Chapter 5.5

Mapping and Map-Based Cloning

Niels Sandal1,*, Katja Krause2, Mirela Frantescu1, Simona Radutoiu1, Lene Krusell2, Lene H Madsen1, Norihito Kanamori1, Esben Madsen1, Leïla Tirichine1, and Jens Stougaard1

1Laboratory of Gene Expression, Department of Molecular Biology, University of Aarhus, Gustav Wieds Vej 10, DK-8000 Aarhus C, DENMARK; 2Molecular Plant Nutrition Group, Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, D-14476 Golm, GERMANY; *Corresponding author

Email: nns@mb.au.dk
Phone: +45 8942 5006 Fax: +45 8612 3178

Keywords: Lotus japonicus, Lotus filicaulis, map-based cloning, mapping, colormap, marker, sym gene, symbiosis

Development of a genetic system and genetic resources is one of the prerequisites for establishing Lotus japonicus as a model legume. Successful identification and characterisation of several key plant genes controlling symbiosis with nitrogen fixing bacteria and mycorrhizal fungi demonstrate the potential of this approach. We provide here an overview of efforts to establish a genetic map and approaches using genetic marker information for positional cloning. In L japonicus, six independent mutant populations in the ecotype Gifu were obtained with EMS, fast neutrons, T-DNA, transposon, or tissue culture. More than 500 independent symbiotic mutants have been isolated and at least 24 independent loci found. General mapping populations are available from crosses between ecotype Gifu and ecotype MG-20 and between Gifu and L filicaulis. These parents are also used for establishing populations for map-based cloning. Five BAC and TAC libraries were constructed. The general strategy for map-based cloning in L japonicus is described and map positions of eleven sym genes are shown. Four L japonicus sym genes have been isolated using map-based cloning and two genes have been isolated with transposon tagging.

Introduction

Part of the rationale for selecting L japonicus as a model legume (Handberg and Stougaard, 1992) was to develop a genetic system enabling characterisation of plant genes required for symbiosis with nitrogen fixing bacteria and mycorrhizal fungi. We have chosen the genetics of symbiosis for illustrating approaches and methodology used in mapping and map based cloning. In order to isolate symbiotic
mutants and genes from a homozygous genetic background the *L. japonicus* ecotype Gifu was inbred for 9 generations to generate the B-129-S9 germplams (Stougard and Beuselinck, 1996).

Six independent mutant populations have been obtained using chemical (EMS) mutagenesis (Perry *et al.*, 2003; Kawaguchi *et al.*, 2002; Szczyglowski *et al.*, 1998; Webb *et al.*, unpublished, Gresshoff *et al.*, unpublished), four using T-DNA or transposon insertion mutagenesis (Thykjær *et al.* 1995, Schauser *et al.* 1998; Webb *et al.*, 2000; Gresshoff *et al.*, unpublished), one using fast neutrons (Gresshoff *et al.*, unpublished), and several using tissue culture-derived mutants (Umehara and Kouchi, unpublished). Slightly different conditions were used for screening of mutant populations but in general the following phenotypes were scored: Absence of nodules (Nod−), nodules not capable of nitrogen fixation (Fix−), decreased or delayed nodulation (Nod+−), increased or hypernodulation (Nod++), Nodulation in the absence of *Rhizobium* (Nar) and absence of mycorrhization.

A summary of the type and number of symbiotic mutant lines so far identified within the *Lotus* community is presented in Table 1. In addition, other developmental or morphological mutants were found. Among these were dwarfs, leaf and root mutants, and flowering mutants, for example *pfo* characterised by excessive development of sepals and indeterminate growth of floral meristems (Zhang *et al.*, 2003).

<table>
<thead>
<tr>
<th>Mutant phenotype</th>
<th>Number of mutant lines</th>
<th>Mycorrhization deficient</th>
<th>Number of loci</th>
<th>Alleles per locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonnodulation (Nod−)</td>
<td>296</td>
<td>31</td>
<td>&gt;10</td>
<td>1-12</td>
</tr>
<tr>
<td>Ineffective nodulation (Fix−)</td>
<td>252</td>
<td>1</td>
<td>&gt;11</td>
<td>1-3</td>
</tr>
<tr>
<td>Hypernodulation (Nod++−)</td>
<td>14</td>
<td>3</td>
<td>3</td>
<td>1-6</td>
</tr>
<tr>
<td>Reduced or delayed nodulation (Nod+−)</td>
<td>61</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

Table 1. Summary of *L. japonicus* symbiotic mutants and alleles. Only few of the Nod− and Fix− mutants have been tested for mycorrhizal symbiosis and most of the mutants have not yet been assigned to complementation group.

Diallelic crosses between all mutants with a similar phenotype to screen F1 plants for genetic complementation is usually next step following mutant isolation. Such an assignment to complementation groups identifies all alleles that contribute to phenotypic characterisation of mutants and genotyping of loci. Equally important it estimates the number of loci involved in a particular process and makes it possible to optimize use of resources required for map-based cloning of genes. Ideally, duplication of efforts between research groups working on *L. japonicus* should be avoided and assignment of all available alleles to loci provides a common knowledge basis for project planning. However, diallelic crossing is a relatively
slow process where progress is determined by generation time and slowed by continuously increasing number of individual crosses necessary to keep up with mutant isolation programs. Given the number of symbiotic mutant lines already available, this approach is unlikely to encompass all alleles in near future. Fortunately, genetic mapping has become easier and it is now possible to quickly define a rough map position for an allele using one or two microsatellite markers from each of the six linkage groups. Only one cross per mutant line is necessary to map a new allele relative to already known loci and together with TILLING of already cloned genes this could decrease the number of diallelic crosses needed to clean up mutant material.

At the end of a map-based cloning project, characterisation of several alleles provides very strong evidence that the correct gene has been identified at the locus. Depending on the strength of such evidence, this alone may serve to prove the identity of the gene underlying the phenotype or trait.

MAPPING AND CLONING IN LOTUS

Map-based cloning

The minimum requirements for accomplishing positional cloning successfully are:

- A mutation with good penetrance ensuring a mutant phenotype that is easily and reliably scored
- A cross to a polymorphic mapping parent producing sufficient seeds to establish a population of F2 plants segregating the mutation, a general genetic map with molecular markers
- A BAC (bacterial artificial chromosome)/TAC (transformation-competent artificial chromosome) library with large inserts
- Access to DNA sequence in the region delimited by recombination end-points
- Several mutant alleles and/or methods allowing complementation of mutants after transformation with the wild type gene.

Mapping partners

At present, two mapping partners, *Lotus filicaulis* (Sandal et al., 2002) and *L. japonicus* ecotype Miyakojima (MG-20) (Kawaguchi et al., 2001), are available for establishing an F2 population segregating mutations obtained in a Gifu background. Each partner has advantages and disadvantages.

*Lotus filicaulis*

+ High level of polymorphism; using a number of AFLP primer combinations, 40-50% of the bands were polymorphic
+ Easy to map a gene
+ A genetic map has been developed (Sandal et al., 2002).

− Low seed yield on F1 plant

− Relatively long generation time

− Several regions with distorted segregation especially middle part of chromosome 1

− Suppression of recombination on upper part of chromosome 3 and middle of long arm of chromosome 1

**Lotus japonicus, MG-20**

+ High seed yield on F1 plant

+ Normal generation time, increasing number of microsatellite markers are developed from sequenced TACs and BACs

+ A genetic map has been developed (Hayashi et al., 2002).

+ This ecotype is getting better and better as a mapping parent due to the genome sequencing of this ecotype at Kazusa DNA Research Institute. Upon completion of the complete MG-20 genome, SNP (single nucleotide polymorphisms) markers for Gifu/MG-20 will be developed on a genome-wide scale.

− Relatively low level of polymorphism; around 6% of AFLP bands polymorphic

− Compared to _L. filicaulis_, more difficult to find linked markers, difficult to map a gene unless it is on a sequenced and mapped TAC clone

− Suppression of recombination on upper part of chromosome 1.

**Crossing**

To cross _L. japonicus_ ecotypes it is important to use female flowers that are at just the right stage of development. The pollen is then collected from a mature flower. Pictures of flowers where the pollen has not yet been liberated and the female parts are ready for fertilisation (Jiang and Gresshoff, 1997) are shown in Fig x. For crosses to _L. filicaulis_ (Grant et al., 1962), older flowers can be used since _L. filicaulis_ is less self fertile than _L. japonicus_, at least in our greenhouse. The F1 plants can be discriminated from both parents based on the morphological characters. The F1 plant has broad leaves like the _L. japonicus_ parent and some of the flower characters are similar to _L. filicaulis_. Alternatively, the F1 plant DNA can be analysed using a PCR marker giving a size difference to confirm cross-fertilisation.

**Mapping**

Efficient mapping relies on high-density genetic maps developed using molecular markers. Evolutionary distant relatives have higher levels of polymorphisms than found between closer relatives and provide a better source of DNA differences for
molecular markers. However there are also disadvantages. Seed yield on F1 plants can be low and there will be regions with distorted segregation and often more inversions and translocations reflecting the evolutionary distance. For *L japonicus* two different maps balancing advantages and disadvantages were established, (Sandal *et al.* 2002, Hayashi *et al.* 2001). One was based on an interspecific cross between *L japonicus* and *L filicaulis*, the other on a cross between the *L japonicus* ecotypes Gifu and MG-20. The nomenclature of the linkage groups was agreed so that linkage group I corresponds to the largest *L japonicus* Gifu chromosome (chromosome 1) and the maps are shown with the short arm on the top (Sandal *et al.*, 2002, Hayashi *et al.*, 2002).

The alignment of the two maps and FISH analysis showed that there is a translocation between the top of chromosome 1 and the bottom of chromosome 2 between Gifu and MG-20 (Hayashi *et al.*, 2002). Furthermore, a large inversion of the short arm of chromosome 3 and a small inversion on the long arm of chromosome 1 were found between Gifu and *L filicaulis* (Pedrosa *et al.*, 2002). These chromosome rearrangements result in problems with genetic mapping in corresponding regions of relevant mapping populations. Alignment of two maps is very helpful in resolving these regional difficulties and in ordering markers. Additional small chromosome rearrangements can not be excluded and it may be important to do physical mapping with single copy DNA from the different chromosomes (and chromosome arms) as previously done by fluorescent in situ hybridisation (FISH) (Pedrosa *et al.*, 2002).

For the Gifu x *L filicaulis*, an F2 mapping population of 165 plants was established and using only 3 selected AFLP primer combinations it was possible to cover much of the *L japonicus* genome for fast mapping (Sandal *et al.*, 2002). The Gifu x MG-20 F2 population consists of 127 plants and the map established from this cross is currently used to support the genome sequencing at the Kazusa. Transformable artificial chromosomes (TACs) corresponding to markers and ESTs are selected for seed point sequencing. After sequencing a TAC clone (app. 100 kb), the sequence is scanned for the presence of microsatellite sequences like, for example, ATATATATAT or GTGTGTGT. This type of sequence changes rapidly during evolution, resulting in small insertion/deletions.

In many cases a size difference between the mapping parents is found that can be used as a microsatellite marker anchoring a sequenced TAC clone on the genetic map of Lotus (Sato *et al.*, 2001, Hayashi *et al.*, 2001, Nakamura *et al.*, 2002, Kaneko *et al.*, 2003, Asamizu *et al.*, 2003). For some of the SNPs found, it was possible to develop codominant derived cleaved amplified polymorphic sequences (dCAPS) (Neff *et al.*, 1998). For each TAC clone, around 10 genes have been sequenced and placed on the genetic map through microsatellite/dCAPS markers. Now, more than 500 TAC clones have been sequenced anchoring more than 5000 genes on the genetic map.

The microsatellite markers developed for the *L japonicus* Gifu x *L japonicus* MG-20 mapping population have also been tested for polymorphism between *L japonicus* Gifu and *L filicaulis* and 83% of the markers can be used in this cross either as dominant or codominant markers. For the remaining 17% of markers sequencing of the PCR products reveal SNPs for many of them. This efficient marker transfer
provides an opportunity for direct and fast alignment of the two genetic maps. After scoring genotypes of individual loci in mapping populations, genetic linkage maps were produced using Mapmaker and Joinmap programs. Usually a manual colormap (Kiss et al., 1998) is also made to provide a visual overview of the linkage groups and to indicate quality of the map and markers. On average, only one recombination is seen per plant per chromosome in Lotus F2 populations. Apparent, close double recombination events in any one F2 plant will be obvious from the colormap. This might be caused by mistakes in the experiment or in the scoring and should therefore be checked again.

BAC libraries

For physical mapping and conversion of genetic linkage into physical distance, large insert libraries are needed. A number of L japonicus BAC libraries are available (Table 2). At Kazusa a number of L japonicus MG-20 BAC and TAC libraries were made from partial MboI or HindIII digestions of genomic DNA to support genome sequencing (Sato et al., 2001). More recent libraries from Kazusa were constructed with EcoRI or Sau3A. A BAC library of L japonicus Gifu carrying very large inserts was made by Kawasaki and Murakami, (2000). A transformation-competent Gifu BAC library was constructed by Men et al., (2001) and we have constructed a L japonicus Gifu BAC library from a partial BamHI digests (Sandal et al., unpublished). This library consists of 30,000 clones with an average insert size of 75 kb. On average two copies of each single copy gene is found in this library. In general, it is a great advantage to have BAC and TAC libraries constructed with different restriction enzymes increasing chances of covering the whole genome. For fine mapping sequencing directly on BACs or TACs, is very important and access to BAC or TAC libraries for both mapping parents is an advantage.

<table>
<thead>
<tr>
<th>Ecotype</th>
<th>Restriction enzyme</th>
<th>Vector</th>
<th>Ave. insert size</th>
<th>Number of clones</th>
<th>Screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gifu (a)</td>
<td>HindIII</td>
<td>PCLD04541</td>
<td>94 kb</td>
<td>33408</td>
<td>Hybridisation</td>
</tr>
<tr>
<td>Gifu (b)</td>
<td>BamHI</td>
<td>pBAC/SacBI</td>
<td>75 kb</td>
<td>32000</td>
<td>Hybridisation + PCR</td>
</tr>
<tr>
<td>Gifu (c)</td>
<td></td>
<td></td>
<td>140 kb</td>
<td>30720</td>
<td>Hybridisation</td>
</tr>
<tr>
<td>MG-20 (d)</td>
<td>HindIII/MboI</td>
<td>pYLTAC7</td>
<td>87, 96, 105, 106 kb</td>
<td>7.7 genomes</td>
<td>PCR</td>
</tr>
</tbody>
</table>

Table 2. L japonicus BAC libraries constructed in different laboratories: (a) Men et al., 2001; (b) Sandal et al., unpublished; (c) Kawasaki and Murakami, 2000; (d) Sato et al., 2001.
Map-based cloning

Most of the *L. japonicus* Gifu mutants originate from chemical mutagenesis programs so map-based cloning is needed to isolate the gene involved. To initiate a positional cloning project a mutant is crossed to MG-20 and/or *L. filicaulis*. MG-20 appears to be more efficient as pollen recipient than the Gifu parent. Alternatively, since most mutations are recessive, the mutant can be chosen as pollen recipient and the F1 plant tested for wild type phenotype. In the next generation, F2 plants segregating mutants are grown under conditions where mutant phenotype can be distinguished from wild type. For recessive mutations, mutants are selected for genotyping because they will be homozygous for the mutant parent genotype (DNA polymorphisms) in the chromosomal region carrying the locus and gene of interest. For dominant mutations, plants with wild type phenotype are genotyped as they are homozygous for the recessive allele from the wt parent. With a mapping population of only 20-30 mutants, it is possible to place the locus on a chromosome and within a region of the linkage map. See Table 3 showing a colormap for 20 mutant plants from a cross of the mutant *har1* to MG-20. Using this approach, a number of symbiotic loci/genees have now a rough position and some were subsequently mapped and cloned (Figure 1).

![Figure 1](image-url)
In practice, microsatellite markers are highly useful for this crude mapping. PCR primers are placed flanking the microsatellite sequence and it is possible to detect the PCR fragment size difference between the mapping parents and both fragments in heterozygotes. To resolve small differences, a PCR product of 150-200 bp is typically chosen. To work efficiently with the mapping population it is a great advantage to run the mapping by PCR in half or whole titertrays using multichannel pipette and running the products in gels with slots that fit a multichannel pipette. The products are separated on 2-4% agarose gels. Generally, 2-2.7% agarose can be boiled in a microwave oven while 4% gels are boiled in an autoclave to dissolve all agarose. For size differences of 1-3 bp it will be necessary to run polyacrylamide gels. Microsatellite markers are codominant which makes them very useful but single nucleotide polymorphisms (SNPs) found by sequencing can be developed into markers using parent specific or dCAPS primers (Neff et al., 1998).

An alternative to microsatellite markers is to use AFLP (Vos et al., 1995) and bulked segregant analysis (Michelmore et al., 1991). For AFLP experiments, a cross to _L. filicaulis_ is often used because of the high level of polymorphism that leads to rapid identification of linked markers. A cross to MG-20 can also be used for AFLP analysis using High Efficiency Genome Scanning (HEGS) (Kawasaki and Murakami, 2000). This approach can provide new markers that might be closer than the already known microsatellite markers. A closely linked AFLP marker band can be extracted from the gel, re-amplified, and cloned. It can then be used to isolate BAC/TAC clones to build a contig of the region.

For regional mapping, it is often an advantage to arrange linkage information in colormaps (Kiss et al., 1998). Different genotypes are indicated in Excel tables (see Table 3). Use the “conditional formatting” option of Excel (see manual for relevant Excel version) to colour the cells according to the genotype.

<table>
<thead>
<tr>
<th>TM marker</th>
<th>Chr</th>
<th>Map pos</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
<th>M9</th>
<th>M10</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM0181</td>
<td>1</td>
<td>0.0</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td>h</td>
<td>B</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>TM0193</td>
<td>1</td>
<td>23.2</td>
<td>B</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>TM0122</td>
<td>1</td>
<td>80.2</td>
<td>B</td>
<td>h</td>
<td>a</td>
<td>h</td>
<td>a</td>
<td>h</td>
<td>a</td>
<td>h</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>TM0076</td>
<td>2</td>
<td>47.9</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td>h</td>
<td>B</td>
<td>B</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>TM0550</td>
<td>2</td>
<td>65.0</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>B</td>
<td>h</td>
<td>B</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM0080</td>
<td>3</td>
<td>10.6</td>
<td>a</td>
<td>B</td>
<td>B</td>
<td>h</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>TM0049</td>
<td>3</td>
<td>54.2</td>
<td>h</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>TM0203</td>
<td>3</td>
<td>64.4</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>h</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>TM0127</td>
<td>3</td>
<td>82.5</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>TM0256</td>
<td>4</td>
<td>5.3</td>
<td>a</td>
<td>B</td>
<td>a</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>TM0030</td>
<td>4</td>
<td>34.7</td>
<td>h</td>
<td>B</td>
<td>a</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td>a</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>TM0095</td>
<td>5</td>
<td>18.6</td>
<td>h</td>
<td>h</td>
<td>B</td>
<td>a</td>
<td>B</td>
<td>h</td>
<td>h</td>
<td>B</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>
Table 3. Colormap of the initial mapping of the har1 mutation. From left to right are the TM marker number, the chromosome number, and the cM position on the chromosome. a: homozygous L japonicus MG-20, B: homozygous L japonicus Gifu (har1), h: heterozygous. The map shows 20 mutant plants tested with microsatellite markers from different chromosomal regions. At the location of the recessive mutant gene, all mutant plants will be homozygous L japonicus Gifu. It is clear that har1 is located on the lower end of chromosome 3. It can also be seen that the use of only one microsatellite marker from the middle of each of the chromosomes would have indicated the correct chromosome. Then different microsatellite markers from this chromosome could further pinpoint the position more precisely.
For fine mapping, large mapping populations are needed (several hundred mutant plants). With additional mutants, the chance of finding a recombination very close to the gene increases. At the end of the process close recombination end-points on either side of the gene are needed. The closest markers are then used to screen BAC/TAC libraries to start building the physical map. See Figure 2 for an overview of the procedure. New markers developed from the BAC/TAC will show the direction towards the gene and are used as probes for isolation of additional clones walking towards the gene (see Figure 2 for details). For marker development, polymorphisms will be found primarily in non-coding regions.

Figure 2. Overview of map-based cloning, fine mapping, and physical contig building.

A. Narrowing down the region containing the gene of interest using molecular markers. This illustrates the genetic composition of the chromosome containing the gene of interest from 5 mutant plants (1 to 5) from an F2 mapping population. In this example, the gene X is recessive, so that all mutant plants must be homozygous for the parental genotype of the mutant (in our case Gifu) in the region surrounding the gene. Black region indicates heterozygosity, white region indicates homozygosity in the region surrounding X. Three markers have been scored: Marker M1 shows heterozygosity in plants 1, 2, and 3, while marker M2 shows heterozygosity in plants 2 and 3 and homozygosity in plant 1, indicating that in plant 1 a breakpoint of recombination lies between M1 and M2. This also shows the direction towards the gene X, since the gene must be located in the region of homozygosity in all five plants. The direction is thus to the right of M1. It is also to the right of M2, since M2 is heterozygous in plants 2 and 3, showing that M2 is outside the region containing the gene. A similar rationale can be used for marker M3. This marker must be on the opposite side of gene X, since it detects recombination breakpoint in plants different from those detected with M1 and M2.
B. The first step of physical contig building. A probe representing M2 is used to isolate a BAC or TAC clone(s). End sequences from this BAC are used to make primers to amplify fragments from the parents in which polymorphisms are used to develop new markers: M4 and M5. Checking the genotype of the five plants with M4 and M5 orients the BAC clone. Only plant 3 is heterozygous for M4 while plants 1, 2, and 3 show heterozygocity for M5 demonstrating that M4 is closer to gene X than M5 (and M2). A similar approach is taken for M3 to isolate BAC clone(s) on this side of the gene, giving markers M6 and M7.

C. Closing the contig. Using the relevant end sequences of the first BACs as probes, new BAC clones are isolated. The orientation of these is determined by doing PCR with end primers on the previous BACs. Primers from the end extending the contig will not be productive on the previous BAC. This is continued until the BACs meet to form a physical contig across the region. During this process, more markers are developed from (partial) BAC sequences to anchor and orient the BACs and to further delimit the region containing X. In this example, diagnostic markers are M8/M9 and M10/M11 defining the closest recombination breakpoints on the left and right side of X, respectively. When the marker information is exhausted, the region between the two closest markers must be sequenced and analysed for candidate genes.

One of the pitfalls in contig building is the presence of repetitive or duplicated regions in the genome. Therefore, overlap between clones must be verified by regional sequence analysis (e.g. using the primers for BAC ends) to assure that the BAC clone contains sequences identical to the probe. It is often useful to Blast the sequence to the DNA databases to exclude repetitive Lotus sequences before use as a hybridization probe.

In the end, only very few mutant plants with recombination breakpoints delimiting the locus are of interest (plants 3 and 5, Figure 2). In order to define these few recombination end points more precisely the genotype of the relevant plants can be determined in polymorphic sites within the delimited region by sequencing of PCR products.

Once the region is narrowed down as much as possible, the sequence of the region between the closest markers must be determined and analysed for candidate genes using BlastN and BlastX searches in public databases, focusing on genes likely to have a function in the process of interest. Mutant alleles are then sequenced to identify premature stop codons, missense mutations, changed exon/intron splice sites, insertions/deletions, or even changes in regulatory elements of the promoter.

The final proof of the identity of a cloned candidate gene is complementation of mutants by the wild type gene. For most of the symbiotic mutants, this can be done by hairy root complementation via *Agrobacterium rhizogenes* (ibid). Some of the hypernodulation mutant phenotypes are determined by the shoot genotype (for example har1) and in this situation stable transformants are required.
General map-based cloning strategy

Select \textit{L. japonicus} Gifu mutant

\( \downarrow \)

Cross to e.g. \textit{L. japonicus} Miyakojima

\( \downarrow \)

Rough mapping

\( \downarrow \)

Fine mapping

\( \downarrow \)

BAC/TAC contig with flanking markers

\( \downarrow \)

Candidate genes

\( \downarrow \)

More alleles/completion

1. Mutant populations have been made in \textit{L. japonicus} Gifu. Decide which gene to clone. The mutant should have a phenotype that is easily scored. It should preferably be caused by a mutation in a single gene (3:1 segregation).

2. Cross to \textit{L. japonicus} MG-20 and/or \textit{L. filicaulis}. Collect seeds from the F1 plant. For recessive mutations screen out F2 mutant plants since they are homozygous for the mutant genotype at the locus of interest. For dominant mutations, wild type plants are used for mapping since they are homozygous for the wild type genotype at the locus of interest.

3. Rough map using microsatellite markers or AFLP and bulked segregant analysis. For recessive mutations, find the region with excess of homozygous \textit{L. japonicus} Gifu alleles. For dominant mutations, find the homozygous region with excess \textit{L. japonicus} MG-20/\textit{L. filicaulis} alleles. This will identify chromosome and chromosomal region.

4. Fine map with a larger population and more microsatellite and/or AFLP markers.

5. When the closest markers are less than 0.5-1 cM from the gene, start to isolate BAC/TAC clones using these flanking markers. Develop new markers from BAC/TAC ends to determine direction to the gene. Use these markers to isolate more BAC/TAC clones. Repeat this procedure until a contig containing markers on both sides of the gene of interest is completed. Determine the sequence of the BAC/TAC clone(s).

6. In this contig, look for candidate genes by BlastN/BlastX database searches in public databases. If a candidate gene is found, sequence the gene from a mutant. If a mutation is found, sequence all allelic mutants. If no mutation is found in the candidate gene, consider other candidate genes. All genes inside the flanking markers could eventually be sequenced or the mapping population could be increased to narrow down the region with closer recombinations.

7. Unequivocal evidence for correct gene assignment can be obtained by complementation of a mutant by \textit{Agrobacterium rhizogenes} or \textit{Agrobacterium tumefaciens} transformation.
From commencement of this undertaking in 1999 until end of 2003, two Lotus genes were cloned using transposon tagging and four using map-based cloning approaches (Table 4). Map-based cloning projects of around ten, mostly symbiotic, genes are close to completion and some will be published in 2004. This rapid progress demonstrates the power of genetics and shows that L japonicus is a prime model legume. With completion of the genome sequence in near future opening the possibility of high throughput SNP mapping positional cloning will be both easier and faster. This success in genetic analysis of primarily legume specific processes will no doubt create increased interest for Lotus as experimental system for scientific questions shared with other plants.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Type</th>
<th>Cloning method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nin</td>
<td>Putative transcription factor</td>
<td>Transposon tagging</td>
<td>Schauser et al., 1999</td>
</tr>
<tr>
<td>Pfo</td>
<td>F-box protein</td>
<td>Transposon tagging</td>
<td>Zhang et al., 2003</td>
</tr>
<tr>
<td>SymRK</td>
<td>Receptor kinase</td>
<td>Map-based cloning</td>
<td>Stracke et al., 2002</td>
</tr>
<tr>
<td>Har1</td>
<td>Receptor kinase</td>
<td>Map-based cloning</td>
<td>Krusell et al., 2002</td>
</tr>
<tr>
<td>Astray</td>
<td>Basic leucine zipper and RING-finger motif</td>
<td>Candidate approach</td>
<td>Nishimura et al., 2002</td>
</tr>
<tr>
<td>Nfr1</td>
<td>Receptor kinase</td>
<td>Map-based cloning</td>
<td>Radutoiu et al., 2003</td>
</tr>
<tr>
<td>Nfr5</td>
<td>Receptor kinase</td>
<td>Map-based cloning</td>
<td>Madsen et al., 2003</td>
</tr>
</tbody>
</table>

Table 4. Genes cloned from Lotus mutants, map-based, Ac, or a candidate approach.

**Mapping resources for the future - recombinant inbred lines**

Stabilised mapping populations, recombinant inbred lines (RILs), are important resources for future mapping of single genes as well as quantitative trait loci (QTLs). Since our interspecific population represents alleles adapted to very different environmental conditions (Algeria and Japan) recombinant inbred lines have been developed by generations of selfing of our F2 plants in the L filicaulis x L japonicus Gifu mapping population. The advantage of recombinant inbred lines is that after 8 generations they are expected to be almost completely (99.22%) homozygous. Plants from seeds of a RIL will thus be phenotypically and genetically identical to the parent and therefore different research groups can use seeds of the same lines and use the same map. Accumulation of marker information will further allow genotyping of each plant. At present (November 2003), 78 lines have reached S8. A number of other lines have reached S5, S6, and S7 so that we expect soon to have around 100 RILs.

Because of the possibility of new recombination events in the heterozygotic regions, the map of the recombinant inbred lines will be expected to have the double length of the F2 map and thereby better resolution (Burr and Burr, 1991). Mapping on 44
RILS has now started with 50 microsatellite markers also used for the *L. japonicus* Gifu x *L. japonicus* MG-20 map. Therefore, the maps can be aligned. To obtain seeds of these RILs send an Email to sandal@biobase.dk.

149 RILs have also been developed for *L. japonicus* Miyakojima MG-20 x *L. japonicus* Gifu B-129: (www.kazusa.or.jp/lotus/RIl/ or http://cryo.naro.affrc.go.jp/sakumotu/mameka/lotus-e.htm). The RILs have been developed by Kazusa DNA Research Institute from F2 seeds provided by M Kawaguchi. The final number of these LjMG RILs will be 210. An initial map on this mapping population can be seen on www.kazusa.or.jp/lotus/RIl/R1_map.html The marker scoring can be downloaded as an Excel file from this website.

Finally, Peter Gresshoff’s laboratory has developed RILs for *L. japonicus* Gifu x *L. japonicus* Funakura. These ecotypes are more closely related than Gifu and Miyakojima. These RILs have been used to examine glutamine synthetase function for plant biomass (Limani *et al*., 1999). To obtain seeds of these RILs send an Email to p.gresshoff@botany.uq.edu.au.

For your level of interest, it is therefore worthwhile to analyze *L. japonicus* ecotypes Gifu, Miyakojima, and Funakura and *L. filicaulis* to look for differences and then map the character in the RILs for which the parents show the difference. As they are separate species *L. filicaulis* and *L. japonicus* would be expected to show more differences of biological (phenotypical) characters than the ecotypes and seeds from these RILs are available on a collaboration basis for analysis of different traits and subsequent mapping of single genes and QTLs.

**ACKNOWLEDGEMENTS**

We would like to thank Martin Parniske, Krzysztof Szczygłowski, Peter Gresshoff, Masayoshi Kawaguchi and Yosuke Umehara for making unpublished results available to us.

**REFERENCES**


Pedrosa A, Sandal N, Stougaard J, Schweizer D, and Bachmair A. (2002). Chromosomal...
map of the model legume *Lotus japonicus*. *Genetics* 161, 1661-1672.


