Chapter 6 Conclusions

Comparative DNA sequencing of *mat*K, a maturase-encoding gene located within the intron of the chloroplast *trn*K gene, was evaluated for phylogenetic utility above the family level and within the grass family (Poaceae). The phylogenetic reconstruction of the Poaceae was conducted using the *mat*K gene sequences representing major grass groups.

6.1 Application of *mat*K above the Family Level

In order to study the potential application of *mat*K to plant systematics above the family level, eleven complete sequences from GenBank representing seed plants and liverworts and nine partial sequences obtained for genera representing the monocot families Poaceae, Joinvilleaceae, Cyperaceae, and Smilacaceae were analyzed.

The study underscored the useful properties of the *mat*K gene for phylogenetic reconstruction: reasonable size (1500 base pairs), high rate of substitution, large proportion of variation at the first and second codon positions, low transition and transversion ratio, and the presence of mutationally conserved sectors. The use of different sectors of the gene and the cumulative inclusion of informative sites showed that the 3' region was the most useful in resolving phylogeny, and that the topology and robustness of the tree reached a plateau after the inclusion of 100 informative sites from that region for the taxa used. The presence of a relatively conserved 3' region and the less conserved 5' region provide two sets of characters that can be used at different taxonomic levels from the tribal to the division levels. The *mat*K gene with its underlying features represents a molecule that has excellent potential for providing insight into evolutionary and systematic problems at various levels. It also has demonstrated the potential of partial sequencing in resolving systematic relationships from the tribe to the division level.

6.2 Preliminary Application of *mat*K to Poaceae

The 3' region of the *mat*K gene from 17 grass species (Poaceae), representing 13 tribes and 6 subfamilies, was used to investigate the potential of the gene in addressing systematic questions in the family. The aligned sequences were analyzed by the Wagner parsimony methods using PAUP and PHYLIP and by the neighbor-joining method.

Out of the 583 bps sequenced, 30% were variable and 14.9% were informative. The strict consensus tree, rooted by *Joinvillea* (Joinvilleaceae), showed well-resolved major clades that represent the grass subfamilies. The bambusoid *Phyllostachys* appeared as a basal clade in the family. *Oryza* appeared either as a sister group to *Phyllostachys* or to the pooid clade. The three members of the Pooideae diverged after the *Oryza* in the parsimony tree or appeared as a sister group to the bambusoid clade in the neighbor-joining tree. The *mat*K-based trees show *Arundo* (Arundinoideae) as a basal taxon to the subfamilies Panicoideae and Chloridoideae. Analysis of the various types of DNA mutations underscores the potential of the *mat*K gene for providing insight into grass systematics and evolution. The relatively high rate of substitution, reasonable size of the coding region, and the low transition/transversion ratios point to the usefulness of the *mat*K gene in grass systematic.

6.3 Characterization of *mat*K in Poaceae

The *mat*K gene in the Poaceae was characterized using the complete sequences from 11 grass genera, representing 7 subfamilies and 11 tribes, and one outgroup (*Joinvillea plicata*) in Joinvillaceae. The alignment of 14 species and 1632 bp yielded a data set of 601 (36.8%) variable sites and 246 (15.1%) informative sites. The variations at nucleic and amino acid levels evenly distribute throughout the entire gene and the 5’ region appears to have more variation than the 3’ region, which is related to its functional domain. The changes at the third codon position are very low as compared to the total (55%) of the first and second positions, which results in a similar variation pattern at nucleic and at amino acid levels. The average transition and transvesion (tr/tv) ratio generated from 14 full length *mat*K sequences is 1.29, which is similar to those in the

previous studies. It was intriguing to find that variation in tr/tv ratios were regionally specific. The ratio tends to become higher toward the 3' end of the *mat*K gene

RASA analysis of the alignment data indicated a relatively high phylogenetic signal in the data set of 14 taxa. In the two-half analyses, while the tRASA of the 5' half of the *mat*K gene is low and is not significant at α=0.05 level, the 3' of the *mat*K gene showed a significant phylogenetic signal. Among the 5 sections of the 14 entire *mat*K sequences, only the fourth sector contains significant phylogenetic signals. These results support the conclusion that the *mat*K gene is phylogenetically valuable, and point to the 3' region of the *mat*K gene, especially the functional fourth sector, to contain strong phylogenetic information. A single most parsimonious tree was obtained from the 246 informative sites of the 14 entire *mat*K sequences. Seven major groups were well resolved in the most parsimonious tree and they correspond to the seven commonly recognized subfamilies: Arundinoideae, Bambusoideae, Centothecoideae, Chloridoideae, Panicoideae, Pooideae and Oryzoideae. The Bambusoid species *Sasa kurilensis* appears at the base of the tree. The two oryzoid species are grouped together and stand out as a separate lineage. Both oryzoid and pooid clades were supported by the bootstrapping value of 100%. The species of PACC group (Panicoideae, Arundinoideae, Centothecoideae, and Chloridoideae) appear to be a well-resolved single lineage supported by the bootstrapping value of 100%.

6.4 Phylogeny of the Grass Family

About 960 base pairs of the *mat*K gene were sequenced from grass species representing 48 genera, 21 tribes, and seven subfamilies to reconstruct a phylogeny for the Poaceae. *Joinvillea plicata* (Joinvilleaceae) was used as an outgroup species. The aligned sequences showed that 495 nucleotides (51%) were variable and 390 (36%) were phylogenetically informative. The Relative Apparent Synapomorphy Analysis (RASA) indicated that a strong significant phylogenetic signal existed in this data set (tRASA=23.4). The cumulative addition of informative sites starting at the internal end of the sequences revealed that at 300 sites, tree topology and bootstrap values matched those

of the consensus tree based on the entire sequence. The Power and Effect test of the RASA showed that after 100 informative sites, the phylogenetic signal reached a plateau and that the signal cannot be improved by adding more characters, i.e., informative sites

The strict consensus tree of the six parsimonious trees showed major lineages supported by 73%-100% bootstrap values. These lineages corresponded to six subfamilies: Bambusoideae, Oryzoideae, Pooideae, Chloridoideae, Panicoideae, and Arundinoideae.

The Bambusoideae, including woody and herbaceous taxa, diverged as a single monophyletic basal lineage. Neither the consensus tree nor the six most parsimonious trees based on the *mat*K gene sequences showed a basal bifurcation in the phylogeny of the Poaceae.

The Oryzoideae genera *Oryza* and *Zizania* formed a monophyletic lineage that included *Ehrharta* as a basal taxon. The lineage was sister to the Bambusoideae. In this study, the oryzoid clade was strongly supported by a 100% bootstrap, 26 apomorphies and a decay index of 4. This molecular information, thus, strongly supports the treatment of the Oryzoideae as a distinct subfamily. In addition, the *Ehrharta*-Oryzoideae clade clearly demonstrates the phylogenetic affinity of *Ehrharta* to the oryzoid grasses

The Chloridoideae, Panicoideae, Arundinoideae, and the centothecoid genus *Zeugitis* (PACC group) emerged as a monophyletic assemblage with a 95% bootstrap support. Within the PACC clade, the Arundinoideae was basal and the Chloridoideae was terminal. The Panicoideae and Arundinoideae were both monophyletic and sister groups.

The monophyly of the Pooideae was supported by a 73% bootstrap value and 18 synapomorphies. *Stipa* emerges as a basal taxon in this subfamily. The remaining eleven pooid genera formed four lineages. The Triticeae clade included *Brachypodium* and was supported by a 98% bootstrap and eight apomorphies. Thus, the *mat*K-based phylogeny supports a subtribal position for *Brachypodium* within the Triticeae (Brachypodiinae).

Bromus appeared as a distinct clade basal to the Triticeae and Poeae lineages. The sample size in this study is too small to address the systematic questions at the tribal level in the Pooideae.

The *mat*K study reflects the phylogenetic affinities of the Centothecoideae and Arundinoideae, and questions the subfamilial status of the Centothecoideae. The Chloridoideae emerged as a monophyletic lineage and a sister group to the Aristideae. The two genera of the Aristedeae (*Aristida* and *Stipagrostis*) branched off before the chloridoid clade as a monophyletic line supported by a 100% bootstrap value and 10 mutations. The *mat*K gene has provided sequence information sufficient for good resolution of the major grass lineages. A more detailed study of some of the subfamilies using the *mat*K gene needs to be conducted and the accumulating data sets should provide more insight into grass systematics and evolution.

HONGPING LIANG

EDUCATION

1993-date: Dept. of Biology, Virginia Polytechnic Institute and State University, Blacksburg, Ph.D. Student, (Plant Systematics)
1984-1987: Dept. of Forestry, Beijing Forestry University, Beijing, China, MS. (Plant Taxonomy)
1978-1982: Dept of Forestry, Sichuan Agricultural University, China, BS. (Forestry)

EXPERIENCE

1987-1993: Lecturer, Dept. of Forestry, Beijing Forestry University, Beijing, China
1982-1984: Assistant Researcher. Sichuan Forestry Institute.

PROFESSIONAL INTERESTS

Plant molecular systematics and evolution, Plant taxonomy

GRANTS

1. 1995: $500 from Sigma Xi, The Scientific Research Society; $500 matched by Dept. of Biology, Virginia Tech
2. 1995: $350 from Graduate Student Assembly, Virginia Tech; $350 matched by Dept. of Biology, Virginia Tech

WORKSHOP AND MEETING

1996: Annual meeting of Virginia Academy of Science

1995: An international workshop on Bilingual (English-Chinese) Botanical Information Management funded by NSF

1994: The annual meeting of American Society of Plant Taxonomists in Knoxville, Tennessee

AWARD

1993: $1000 award from Chinese Education Committee for outstanding young scholar in China (Huo YingDong Award).

PUBLICATIONS

1. Hilu, K. W. and H. Liang. In preparation. Molecular characterization of the *mat*K gene in the Poaceae.
2. Hilu, K. W. and H. Liang. 1997. The *mat*K gene: sequence variation and application in plant systematics. *American Journal of Botany*. 84: 830-839.
3. Hilu, K. W., K. M. M'Ribu, H. Liang, and C. Mandelbaum. 1997. Genetic diversity and evolution of fonio millets. *South African of Botany*. 63:185-190.
4. Liang, H. and Hilu, K. W. 1996. Application of the *mat*K gene sequences to grass systematics. *Canadian Journal Botany*. 74:125-134.

5. Liang, H. 1990. A study on the morphology of foliar trichomes of evergreen oaks (*Quercus L.*) in China and its implication in taxonomy. *Acta Phytotaxonomica Sinica*. 28(2):112-121.
6. Liang, H. 1990. Karyotype Analyses in *Cupressus gigantea*. *Acta Botanica Sinica*. 32(8):653-656.
7. Liang, H. 1990. Studies on plant floristics of Songsan Nature Reserve in Beijing. *In* Essays on the investigation on the plants and animals of Songsan Nature Reserve, edited by Ren Xianwei et al. Chinese Forestry Press.
8. Ren Xianwei, Liang, H., and Xie, W. 1991. The mean temperature method for forecasting phenological phases of trees. *Journal of Beijing Forestry University*. 13(2):87-93.
9. Liang, H. 1991. Karyotype analyses of eight alpine oaks (*Quercus L.*). *Beijing Forestry Technology*. 25:137-145.
10. Liang, H. 1987. The systematic anatomy of twig and leaf and clustering analyses of the alpine oaks in China. Thesis, Beijing Forestry University.
11. Liang, H. 1991. The ornamental value of the wild flowers in Songsan Nature Reserve in Beijing. *Journal of Beijing Landscape and Architecture*. 5:7-14.
12. Liang, H. 1991. Enumeration of the plants in Baihuashan Nature Reserve, Beijing. Chinese Forestry Press.

ABSTRACTS

1. Liang, H. and Hilu, K. W. 1996. Molecular characterization of the *mat*K gene in the grass family (Poaceae). *Virginia Journal of Science.*
2. Liang, H. and Hilu, K. W. 1996. Application of the *mat*K gene sequence to phylogeny of the grass family (Poaceae). *Virginia Journal of Science*. 47:103.
3. Hilu, K. W. and Liang, H. 1996. Phylogenetic construction with the *mat*K gene: walking along the gene. *Virginia Journal of Science*. 47:103.
4. Liang, H. and Hilu, K. W. 1995. Application of the *mat*K gene to the phylogenetic reconstruction of the Poaceae. *American Journal Botany*. 82:146.