Transcriptome analysis of an invasive weed *Mikania micrantha*

Y.-L. HUANG[1]*, X.-T. FANG[1], L. LU[1], Y.-B. YAN[1], S.-F. CHEN[1], L. HU[1], C.-C. ZHU[1], X.-J. GE[2] and S.-H. SHI[1]*

State Key Laboratory of Biocontrol and Guangdong Key Laboratory of Plant Resources, Sun Yat-Sen University, Guangzhou 510275, P.R. China[1]
South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, P.R. China[2]

Abstract

As an initial step towards understanding the molecular mechanisms by which plants become invasive, we present here the first transcriptome analysis for an invasive weed *Mikania micrantha*. The analysis was based on the 75-nucleotide short reads data generated by the *Illumina Genome Analyzer II* system. A total of 31 131 unique sequences were assembled *de novo* based on 8.7 million filtered paired-end sequence reads for the transcriptome of an individual *M. micrantha* growing in the field. 73 % of the unique sequences showed significant similarity to existing proteins in the *NCBI* database, and 21 448 could be grouped based on gene ontology assignments. Of particular interest are the sequences that share homology with genes involved in genome evolution, plasticity, secondary metabolism and defense responses.

Additional key words: gene ontology, *Illumina Genome Analyzer II* system, sequence annotation.

Introduction

Invasive weeds are of primary interest in evolutionary ecology. They not only have important impacts on ecosystems and biodiversity, but also illustrate a fundamental biological process: that is, the establishment and expansion of new populations in a short period of evolutionary time (Levine 2003, Sax *et al.* 2007). Studying a species that recently invaded a new territory allows one to potentially determine if rapid evolutionary change occurred in the invasive populations, and to reveal which particular traits promoted invasiveness. It is not surprising, therefore, that invasive species are being increasingly used as models for the study of rapid evolutionary change (Lee 2002). Despite an increasing number of hypotheses that unite various arms of invasion ecology (Richardson and Pysek 2006), it remains unclear why some plants become problematic invaders and others do not. One aspect that is rarely investigated in relation to invasive weeds and their native counterparts is the potential for modified gene regulation in the introduced range (Broz *et al.* 2007). The development of genetic, especially genomic, resources for weed species would increase our ability to answer fundamental questions about how weeds survive and adapt to natural and human-made stresses, and what evolutionary or environmental changes cause invasiveness.

Mikania micrantha H.B.K. (*Eupatorieae*, *Asteraceae*), a perennial creeping vine commonly called 'mile-a-minute', is one of the 10 worst weeds and one of the 100 worst invasive alien species in the world (Lowe *et al.* 2001). The species is native to Central and South America, where it is a weed of minor importance (Holm *et al.* 1977). The first record of *M. micrantha* in Asia is from the Hong Kong Zoological and Botanical Gardens, dates back to 1884 (Li *et al.* 2000). By the late 1980s and early 1990s, the weed had invaded large areas of the coastal Guangdong Province in South China. Invasive species of recent origin, such as *M. micrantha*, are ideal model systems with which to investigate the early evolutionary mechanisms associated with their ecological success, by comparisons to counterparts in its native range (Ainouche *et al.* 2009).

Although some research has been conducted on *M. micrantha* ecophysiological aspects, genetic variation

Received 14 August 2010, *accepted* 7 February 2011.
Abbreviations: EST - extensive expressed sequence tag; PCR - polymerase chain reaction; GO - Gene ontology.
Acknowledgments: The work was supported by National Natural Science Foundation of China (40876075, 31070290, 30730008 and 40976081), National Basic Research Program of China (2007CB815701, 2009CB119204), the Natural Science Foundation of Guangdong Province (8151027501000089), Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, the Fundamental Research Funds for the Central Universities (09lgpy35), SKLBC09A06 and SKLBC09F06.
* Corresponding authors; fax: (+86) 20 3402 2356, e-mail: lsshyl@mail.sysu.edu.cn; lssssh@mail.sysu.edu.cn

and structure (Zhang *et al.* 2004, Wang *et al.* 2008, Zhang *et al.* 2009, Yan *et al.* 2011), comprehensive analyses of the molecular genetic composition of introduced populations are lacking. Whether the goal is to address the lack of sequence data available for the *M. micrantha* or to elucidate the genetic basis for adaptation, much greater genome-level understanding is required. The need for such studies has been increasingly recognized, particularly in this era of rapid global change (Reusch and Wood 2007, Karrenberg and Widmer 2008). Large-scale sequencing is now possible in non-model organisms, thanks to the development of second-generation sequencing technology, which includes the *Roche GS FLX10*, *Illumina-GAII* and *AB-SOLiD12* systems (Bentley 2006, Schuster 2008). Nevertheless, sequencing complex genomes remains a challenge (Pop and Salzberg 2008), and can be proble-matic during assembly. This applies especially to non-model species such as *M. micrantha* for which genomic information is scarce. As an alternative, sequencing transcriptomes entails less complexity during assembly, while specifically identifying expressed genes. Several recent papers describe the successful application of these second-generation sequencing technologies to deeply sequence transcriptomes (Dassanayake *et al.* 2009, Flicek and Birney 2009, Jackson *et al.* 2009). Compared with other sequencing platforms, *Illumina-GAII* seems to be a more cost-effective option, since it has recently begun supporting 75-nucleotide (nt) paired-end reads with a cost that small laboratories can afford. This technology has now been applied for sequencing several crop species, including cucumber and *Brassica rapa* (Imelfort and Edwards 2009). As an initial step towards understanding the molecular mechanisms by which plants become invasive, we have begun an in-depth analysis of the transcriptomes of *M. micrantha* using the *Illumina-GAII* system. Here, we describe candidate genes that could be used in future experiments to correlate plant gene expression and ecological hypotheses proposed for invasive success. To the best of our knowledge, this is the first published *de novo* assembly of transcriptomes derived from an invasive weed that will be targeted to study invasive behaviour.

Materials and methods

Plants were harvested from the field in the Sun Yat-Sen University Botanic Garden. The source population of this plant comes from Yinhu, Shenzhen, China (N 22°34', E 114°4') where *M. micrantha* was first found in mainland China (Kong *et al.* 2000). The plant was in the form of young seedlings, lacking floral tissue. The entire plant, including roots, was used for RNA extraction. Total RNA was immediately isolated after harvest using the improved CTAB method (Fu *et al.* 2004).

Total RNA integrity was checked by 1.0 % agarose gel electrophoresis and *Bioanalyzer 2100* (*Agilent*, Santa Clara, USA) analysis. Aliquots from the total RNA samples were subjected to library construction according to the protocol offered by *Illumina* (Beijing Genomics Institute, Shenzhen, China). In brief, poly (A) mRNA was isolated and fragmented from the total RNA sample, followed by first- and second-strand cDNA synthesis using random hexamer primers. The double-stranded cDNA fragments were then end-repaired, and a single 'A' base was added to ligate the *Illumina* paired-end adapters. After fragment size selection (in this study, the insert fragment length was ≈140 bp) and a few cycles of amplification through PCR, the final cDNA fragment library was hybridized to a sequencing flow cell. Paired-end short read sequencing was done on one lane of the *Illumina GAII* system.

Image analysis, base-calling, sequence analysis and reads filtering were sequentially done using *Illumina Solexa Pipeline* filters. The first 75 bases of the paired-end read sequences were extracted as raw reads. Reads that either could not pass the *Pipeline* quality filter or contained N bases were excluded from further analysis. Remaining read pairs were regarded as quality control (QC) filtered paired-end reads and subjected to the following assembly.

Three pioneering short-read assembly programs (*Velvet*, *ABySS* and *Edena*) were used for contig assembly and their individual results were further compared and integrated (Hernandez *et al.* 2008, Zerbino and Birney 2008, Simpson *et al.* 2009). *Edena* implements the overlap-layout paradigm that can be used in effective small-size contigs assembly, while Velvet and AbySS use a methodology based on de Bruijn graphs to construct an assembly. Two important parameter effects were investigated during assembly: the trimming effect on assemblies of different read length, and the K-mer effect on assemblies when K-mer ranges from 31 to 41. The analysis process is detailed in the following three steps:

1) Besides the original 75-nt subset (subset 1), two 60-nt data subsets were created separately according to GC content and base quality distribution. In subset 2, the first 15 bases of reads were trimmed off. In subset 3, the last 15 bases of reads were trimmed off.

2) Common approaches to analyze an *Illumina GAII*-generated transcriptome of a virtually unknown genome include mapping the reads to a closely related model genome that has been sequenced, or using existing extensive expressed sequence tag (EST) databases. In the absence of either of these for *M. micrantha*, *de novo* assembly was carried out using three pioneering short-read assembly programs: the *Velvet* assembly program, which is the most widely used program for *de novo* assembly of genomes (http://www.ebi.ac.uk/~zerbino/velvet/); the *ABySS* assembler, which was recently demonstrated to successfully assemble both genomes and transcriptomes using *Illumina-GA* data (http://www.bcgsc.ca/platform/bioinfo/software/abyss); and the *Edena* assembler (http://www.genomic.ch/edena.php), which is based on

the traditional overlap layout paradigm. Each assembly with a distinct K value was first performed to investigate the K-mer effect on assemblies when K-mer ranges from 31 to 41. Subsequently, an optimized k value was used in the assembly with *ABySS* and *Velvet*, and an optimized overlaps length was used in the assembly with *Edena*. We used an optimized k value of 39 in the assembly with *ABySS* and *Velvet*. *Edena* was optimally parameterized to consider overlaps displaying a minimum length of 38 nt for the strict mode.

3) All assemblies from the three subsets using different programs with corresponding optimum K-mers were stored for the further analysis. *CAP3* software (Huang and Madan 1999) was adopted to reassemble the assemblies from different sources, to remove redundancy and generate consensus contigs.

The additional redundancy removal step was performed to address the concerns that some of non over-lapping unique contigs might effectively correspond to different regions of the same expressed mRNA. In this step, contigs assembled from experimental data were further screened for redundancy using a *BLASTX* search against a download *NCBI* protein database consisting of three model plant: *Vitis vinifera*, *Arabidopsis thaliana* and *Populus trichocarpa* (*p*-value cutoff: 1.0E-10). Different contigs aligning to the same *Refseq* on the condition (≥ 40 bp overlap and at least 90 % identity) are considered to be redundant (cutoff values are same with those adopted by CAP3). All redundant contigs identified in this step were removed before the annotation process.

Computational annotation of transcriptome data was performed using *Blast2GO* software *v2.4.5* (http://www.blast2go.org, Conesa *et al.* 2005, Aparicio *et al.* 2006) as described in Galla *et al.* (2009) with minor modifications. Briefly, a sequence length threshold of 200 base pairs was used. The *BLAST* expectation value (E-value) threshold was constantly set to 1.0E-3. Similarly, the *Blast2GO* software *v2.4.5* was used to obtain gene ontology (GO) information from retrieved database matches. Annotation of all sequences was performed using the software default parameters. *InterPro Scan* (Berardini *et al.* 2004) was performed to find functional motifs and related GO terms using the specific tool implemented in the *Blast2GO* software with the default parameters. Lastly, the 'Augment Annotation by ANNEX' function was used to refine annotations (http://www.goat.no, Zdobnov *et al.* 2001). The *GOslim* 'goslim_plant.obo' was used to achieve specific GO terms by means of a plant-specific reduced version of gene ontology (http://geneontology.org).

Results

One lane of one *Illumina GAII* run resulted in 15.5 million paired-end sequence reads (75 nt each read) for the transcriptome of *M. micrantha*. After removing low quality sequence reads through *Illumina Solexa Pipeline* filters, 8.7 million paired-end sequence reads were retained as qualified reads. To investigate the effect of error rate and GC content on the assembly, three subsets were created and assembled (Table 1). Longer reads often resulted in a higher number of contigs and longer contig length during assembly. The subsequent *BLAST* result also indicated higher *BLAST* or annotation rate when using untrimmed reads compared with trimmed reads (result not shown). In our study, the quality score of all filtered reads exceeded 26, which means the error rate was below 0.23 % in each site. Therefore, untrimmed sequence reads were selected for our analysis.

All programs generated comparable contigs. However, for a given subset where *Velvet* often generated a higher contig number, *ABySS* generated longer contigs and higher N50. To obtain representative transcripts as much as possible, we pooled the assembly results from all three assemblers. Pooled data were further reassembled using *CAP3* to remove redundancy. Following assembly, the average contig size was 354 bp (Table 1), which is

Table 1. Comparison of transcriptome assemblies of *Mikania micrantha* obtained by *ABySS*, *Edena*, *Velvet*, respectively and reassemblies by *CAP3* (The contigs in each assemblies are retained when the length ≥ 200 bp)

Data subset	Programs	Number of contigs	Mean	Max	N50	Sum
75 nt	*ABySS*	10047	421	2756	417	423E+06
	Edena	13658	335	1677	340	458E+06
	Velvet	32218	334	2128	335	1.08E+06
	CAP3	32135	354	3048	358	1.14E+06
First 60 nt	*ABySS*	9317	314	2008	304	2.93E+06
	Edena	9859	303	1384	299	2.99E+06
	Velvet	21701	306	1986	300	6.64E+06
	CAP3	22057	320	2310	316	7.06E+06
Last 60 nt	*ABySS*	11659	343	1824	338	4.00E+06
	Edena	11692	329	2279	328	3.85E+06
	Velvet	23097	327	2130	324	7.55E+06
	CAP3	23057	345	2317	345	7.95E+06

sufficient to effectively assign functional annotations. The additional redundancy removal step identified 1 004 (*ca*. 3.1 % of the total) non overlapping unique contigs which correspond to different regions of the same mRNA. A total of 31 131 unique sequences (contigs) after redundancy removal were used for annotations and further analysis.

These 31 131 unique sequences were used as *BLASTX* queries to search for structural homologies and significant similarities. A total of 22 704 sequences (72.9 %) revealed significant similarity with deposited records, showing an average nucleotide similarity estimate of 84.3 % and a median E-value of 1.0 E-6.

GO term mapping allowed the identification of 107 869 GO terms based upon 21 448 matches retrieved by *BLASTX*. The computational analysis of all 31 131 unique sequences using the software *Blast2GO* allowed the annotation of the expressed sequences according to the terms of the three main gene ontology vocabularies, *i.e*. cellular compartment, molecular function, and biological process. Annotations were simplified using a plant specific *Goslim*. The distribution of annotations among GO vocabularies and the number of unique sequences with GO terms belonging to only one, a combination of two and all three vocabularies was organized as a Venn diagram (Fig. 1).

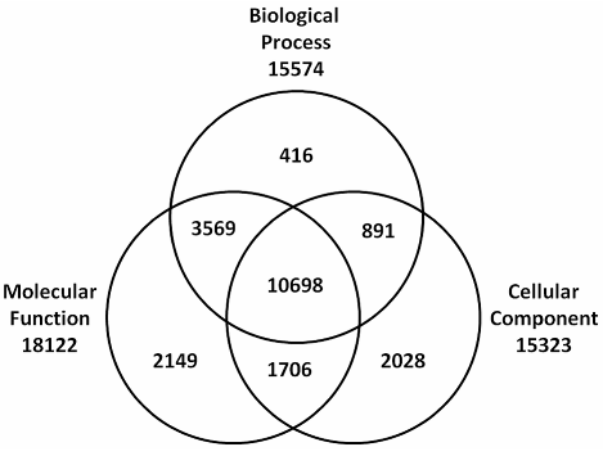

Fig. 1. Venn diagram showing the number of annotations in each gene ontology (GO) category for *Mikania micrantha* and the overlap of their representation. The numbers outside the diagram report the total number of annotations in each GO category.

As far as molecular functions are concerned, 42.1 % (6 219) of the unique sequence annotations fell into the 'catalytic activity' category (Table 2). Specific catalytic activities associated with these sequences covered a range of GOs, with the largest amount of unique sequences falling into the ‘transferase’ and ‘hydrolase’ categories. The next largest category under molecular function was 'binding' (41.2 %, 6 076 unique sequences), with the majority of unique sequences (2 315) being associated with protein binding. Transporters accounted for 6.3 % of all unique sequence annotations. However, most unique sequences could not be subcategorized with a specific transport role.

A total of 578 unique sequence annotations were designated as transcriptional regulators, with the majority (428) being subcategorized as having 'transcription factor activity' (Table 1). Transcription factors are responsible for modulating cellular responses to biotic and abiotic stimuli, and they may play important roles in plant invasion by up- or down-regulating the expression of genes involved in defense and growth responses. Transcription factors identified in the transcriptome of *M. micrantha* made up 3.9 % of the GO annotations for molecular function, whereas ≈ 4 % of the *Arabidopsis* genome sequences are annotated as transcription factors (the *Arabidopsis* information resource).

Table 2. Gene ontology annotation of the unique sequences of *Mikania micrantha* transcriptome (level 2).

		Unique sequences	[%]
Molecular function	catalytic activity	6219	42.1
	binding	6076	41.2
	transporter activity	930	6.3
	transcription regulator activity	578	3.9
	molecular transducer activity	373	2.5
	structural molecular activity	320	2.2
	translation regulator activity	148	1.0
	enzyme regulator activity	122	0.8
Biological process	cellular process	5969	30.2
	metabolic process	5685	28.7
	response to stimulus	1866	9.4
	localization	1241	6.3
	developmental process	1117	5.6
	multicellular process	1093	5.5
	biological regulation	1046	5.3
	cellular organization	744	3.8
	reproduction	623	3.2
	growth	237	1.2
	death	114	0.6
	multi-organism process	63	0.3
Cellular component	cell	7725	54.8
	organelle	5468	38.8
	membrane-enclosed lumen	274	1.9
	extracellular region	323	2.3
	macromolecular complex	304	2.2

Concerning the biological process, 30.2 % of the unique sequence annotations were grouped into the 'cellular process' category (Table 2), which includes subcategories such as cell communication, cell cycle, and differentiation. The next largest category of unique sequence annotations (28.7 %), 'metabolic process', has some overlap of subcategories with cellular process (*i.e*. cellular biosynthetic process and photosynthesis), but includes unique subcategories such as catabolic process and secondary metabolic process. Approximately 9.4 % of unique sequence annotations fell in the 'response to stimulus' category, which includes subcategories that relate to plant responses of abiotic and biotic stimulus. An additional 1 317 unique sequences were subcategorized as

responding to stress.

Concerning cellular compartments, the most represented categories were those of cell (54.8 %), followed by proteins localized in organelles (38.8 %). The other cellular compartments were represented at a lower level (Table 2).

Discussion

Understanding the genetic basis of traits involved in rapid adaptation to novel environmental conditions is a major goal in invasion biology (Prentis *et al.* 2008). However, limited genetic resources are available for most invaders and other weedy species, including *M. micrantha,* which has limited available genetic data. Therefore, our main goals in this study were to assemble the representation and annotation of the transcriptome of *M. micrantha* as much as possible, and to examine its characteristics by investigating unigene candidates for testing ecological hypotheses. Although molecular markers have been used successfully to study plant species (including invasive species) in detecting temporal and spatial patterns (Ma *et al.* 2010, Sikdar *et al.* 2010, Viana *et al.* 2010), DNA sequence have much more potential to test the functional hypotheses (Reusch and Wood 2007, Karrenberg and Widmer 2008). The recent development of next generation sequencing technologies has significantly improved sequencing throughput at a markedly reduced cost (Shendure and Ji 2008). However, the application of these technologies to population studies is not straightforward. In mixed samples from multiple individuals, the error rate may sometimes be higher than true polymorphisms of biological interest (Dohm *et al.* 2008, Shendure and Ji 2008). For this reason, rather than acquiring the most complete representation of its transcriptome by pooling RNA samples from different developmental stages, tissues types or microhabitats, we focused on the transcriptome profile of an individual seedling growing in the field. The success of this approach is indicated by the fact that we were able to generate *de novo* a transcriptome to a depth of 31 131 unique sequences, and 21 448 unique sequences with GO functional interpretations. We have also shown that next-generation DNA sequencing technologies can be used effectively for de novo sequencing of plant transcriptomes, making it possible to carry out rapid and low-cost sequencing for other important plant species.

The GO annotation in this study indicated that a wide variety of functional categories were represented in the transcriptome of *M. micrantha*. Evolution of invasive species through random mutation, movement of transposable elements, and genetic recombination may facilitate changes in plant genes or gene expression that give them a competitive and evolutionary advantage. On some occasions, invasive populations possess advantageous phenotypes due to adaptive or developmental plasticity under the novel environments. Candidate unique sequences from the transcriptome of *M. micrantha* potentially involved in genome evolution and plasticity were annotated as mobile elements, heat-shock proteins, *etc* (data not shown). In addition, members of the aster family can synthesize a broad spectrum of secondary metabolites that may aid in basal and induced defense responses, as well as in competition against other plants. Candidate unique sequences potentially involved in defense responses or secondary metabolism are also identified in this transcriptome (data not shown). Many of these unique sequences could be used as candidates to look at more specific changes in regulation related to invasion.

The transcriptome sequencing of *M. micrantha* represents an initial step towards looking at gene-specific expression in this species. Other resources, such as microarray chips can provide a view of global gene expression in invasive *M. micrantha* and its native counterparts. By comparing the transcriptomes of populations from the native and introduced ranges of invasive species, we will be able to determine both the severity of bottlenecks at population founding, and potential candidate loci that have been the target of selection.

References

Ainouche, M.L., Fortune, P.M., Salmon, A., Parisod, C., Grandbastien, M.A., Fukunaga, K., Ricou, M. Misset, M.T.: Hybridization, polyploidy and invasion: lessons from *Spartina* (*Poaceae*). - Biol. Invasions **11**: 1159-1173, 2009.

Aparicio, G., Götz, S., Conesa, A., Segrelles, D., Blanquer, I., García, J.M., Hernandez, V., Robles, M., Talon, M.: *Blast2GO* goes grid: developing a grid-enabled prototype for functional genomics analysis. - Stud. Health Technol. Inform. **120**: 194-204, 2006.

Bentley, D.R.: Whole-genome re-sequencing. - Curr. Opin. Genet. Dev. **16**: 45-552, 2006.

Berardini, T.Z., Mundodi, S., Reiser, L., Huala, E., Hernandez, M.G., Zhang, P., Mueller, L.A., Yoon, J., Doyle, A., Lander, G., Moseyko, N., Yoo, D., Xu, I., Zoeckler, B., Montoya, M., Miller, N., Weems, D., Rhee, S.Y.: Functional annotation of the *Arabidopsis* genome using controlled vocabularies. - Plant. Physiol. **135**: 745-755, 2004.

Broz, A.K., Broeckling, C.D., He, J.B., Dai, X.B., Zhao, P.X., Vivanco, J.M.: A first step in understanding an invasive weed through its genes: an EST analysis of invasive *Centaurea maculosa*. - BMC Plant. Biol. **7**: 25, 2007.

Conesa, A., Götz, S., Garcìa-Gòmez, J.M., Terol, J., Talón, M., Robles, M.: *Blast2GO*: a universal tool for annotation, visualization and analysis in functional genomics research. - Bioinformatics **21**: 3674-3676, 2005.

Dassanayake, M., Haas, J.S., Bohnert, H.J., Cheeseman, J.M.: Shedding light on an extremophile lifestyle through transcriptomics. - New Phytol. **183**: 74-75, 2009.

Dohm, J.C., Lottaz, C., Borodina, T., Himmelbauer, H.: Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. - Nucleic Acids Res. **36**: e105, 2008.

Flicek, P., Birney, E.: Sense from sequence reads: methods for alignment and assembly. - Natur. Methods **6**: s6-s12, 2009.

Fu, X., Deng, S., Su, G., Zeng, Q., Shi, S.: Isolating high-quality RNA from mangroves without liquid nitrogen. - Plant. mol. Biol. Rep. **22**: 1-5, 2004.

Galla, G., Barcaccia, G., Ramina, A., Collani, S., Alagna, F., Baldoni, L., Cultrera, N.G., Martinelli, F., Sebastiani, L., Tonutti, P.: Computational annotation of genes differentially expressed along olive fruit development. - BMC Plant. Biol. **9**: 128-145, 2009.

Hernandez, D., Francois, P., Farinelli, L., Østerås, M., Schrenzel, J.: *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. - Genome Res. **18**: 802-809, 2008.

Holm, L.G., Plucknett, D.L., Pancho, J.V., Herberger, J.P.: The World's Worst Weeds: Distribution and Biology. - East-West Center/University Press of Hawaii, Honolulu 1977.

Huang, X., Madan, A.: CAP3: A DNA sequence assembly program. - Genome Res. **9**: 868-877, 1999.

Imelfort, M., Edwards, D.: *De novo* sequencing of plant genomes using second-generation technologies. - Brief. Bioinform. **10**: 609-618, 2009.

Jackson, B.G., Schnable, P.S., Aluru, S.: Parallel short sequence assemble of transcriptomes. - BMC Bioinform. **10**: S14, 2009.

Karrenberg, S., Widmer, A.: Ecologically relevant genetic variation from a non-*Arabidopsis* perspective. - Curr. Opin. Plant Biol. **11**: 156-162, 2008.

Kong, G.H., Wu, Q.G., Hu, Q.M.: Appearing of exotic weed *Mikania micrantha* H.B.K. in China. - J. trop. subtrop. Bot. **8**: 27, 2000.

Lee, C.E.: Evolutionary genetics of invasive species. - Trends Ecol Evol **17**: 386-391, 2002.

Levine, D.A.: Ecological speciation: lessons from invasive species. - Syst. Bot. **28**: 643-650, 2003.

Li, M.G., Zhang, W.Y., Liao, W.B.: The history and status of the study on *Mikania micrantha*. - Ecol. Sci. **19**: 41-45, 2000.

Lowe, S., Browne, M., Boudjelas, S.: 100 of the World's Worst Invasive Alien Species, a Selection from the Global Invasive Species Database. IUCN/SSC Invasive Species Specialist Group (ISSG), Auckland 2001.

Ma, K.H., Kim K.H., Dixit, A., Chung, I.M., Gwag, J.G., Kim, T.S., PARK, Y.J.: Assessment of genetic diversity and relationships among *Coix lacryma-jobi* accessions using microsatellite markers. - Biol. Plant. **54**: 272-278, 2010.

Pop, M., Salzberg, S.L.: Bioinformatics challenges of new sequencing technology. - Trends Genet. **24**: 142-149, 2008.

Prentis, P.J., Wilson, J.R., Dormontt, E.E., Richardson, D.M., Lowe, A.J.: Adaptive evolution in invasive species. - Trends Plant Sci. **13**: 288-294, 2008.

Reusch, T., Wood, T.: Molecular ecology of global change. - Mol. Ecol. **16**: 3973-3992, 2007.

Richardson, D.M., Pyšek, P.: Plant invasions: merging the concepts of species invasiveness and community invasibility. - Prog. Phys. Geogr. **30**: 409-431, 2006.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I., Rice, William, R.: Ecological and evolutionary insights from species invasions. - Trends Ecol. Evol. **22**: 465-471, 2007.

Schuster, S.C.: Next-generation sequencing transforms today's biology. - Natur. Methods **5**: 6-18, 2008.

Shendure, J., Ji, H.: Next-generation DNA sequencing. - Natur. Biotechnol. **26**: 1135-1145, 2008.

Sikdar, B., Bhattacharya, M., Mukherjee, A., Banerjee, A., Ghosh, E., Ghosh, B., Roy, S.C.: Genetic diversity in important members of *Cucurbitaceae* using isozyme, RAPD and ISSR markers. - Biol. Plant. **54**: 135-140, 2010.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J., Birol, I.: *AbySS*: a parallel assembler for short read sequence data. - Genome Res. **19**: 1117-1123, 2009.

Viana, A.K.C., Souza, M.M., Araújo, I.S., Corrêa, R.X., Ahnert, D.: Genetic diversity in *Passiflora* species determined by morphological and molecular characteristics. - Biol. Plant. **54**: 535-538, 2010.

Wang, T., Su, Y., Chen, G.: Population genetic variation and structure of the invasive weed *Mikania micrantha* in Southern China: Consequences of rapid range expansion. - J. Hered. **99**: 22-33, 2008.

Yan, Y.B., Huang, Y.L., Fang, X.T., Lu, L.,Zhou, R.C., Ge, X.J., Shi, S.H.: Development and characterization of EST-SSRs in an invasive weed *Mikania micrantha* (Asteraceae). - Amer. J. Bot. **98** (Suppl.): e1-e3, 2011.

Zdobnov, E.M., Apweiler, R.: *InterProScan* - an integration platform for the signature-recognition methods in *InterPro*. - Bioinformatics **17**: 847-884, 2001.

Zerbino, D.R., Birney, E.: *Velvet*: algorithms for *de novo* short read assembly using de Bruijn graphs. - Genome Res. **18**: 821-829, 2008.

Zhang, L.L., Wen, D.Z., Fu, S.L.: Responses of photosynthetic parameters of *Mikania micrantha* and *Chromolaena odorata* to contrasting irradiance and soil moisture. - Biol. Plant. **53**: 517-522, 2009.

Zhang, L.Y., Ye, W.H., Cao, H.L., Feng, H.L: *Mikania micrantha* H. B. K. in China - an overview. - Weed Res. **44**: 42-49, 2004.