flanking sequences and show sequence exchange and diversifying selection among homologues. Theor. Appl. Genet. 22 Feb 2007 Epub ahead of print.

Shaiq, A.M., M. Saleem, W. Ahmad, M. Parvez, and R. Yamdagni. 2002. A chlorinated monoterpene ketone, acylated beta-sitosterol glycosides and a flavanone glycoside from *Mentha longifolia* (Lamiaceae). Phytochemistry. 59(8):889-895.

Shen, K.A., B.C. Meyers, M. Nurul Islam-Faridi, D.B. Chin, D.M. Stelly, and R.W. Michelmore. 1998. Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. Mol. Plant-Microbe Interact. 11(8):815-823.

Simko, I., S. Costanzo, K.G. Haynes, B.J. Christ, and R.W. Jones. 2004. Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in a tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor. Appl. Genet. 108(2):217-224.

Simko, I., K.G. Haynes, and R.W. Jones. 2004. Mining data from potato pedigrees: tracking the origin of susceptiility and resistance to *Verticillium dahliae* in North American cultivars through molecular marker analysis. Theor. Appl. Genet.. 108(2):225-230.

Sinclair, W.A., H.H. Lyon, and W.T. Johnson. 1987. Diseases of trees and shrubs. Comstock Publishing Associates, Cornell University Press, Ithaca and London.

Sink, K.C.and W.E. Grey. 1999. A root-injection method to assess verticillium wilt resistance of peppermint (*Mentha xpiperita* L.) and its use in identifying resistant somaclones of cv. Black Mitcham. Euphytica. 106(3):223-230.

Smith, L.D.and D. Neely. 1979. Relative susceptibility of tree species to *Verticillium dahliae*. Plant Dis. Rptr. 63(4):328-332.

Soriano, J.M., S. Vilanova, C. Romero, G. Llácer, and M.L. Badenes. 2004. Characterization and mapping of NBS-LRR resistance gene analogs in apricot (*Prunus armeniaca* L.). Theor. Appl. Genet. 110(5):980-989.

Steventon, L.A., P. Okori, and C. Dixelius. 2001. An investigation of the susceptibility of *Arabidopsis thaliana* to isolates of two species of *Verticillium*. J. Phytopathology. 149(7-8):395-401.

Theodoulou, F.L. 2000. Plant ABC transporters. Biochim. Biophys. Acta. 1465(1-2):79-103.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25(24):4876-4882.

Timmerman-Vaughan, G.M., T.J. Frew, and N.F. Weeden. 2000. Characterization and linkage mapping of R-gene analogous DNA sequences in pea (*Pisum sativum* L.). Theor. Appl. Genet. 101(1-2):241-247.

Tjamos, E.C.and I.M. Smith. 1975. The expression of resistance to *Verticillium albo-atrum* in monogenically resistant tomato cultivars. Physiol. Plant Pathol. 6(1-2):215-225.

Todd, W.A., R.J. Green, and C.E. Horner. 1977. Registration of Murray Mitcham peppermint. Crop Sci. 17(1):188.

Torres, A.M., N.F. Weeden, and A. Martin. 1993. Linkage among isozyme, RFLP, and RAPD markers in *Vicia faba*. Theor. Appl. Genet. 85(8):937-945.

Traut, T.W. 1994. The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. Eur. J. Biochem. 222(1):9-19.

Tucker, A.O., R.M. Harley, and D.E. Fairbrothers. 1980. The Linnaean types of *Mentha* (Lamiaceae). Taxon. 29(2-3):233-235.

Tucker, A.O., H. Hendricks, R. Bos, and D.E. Fairbrothers. 1990. The origin of *Mentha x gracilis* (Lamiaceae). I. Chromosome numbers, fertility, and three morphological characters. Econ. Bot. 44(2):183-213.

Tucker, A.O., H. Hendriks, R. Bos, and D.E. Fairbrothers. 1991. The origin of *Mentha x gracilis* (Lamiaceae). II. Essential oils. Econ. Bot. 45(2):200-215.

Tucker, A.O.and R.F.C. Naczi. 2005. *Mentha*: An overview of its classification and relationships. In: B.M. Lawrence (ed.). Mints: The genus *Mentha*. Taylor & Francis, London.

Udo, S., S. Shimizu, and N. Ikeda. 1962. Studies on the origin of *Mentha piperita* L. Sci. Rep. Fac. Agric. Okayama Univ. 201-12.

United States Department of Agriculture. National Genetic Resources Program. Germplasm Resources Information Network (GRIN) National Germplasm Resources Laboratory, Beltsville, Maryland.

University of Illinois at Ubana-Champaign, Department of Crop Sciences. 1997. Report on plant disease: verticillium wilt disease. RPD 1010. Accessed December 11, 2006. <http://web.aces.uiuc.edu/vista/pdf_pubs/1010.pdf>

van der Hoorn, R.A.L., R. Roth, and P.J.G.M. De Wit. 2001. Identification of distinct specificity determinants in resistance protein Cf-4 allows construction of a Cf-9 mutant that confers recognition of avirulence protein AVR4. Plant Cell. 13(2):273-285.

van der Vossen, E.A.G., J.N.A.M.R. van der Voort, K. Kanyuka, A. Bendahmane, H. Sandbrink, D.C. Baulcombe, J. Bakker, W.J. Stiekema, and R.M. Klein-Lankhorst. 2000. Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. Plant J. 23(5):567-576.

Venskutonis, P.R. 1996. A chemotype of *Mentha longifolia* L. from Lithuania rich in piperitenone oxide. Journal Essent. Oil Res. 8(1):91-95.

Veronese, P., M.L. Narasimhan, R.A. Stevenson, J.-K. Zhu, S.C. Weller, K.V. Subbarao, and R.A. Bressan. 2003. Identification of a locus controlling verticillium disease symptom response in *Arabidopsis thaliana*. Plant J. 35(5):574-587.

Vining, K.J., Q. Zhang, S.C. A., T.A. O., and T.M. Davis. 2005. *Mentha longifolia* (L.) L.: a model species for mint genetic research. HortScience. 40(5):1225-1229.

Vining, K.J., Q. Zhang, C.A. Smith, and T.M. Davis. 2007. Resistance gene analogs and *Verticillium* resistance-like sequences from *Mentha longifolia*. Journal of the American Society for Horticultural Science (in press).

von Rudloff, E. 1969. Scope and limitations of gas chromatography of terpenes in chemosystematics studies. Rec. Adv. Phytochem. 2(1):127-162.

Wei, F., R.A. Wing, and R.P. Wise. 2002. Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. Plant Cell. 14(8):1903-1917.

Williams, J.S., S.A. Hall, M.J. Hawkesford, M.H. Beale, and R.M. Cooper. 2002. Elemental sulfur and thiol accumulation in tomato and defense against a fungal vascular pathogen. Plant Physiol. 128(1):150-159.

Yaish, M.W.F., L.E. Sáenz de Miera, and M. Pérez de la Vega. 2004. Isolation of a family of resistance gene analogue sequences of the nucleotide binding site (NBS) type from *Lens* species. Genome. 47(4):650-659.

Yu, Y.G., G.R. Buss, and M.A. Saghai Maroof. 1996. Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc. Natl. Acad. Sci. USA. 93(21):11751-11756.

# APPENDICES

# APPENDIX A

Examples of verticillium wilt symptoms with associated numerical ratings. This rating system is a slight modification of a rating scale provided by Dennis Johnson[1] at Washington State University.

ZERO
(0)

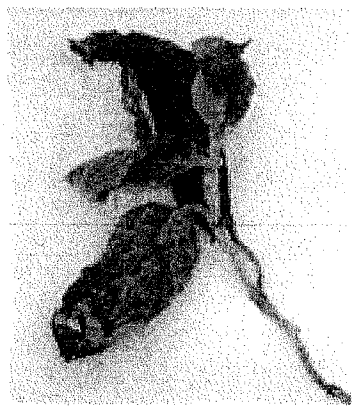
healthy

THREE
(3)

severe chlorosis, severe stunting, 10-60% necrotic

ONE
(1)

mild chlorosis or crescent leaf



FOUR
(4)

dead plant

TWO
(2)

distinct crescent leaf, chlorosis, mild stunting

[1] Dennis Johnson, Washington State University, Department of Plant Pathology, P.O. Box 646430, Pullman, WA 99164-6430.

APPENDIX B

## *Mentha longifolia* Genomic Library Construction

A CopyControl™ Fosmid Library Production Kit (Epicentre, Madison, WI) was used to construct a genome library from a verticillium wilt-resistant *Mentha longifolia* accession. The manufacturer's protocol was followed, with exceptions noted below.

### DNA extraction

DNA was isolated from unexpanded leaves of accession CMEN 585 (PI557767) using the CTAB method described in Chapter I with volumes scaled up to accommodate 1.8 grams of starting tissue.

### Library host cell preparation

*Eschericia coli* plating strain EPI300™-T1R was provided with the Epicentre CopyControl™ Fosmid Library Production Kit. EPI300™-T1R cells were streaked on LB agar medium and incubated at 37 °C overnight. Plates were sealed and stored at 4 °C until the day prior to clone packaging and cell transformation. Fifty milliliters LB containing 10 mM $MgSO_4$ was inoculated with a single colony of the cultured cells and incubated at 37 °C for 24 h at 180 rpm on an orbital shaker. The day of the packaging reactions, 50 ml fresh enriched broth medium was

inoculated with 5 ml of the liquid cell culture and incubated at 37 °C for 4 h at 180 rpm on an orbital shaker to attain an OD600 = 0.8 - 1.0.

**End-repair of sheared DNA**

The end-repair reaction was prepared as described in the Epicentre protocol, and consisted of 20.2 µg genomic DNA (23.5 µl @ 861 ng/µl), 1X End-Repair buffer, 0.25 mM of each dNTP, 1 mM ATP, and 4 µl End-Repair enzyme mix in a final volume of 80 µl. The reaction was incubated at room temperature for 45 min. The enzyme was inactivated by incubation at 70 °C for 10 min.

**Size-selection of end-repaired DNA**

Size-selection was performed as described in Method 1 of the Epicentre protocol. End-repaired insert DNA was electrophoresed on a 0.8% SeaPlaque® GTG® low melting point agarose (Cambrex, East Rutherford, NJ) gel in 1X TAE buffer. A T7 DNA Size Marker (Epicentre) was loaded in an adjacent lane of the gel. Subsequent to electrophoresis, the marker lane was cut from the gel, stained with ethidium bromide and visualized on a uv light box. The T7 marker was identified, and its position in the gel marked by cuts made with a clean glass cover slip above and below the marker fragments. The stained gel piece was covered with plastic wrap to protect the unstained gel piece from ethidium bromide contamination. The stained and unstained gel pieces were then realigned without uv light and the 40 kb region of the insert DNA was excised from the unstained gel and transferred to a 1.5 ml microfuge tube.

**Recovery of insert DNA**

Size-selected insert DNA was recovered from the gel slice using slight modifications to the Epicentre protocol. GELase Buffer was not added to the gel slice to be melted. The gel slice (total 320 mg) was melted in a hybridization oven at 70 °C for 7 min. Digestion was performed by incubation of the melted gel with 6.4 U GELase™ (Epicentre) in a 1.5-ml microfuge tube for 1 h 10 min at 45 °C. The enzyme was inactivated by incubation at 70 °C for 10 min. The tube was then chilled on ice for 5 min and centrifuged at 14,000 x g 20 min at 4 °C to pellet insoluble polysaccharides. The supernatant (~250 mg) was transferred to a fresh 1.5 ml microfuge tube and DNA was precipitated with the addition of 1/10 volume 3M sodium acetate, pH 7.0, and 2.2 volumes 95% ethanol at room temperature for 10 min. DNA was pelleted by centrifugation at 14,000 x g for 30 min at 4 °C. Pelleted DNA was washed twice with 700 µl 70% ethanol, allowed to air dry in a laminar flow hood for 15 min, and then resuspended in 15 µl TE overnight at 4 °C. A fluorometer reading of the recovered DNA was 160 ng/µl.

**Fosmid vector ligation**

A single ligation reaction was prepared with 272 ng (1.7 µl) of the recovered insert DNA in 1X Fast-Link™ Ligation buffer (Epicentre), 1mM ATP, 0.5 µg pCC1FOS™ vector and 2 U Fast-Link™ DNA ligase (Epicentre) in a 10 µl final volume, as described in the Epicentre protocol. The reaction was incubated at room temperature for 2 h and the enzyme inactivated at 70 °C for 10 min. A parallel control ligation was prepared with 500 ng pCC1FOS™ vector and 250 ng Fosmid Control Insert DNA (Epicentre).

**Vector packaging and titering of packaged clones**

Packaging and titering of fosmid clones were performed as described in the Epicentre protocol. MaxPlax™ Lambda Packaging Extracts were thawed on ice just prior to use. For each packaging reaction, 25 µl of the thawed packaging extract was transferred to a sterile 1.5 ml microfuge tube into which a completed 10 µl ligation reaction was added. Each reaction was mixed by pipetting. Reactions were then incubated at 30 °C for 1.5 h. An additional 25 µl of thawed packaging extract was added and the reactions were incubated a further 1.5 h at 30 °C. After the final incubation period, 940 µl Phage Dilution Buffer was added to bring the total volume to 1 ml. Twenty-five microliters chloroform was added and the suspension was mixed by inversion.

To titer packaged clones, serial dilutions of 100 µl of the packaging reaction were made in 10-fold increments to $10^{-4}$ with Phage Dilution Buffer. Ten microliters of each dilution was then transferred to 100 µl of the prepared EPI300™-$T1^R$ cell culture and incubated at 37 °C for 20 min for transfection. These cells were then plated on LB agar containing 12.5 µg/ml chloramphenicol and incubated at 37 °C overnight to select for successful transfections. All remaining packaged phage particle dilutions were stored at 4 °C.

The entire control ligation reaction was packaged as described above. The packaged recombinant control DNA was diluted 1:1000 in Phage Dilution Buffer and 10 µl of diluted control was transfected into 100 µl of the prepared EPI300™-$T1^R$ cell culture.

Colonies were counted for each dilution plate and the number of colony forming units (cfu) per milliliter of culture was calculated. From this calculation, the total number of clones in the library was estimated to be 114,000, which represents ~8.5x coverage of the *M. longifolia* genome.

**Library storage**

A fresh culture of EPI300™-T1$^R$ cells was grown in LB + 10 mM $MgSO_4$ to $OD_{600}$ = 0.9 (~3 h). Ten microliters of phage-packaged clones were diluted with 990 µl phage dilution buffer, and the entire aliquot (1ml) was added to 10 ml cells and allowed to adsorb to the cells for 20 min at room temperature. The cells were then incubated at 37 °C for 30 min at 125 rpm on an orbital shaker. Cells were pelleted at 8000 x g for 5 min. The supernatant was discarded and the pellets were resuspended in 10 ml of LB + 20% glycerol. One hundred microliter aliquots of the suspension were pipetted into 1.5 ml microfuge tubes which were stored at -80°C.

**Picking and arraying clones**

A total of 18,432 fosmid clones was manually picked in a laminar flow hood using sterile toothpicks into 384-well plates (Genetix, Boston, MA) containing 0.2 ml LB + 20% glycerol per well. These clones were spotted in duplicate onto Performa II high-performance, positively charged, 22 cm x 22 cm nylon high-density filters (Genetix) using a Genetix Qbot at the Hubbard Center for Genome Studies (HCGS), UNH. Plates were stored at -80°C.

# APPENDIX C

*Mentha longifolia mVe1* draft sequence, assembled from degenerate PCR, IPCR and TAIL PCR products. Start and stop codons are underlined.

GGTCGAGAGAGAAGAAAAACTAACAGAAAAAAAGATAAACTAAAGAAGAAAA
AATCTAAAATATGGGACCAAATTAATTGTGGAATCTGCCACGAAATTGAAGA
GAAGAAATCTGAGAAGCAGCAGCCATATGCATATATACATGGATACAAATAC
GTATATACCACAGCTAGCTGTAGTTTGCACACATCATCACTCATCTCCTTCA
CAATGGCGAACTTGTTCTTATCTGTACTCATGATCTCTATTATTACAGCAACA
ACTTTTACTACTCTTTCCTACAGCCAACAGTGTCTCCACCATCAAAAAACTTC
GTTGCTTCAACTGAAGAATGAGTTGAAATTCGATTCTTCTAATTCAACAAAAC
TGGTGCAATGGAATCGAAAAAACAACGACTGCTGCAACTGGTACGGGGTGG
GATGTGACGGTGCAGGCCACGTCACGAGTTTGCAGCTCGACCATGAGGCC
ATTTCCGGTGGAATCGATGATTCGTCGAGTCTGTTCAGACTCGAGTTTCTTG
AGAAGCTCAACCTAGCCTACAACGTCTTCAACCGCACTCAGATTCCAAGAG
GTATTCAGAATCTCACGTATTTGACACACTTGAATTTGTCGAATGCTGGTTTC
ACTGGGCAGGTTCCACTTCAACTTTCCTTCTTGACAAGATTAGTTAGTCTCG
ACATCTCCAAGTTCCGTAGGGGCATCGAGCCTCTAAAACTCGAGCGCCCAA
ATTTGGAGACGCTTCTCCAGAATCTGAGTGGACTCAGAGAGCTCTGTCTTGA
TGGTGTCGATGTTTCGTCTCAAAAGAGTGAGTGGGGTCTTATTATCTCTTCC
TGTTTACCAAACATTAGGAGTTTGAGCTTGCGTTATTGTAGTGTTTCTGGGC
CTTTGCATGAGAGCCTTTCGAAACTTCAATCCCTTTCCATTCTTATACTAGAT
GGGAATCACCTTTCCTCAGTTGTCCCCAACTTCTTTGCAAATTTCTCAAGTTT
GACAACTTTGAGTCTTAAAAATTGCTCTTTAGAAGGTTCTTTCCCTGAGATGA
TCTTCCAAAAACCCACTCTACAAAATCTCGATTTATCTCAGAATATGCTACTC
GGTGGAAGCATACCGCCATTTACTCAAAATGGATCTCTTAGGTCTATGATCC
TCAGCCAGAACAACTTCTCGGGGTCGATACCAAGTTCAATTAGCAATCTCAA
ATCGTTGTCGCATATAGACCTATCTTATAATAGATTCACAGGACCTATTCCAT
CCACATTGGGTAACCTATCTGAGTTGACATATGTCCGTTTATGGGCTAACTT
CTTCACCGGCTCACTTCCTTCCACGTTGTTTCGAGGTCTATCCAATCTTGAT
AGTTTAGAATTAGGGTGTAACTCATTCACCGGTTACGTACCCCAGTCTCTCT
TTGATCTCCCTTCACTGCGGGTAATTAAGCTTGAAGACAACAAATTTATTGG
GCAAGTTGAAGAATTTCCCAACGGAATTAATGTCTCTAGCCACATTGTTACT
TTAGATATGAGCATGAATCTATTAGAAGGGCATGTTCCTATCTCTCTCTTCCA
AATTCAAAGCCTTGGAGAACCTTGTACTATCCCACAACTCTTTTCCGGCACT
TTCCAAATGAAAAACGTTGGGAGCCCCAATCTTGAAGTTCTTGATTTATCTTA
CAACAACTTGTCAGTAGATGCCAACGTGGATCCAACTTGGCATGGATTTCCC
AAGCTGAGAGAGTTAAGCCTAGCTTCGTGTGACTTGCATGCCTTCCCTGAAT
TCTTGAAACATTCTGCTATGATTAAATTGGACCTTTCAAACAATCGGATTGAT
GGGCAGATACCTAGATGGATTTGGGGAACAGAGCTTTATTTTATGAACCTCT
CTTGTAATCTTCTGACAGATGTGCAAAAGCCTTACCATATCCCAGCTTCTCTT
CAATTACTAGACTTACACTCTAACCGGTTCAAGGGCGACCTGCACCTCTTTA
TTTCTCCCATCGGAGACCTCACGCCCTCTCTTTACTGGTTATCTCTCGCAAA

CAATAGTTTCAGTGGATCGATTCCAACCTCCCTATGCAACGCCACACAACTT
GGAGTTATTGATCTGTCTCTGAATCAATTGAGTGGAGATATAGCCCCTTGTC
TGCTTGAAAACACTGGGCATATCCAAGTGCTTAATCTGGGGCGAAACAACAT
CAGCGGACACATCCCGGACAATTTCCCTTCCCAATGTGGGTTACAGAACTT
GGATCTTAACAACAATGCTATACAAGGGAAAATCCCAAAGTCCCTTGAGAGT
TGCATGTCGTTGGAGATCATGAACGTCGGCGACAACAGCATCGACGATACT
TTCCCATGCATGCTACCGCCGAGCTTGTCCGTCCTTGTTCTGCGCTCCAAC
CGGTTCCACGGAGAGGTTACCTGTGAGAGAAGGGGCACGTGGCCGAATCT
CCAGATCATCGATATATCTTCCAACAATTTCAATGGAAGTCTTGAATCAATAA
ACTTCTCTAGCTGGACAGCAATGGTGCTAATGAGCGATGCACGTTTCACGC
AGCGCCACTGGGGGACTAACTTCCTGTCGGCTTCCCAATTCTACTATACGG
CCGCGGTGGCGCTGACCATCAAAAGGGTGGAGTTGGAGCTCGTCAAGATT
NGGCCGGACTTTATTGCCGTTGATTTGTCCTGCAATGACTTCCATGGAGATA
TACCAGATGCAATAGGCGATCTGACCTCACTCTATGTTCTCAACATATCTCA
CAACGCTCTCAGTGGAAGCATCCCGAAGTCGATGGGTCATTTGTCAAAGCT
CGAATCACTCGATCTCTCGCGAAACCGGCTGTCAGGGCATGTACCGACGGA
GCTCGGAGGTCTGACATTCCTCTCGGTCTTGAACCTGTCGTACAATGAGCT
GGTTGGAGAGATCCCGAATGGGCGTCAGATGCATACATTTTCAGCTGATGC
CTTCAAAGGTAACGCGGGATTATGTGGTCGCCATCTCGAGAGAAATTGCAG
CGATGATCGATCGCAGGGGGAGATTGAGATTGAGAATGAGATTGAGTGGGT
GTACGTTTTTGTTGCATTGGGATATGTTGTGGGCTCAGGAATCATTGTGTGG
CTACTTTTGTTCTGCCGAAGCTTCAGATACAAATACTTCGACAAAATAGATAA
GGTCGTTCAAGAGACATTTGATGCCAGAGACAGGAGAAGAAGAAGACCACG
CGGAACAAGAATAGTAAGGAATCAAGTGGTCAGGAGATCACACTAATGATG
AGGCATGTTATTGTCAAAGATTATTCAAGTTTTGTAGGTTTCTGCGCAACATT
ATTTTA