# Microarray-based rapid cloning of an ion accumulation deletion mutant in *Arabidopsis thaliana*

**Ji-Ming Gong, David A. Waner*, Tomoaki Horie*, Shi Lun Li, Rie Horie, Khush B. Abid, and Julian I. Schroeder†**

Division of Biological Sciences, Cell and Developmental Biology Section and Center for Molecular Genetics, University of California at San Diego, La Jolla, CA 92093-0116

Edited by Christopher R. Somerville, Carnegie Institution of Washington, Stanford, CA, and approved September 14, 2004 (received for review July 2, 2004)

**Here we describe the development of a microarray-based mapping strategy to rapidly isolate deletion mutant genes. The presented approach is particularly useful for mapping mutant genes that are difficult to phenotype. This strategy uses masking bulk segregant analysis to mask unrelated deletions, thus allowing identification of target deletions by microarray hybridization of pooled genomic DNA from both WT and mutant $F_2$ populations. Elemental profiling has proven to be a powerful tool for isolation of nutrient and toxic metal accumulation mutants in *Arabidopsis*. Using microarray mapping, a sodium overaccumulation mutant FN1148 was identified as having a 523-bp genomic deletion within the second exon and intron of the *AtHKT1* gene. Further cosegregation, complementation, and comparative analyses among different salt-sensitive mutants confirmed that the deletion within the *AtHKT1* gene is responsible for the sodium overaccumulation in shoots and leaf sodium sensitivity of the FN1148 mutant. These results demonstrate that microarray-based cloning is an efficient and powerful tool to rapidly clone ion accumulation or other genetic deletion mutants that are otherwise difficult to phenotype for mapping, such as metabolic or cell signaling mutants.**

Map-based cloning has been widely used to map and clone mutant genes and is the standard mutant gene isolation strategy at present. Map-based cloning in *Arabidopsis* requires crosses between two genetically variant accessions (ecotypes), and genotyping of ≈500–1,000 $F_2$ mutants is normally required for fine mapping (1). However, for mutants with phenotypes that require intricate analyses, such as ion accumulation mutants (2), traditional map-based cloning is cumbersome because of the following considerations.

1. Phenotyping of individuals in the mapping population is time- and labor-intensive. Ion accumulation mutants have an altered ion accumulation profile as the scorable phenotype, and therefore, map-based cloning requires elemental profiling of each individual by inductively coupled plasma (ICP) analysis.
2. Furthermore, because ion profile phenotypes are quantitative, ion accumulation changes in many mutants will require statistical analyses. Ion profile phenotype scoring in the $F_3$ population of those mutants could be necessary to confirm each $F_2$ mutant. Assuming that 2,000 $F_2$ plants are screened to obtain 500 $F_2$ mutants for fine mapping of a single recessive locus (1), and five $F_3$ progeny of each $F_2$ mutant are scored to obtain statistical confidence (2), a total of 4,500 plants (2,000 $F_2$ plants and 2,500 $F_3$ mutants) would need to be scored by ICP.
3. In addition, parental varieties [in *Arabidopsis*, usually Columbia 0 (Col-0) and Landsberg erecta (Ler), or Wassilevskija (Ws)] used to generate $F_2$ mapping populations may show natural variation in ion accumulation (3). For example, Ws shoots accumulate more cadmium than Col-0 shoots (J.-M.G., unpublished data). Crossing of ion profile mutants to another ecotype can complicate later genetic analyses because of the segregation of natural variation in ion accumulation (3). Together, these factors led us to develop a rapid oligonucleotide microarray-based cloning approach.

Bulked segregant analysis was developed to establish the genetic linkage between a molecular marker and a target mutation by comparing the genetic polymorphisms between two pooled genomic DNA populations: an $F_2$ mutant pool and an $F_1$ pool (4). Recent research has applied this approach to identify single feature polymorphisms (5) and map QTLs by hybridizing pooled genomic DNA from 15 individuals to microarrays (6). However, this QTL mapping approach is based on genetic polymorphisms between two different ecotypes and typically maps a QTL to a possible genetic interval of, for example, 6 centimorgan (cM) (6) (1 cM corresponds to a physical distance of ≈250 kb on average in the *Arabidopsis* genome; ref. 1). In contrast, the modified masking bulked segregant (MABS) analysis developed and applied here can allow rapid and direct identification of a genomic deletion responsible for a phenotype within one or two backcrosses. Oligonucleotide arrays with probes that span genes are used to identify genomic deletions by hybridization of pooled genomic DNA samples.

Understanding how plants take up and process essential nutrients and nonessential toxic elements will facilitate development of nutrient-enriched crops and may lead to novel bioremediation strategies. However, many of the rate-limiting mechanisms controlling ion accumulation in plants remain to be identified. A recent genetic screen that profiled nutrient and trace elements in *Arabidopsis* has led to isolation of ion accumulation mutants (2). By elemental profiling of ≈6,000 fast neutron-mutagenized *Arabidopsis* plants using ICP, 51 mutants were identified with significantly altered element accumulation; these are referred to as ion profile mutants. Further mapping and cloning of these mutants will lead to a better understanding of the underlying molecular mechanisms by which plants control the transport and accumulation of nutrient elements and toxic metals. To our knowledge, direct isolation of a eukaryotic deletion mutant gene by hybridization of genomic DNA to microarrays has not been previously published. Here, we demonstrate the utility of a microarray-based cloning technique to successfully clone a sodium ($Na^+$) overaccumulation mutant gene by using two mutant microarray hybridizations.

## Materials and Methods

**Plant Growth and ICP Analyses.** FN1148 $M_3$ plants were grown in Sunshine Mix soil (Sun-Gro Horticulture, Bellvue, WA) supplemented with trace amounts of nutrient and toxic metals as described (2). When indicated, FN1148 $M_3$ plants were grown in soil without any added heavy metals, but flooded from the bottom with 10 mM NaCl once for 60 h at 3–4 weeks of age. Two to three leaves were sampled from soil-grown plants and dried overnight at 80°C in an oven. Alternatively, when indicated,

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ICP, inductively coupled plasma; Col-0, Columbia 0; MABS, masking bulk segregant; 5′-MP, 5′ most-located probes; 3′-MP, 3′ most-located probes.

*D.A.W. and T.H. contributed equally to this work.

†To whom correspondence should be addressed. E-mail: julian@biomail.ucsd.edu.

© 2004 by The National Academy of Sciences of the USA



www.pnas.org/cgi/doi/10.1073/pnas.0404780101

plants were grown in sterile hydroponic solution containing 1.25 mM $KNO_3$, 0.625 mM $KH_2PO_4$, 0.5 mM $MgSO_4$, 0.5 mM $Ca(NO_3)_2$, and the additional micronutrients: 17.5 $\mu$M $H_3BO_3$, 0.125 $\mu$M $CuSO_4$, 3.5 $\mu$M $MnCl_2$, 0.25 $\mu$M $ZnSO_4$, 0.05 $\mu$M $NaMoO_4$, 2.5 $\mu$M NaCl, and 0.0025 $\mu$M $CoCl_2$ (7), at 24°C with 16-h light/8-h dark cycles. At ≈3–4 weeks of age, 2.5 mM NaCl was applied to the hydroponic solution; 6 days after $Na^+$ addition, stems, rosette leaves, and roots were harvested separately and dried. Before harvesting, root samples were rinsed in deionized water and then washed at room temperature in 25 mM $CaCl_2$ (pH $<$ 5) for 7 min. Dried samples were digested in ultrapure nitric acid (Sigma) as described (8) and analyzed by ICP optical emission spectroscopy (Perkin-Elmer Optima 3000XL, Applied Biosystems) at the Scripps Institution of Oceanography (University of California at San Diego, La Jolla).

**Phenotypic Growth Analyses.** Seeds were germinated on 0.5× Murashige and Skoog medium containing plates for 3 days, and then transferred to minimal medium plates containing 1 mM $H_3PO_4$, 0.5 mM $Ca(NO_3)_2$, 1 mM $MgSO_4$, 50 $\mu$M KCl, 3 mM Mes, 100 $\mu$M FeNa-EDTA, and the additional micronutrients 7 $\mu$M $H_3BO_3$, 1.4 $\mu$M $MnSO_4$, 1 $\mu$M $ZnSO_4$, 4.5 $\mu$M KI, 0.1 $\mu$M $CuSO_4$, 0.2 $\mu$M $Na_2MoO_4$, 10 nM $CoCl_2$, and 1% agarose, supplemented with NaCl as indicated (pH 5.7, adjusted by NaOH). Plants were grown vertically at 24°C with 16-h light/8-h dark cycles. Phenotypes were evaluated at the indicated times.

**Chip-Based Cloning and Related Methods.** Two pools of genomic DNA were extracted from young leaves of either WT or mutant plants by using DNAeasy plant mini kit (Qiagen, Valencia, CA). Three hundred nanograms of each pool was randomly biotin-labeled by using Bioprime DNA-labeling system (Invitrogen), and the derived biotin-labeled DNA fragments (≈100 bp, data not shown) were purified by using the protocol suggested by the manufacturer. To check the yield and quality, 1/20 of each labeled sample was separated on 1% agarose gel, and the rest was subjected to oligonucleotide microarray-based chip hybridization by the University of California at San Diego Genechip Core. *Arabidopsis* ATH1 oligonucleotide-based genome arrays were used, and hybridization and washing were performed by using the standard Affymetrix protocol for RNA analyses (www.affymetrix.com/products/arrays/specific/arab.affx).

Raw chip hybridization data were subjected to background subtraction using the "mas" background correction method, and normalization was performed by using the "quantiles" method of the Bioconductor "Affy" package for the R statistical language (9, 10). Mismatch probes were not analyzed, because these have been reported not to render more accurate analyses of genomic DNA hybridization signals compared to perfect match probes (5). Potential deletions were identified by comparing differences between the individual perfect match probe intensities in corresponding mutant and WT arrays by developing a PYTHON program. These programs are available upon request. Probe sets with potential deletions were selected by using adjustable criteria, including that the probe set has at least two adjacent probes that show hybridization intensities decreased by ≥30% in mutant compared to WT controls in two independent mutant microarray experiments.

PCR analyses were performed to confirm deletions and to determine the exact deletion site and size. Genomic fragments carrying a target deletion were recovered by PCR and subcloned into the pGEM-T easy vector (Promega). Sequencing was performed by Retrogen (La Jolla, CA). DNA for regular PCR analysis was extracted with a simplified CTAB method as described (11). To determine the effect of genomic deletions on transcripts, total RNA was extracted from both the FN1148 mutant and WT plant roots, and RT-PCR was performed as described (8). PCR fragments were subcloned and sequenced (Retrogen).

**MABS Analysis.** MABS analysis was performed by backcrossing the FN1148 mutant once to the same ecotype (Col-0) from which it originated to circumvent natural variation in ion accumulation between different ecotypes (see Introduction). Backcrossing into the parental ecotype also circumvented problems that could result from naturally occurring genomic deletion polymorphisms between two different ecotypes (12). $F_2$ mutants with $Na^+$ overaccumulation phenotypes were identified by ICP, and a DNA pool was extracted from five of these individual $F_2$ mutant plants. Control DNA pools were extracted from WT (Col-0) plants for control microarray hybridizations.

**Complementation Analyses of FN1148 Mutant.** The *AtHKT1* cDNA with 0.8 kb of the native promoter region was subcloned into the pBI101-Hm vector to complement the FN1148 mutant. The construct was transformed into the FN1148 mutant by direct *Agrobacterium*-mediated transformation using the floral dip method (13). Three transgenic lines showing *AtHKT1* mRNA expression were selected and analyzed. Heterozygous *AtHKT1* complemented FN1148 transgenic lines were subjected to hygromycin selection before transfer to 0.5× Murashige and Skoog plates for recovery for 5 days.

**Bioinformatic Analyses.** Bioinformatic analyses were pursued to analyze the Affymetrix probe coverage of *Arabidopsis* genes. The *Arabidopsis* gene coding sequences were obtained from ftp://ftp.Arabidopsis.org/home/tair/Sequences/blast_datasets. The Affymetrix ATH1 oligonucleotide microarray index file to match the probe set IDs and the corresponding *Arabidopsis* Genome Initiative gene numbers was obtained from ftp://tairpub:tairpub@ftp.Arabidopsis.org/home/tair/Microarrays/Affymetrix. The Affymetrix ATH1 probe set information file containing probe set ID numbers, spatial position of each probe on the chip, probe interrogation positions along ORFs, and 25-bp sequence of each probe was obtained from www.affymetrix.com/support/technical/byproduct.affx?product=arab. The above three files were then reformatted and subjected to analyses to determine the probe set coverage for the corresponding gene coding sequences. For these analyses, ≈80% of the 22,746 probe sets were analyzed, for which an unequivocal annotation was obtained.

## Results

**Oligonucleotide Microarray-Based Mapping of Fast Neutron Bombarded Mutants.** The Affymetrix ATH1 genome microarray contains 22,746 probe sets representing ≈23,750 *Arabidopsis* genes (14). Each probe set has an average of 11 perfect match probes (www.affymetrix.com/products/arrays/specific/arab.affx). These probe sets were used to identify genomic deletions by hybridization of both mutant and WT genomic DNA to oligonucleotide-based microarrays. Note that genomic DNA, instead of mRNA, was subjected to biotin labeling and hybridization to microarray genechips in this microarray-based cloning strategy, because the mRNA abundance of many genes is too low to be efficiently and reliably detected upon hybridization to oligonucleotide microarrays. Furthermore, mRNA level changes could be the result of secondary effects of a mutant. If one or more probes show decreased hybridization intensity in the mutant array compared to the WT array, a deletion may exist within the genomic DNA corresponding to that probe set. To determine the efficiency of ATH1 genome arrays to identify a genomic deletion by this microarray-based cloning method, we first analyzed the locations of probe sets covering the ORFs of annotated genes (Fig. 1*A*). The 5′ most-located probes (5′-MP) were analyzed (Fig. 1 *A* and *B*) for each gene. The average 5′-MP is located 42%

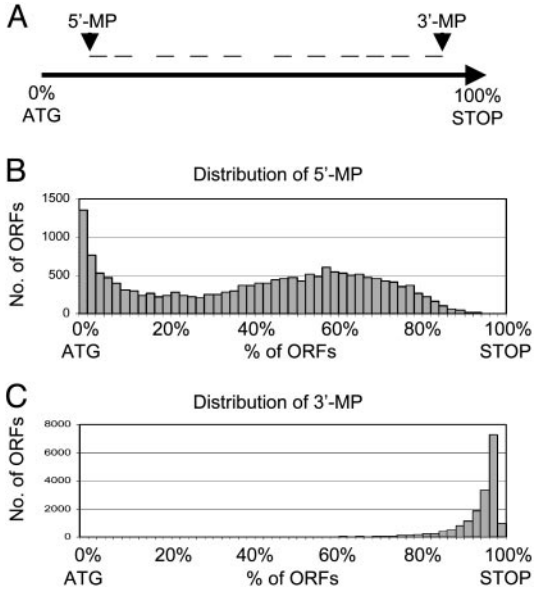


**Fig. 1.** ATH1 probe set distribution along ORFs of *Arabidopsis* genes. (*A*) Illustrates an example of a probe set distributed along a corresponding ORF (average 11 perfect match probes per gene). ATG represents the start codon position of an ORF and STOP represents the stop codon position. 5′-MP represents the 5′-most located probe(s) along an ORF and 3′-MP represents the 3′-most located probe(s). (*B*) The distribution of the 5′-most located probes (5′-MP) is illustrated for ORFs represented on the ATH1 GeneChip. (*C*) The distribution of the 3′-most located probes (3′-MP) is illustrated for ORFs represented on the ATH1 GeneChip. Locations of 5′-MP (*B*) and 3′-MP (*C*) are illustrated as percentage of the entire ORF of each analyzed gene on the *x* axis.

downstream of the ORF start codons and shows a broad distribution for the genes represented on the ATH1 genechip (Fig. 1*B*). The average 3′ most-located probe (3′-MP) of a probe set (Fig. 1 *A* and *C*) is located 95.4% downstream of the ORF start codons and is mainly close to the stop codon of genes (Fig. 1*C*). Assuming an average fast neutron deletion size of 1–4 kb (15, 16), we estimated that a substantial number of deletion mutant genes could be cloned, considering that the average *Arabidopsis* gene size is ≈2 kb (17). In some cases, deletion sizes can be >5 kb (18–20), which may cover more than one gene, thus increasing the deletion identification efficiency using ATH1 microarray-based mapping. These bioinformatic analyses indicated that the Affymetrix ATH1 microarrays could be used for developing rapid microarray-based cloning of fast neutron bombarded mutants.

**Recessive Mutant FN1148 Has Strong Shoot $Na^+$ Overaccumulation and Salt-Sensitive Phenotypes.** A $Na^+$ overaccumulation mutant FN1148 was isolated in our laboratory within the ICP elemental profiling screen (2) (http://www-biology.ucsd.edu/labs/schroeder/supplemental.html). As shown in Fig. 2*A*, ICP profiling of individual $M_3$ mutant progeny derived from the original FN1148 mutant showed significant $Na^+$ overaccumulation in shoots ($P = 4.56e^{-10}$, $n = 10$ $M_3$ plants). ICP profiling of $M_3$ FN1148 mutant plants was also performed in potting soil without added heavy metals, but with one time flooding of trays with 10 mM NaCl at 3 weeks of age (see *Materials and Methods*). Similar $Na^+$ overaccumulation in shoots of the $M_3$ progeny of the FN1148 mutant was observed ($n = 10$). Further analyses showed that, when grown on minimal medium containing agar plates supplemented with 40 mM NaCl, the FN1148 mutants showed curled, smaller size rosette leaves (Fig. 2*B*) compared to WT plants (Fig. 2*C*). The salt sensitive phenotype of the FN1148 mutant (Fig. 2*B*) correlates with $Na^+$ overaccumulation in

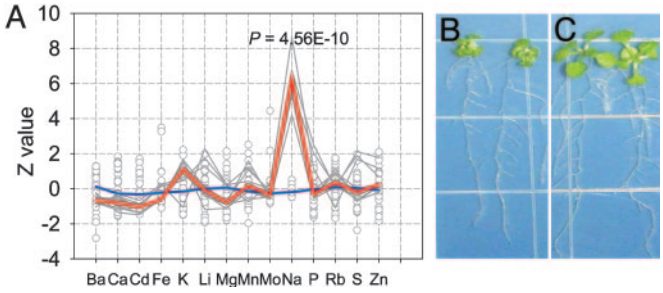


**Fig. 2.** FN1148 mutant overaccumulates $Na^+$ in shoots and is salt sensitive. (*A*) Fourteen different elements in FN1148 mutant $M_3$ plants were determined by ICP (*x* axis). The blue line represents the median *Z* value (standard deviation from the mean value of control WT, *y* axis) of WT, and the red line represents the median *Z* value of the FN1148 mutant plants ($n = 32$ plants for the WT, and $n = 10$ plants for the FN1148 mutant). Gray open circles in *A* represent *Z* values of individual WT control plants; gray lines in *A* represent *Z* values of individual FN1148 mutant plants. (*B* and *C*) Salt sensitivity analyses of FN1148 mutant (*B*) and WT (*C*) plants were performed on minimal medium containing plates (see *Materials and Methods*) supplemented with 40 mM NaCl. Phenotypes were evaluated at 10 days of age.

rosette leaves (Fig. 2*A*). No dramatic visible morphological phenotype of FN1148 was observed without exposure to excess $Na^+$ under the imposed conditions (data not shown).

After a single backcross of FN1148 to its WT parent (Col-0), the $Na^+$ overaccumulation phenotype was scored in the backcrossed $F_1$ and $F_2$ generations. No ion profile phenotype was observed in 22 $F_1$ plants (data not shown). In the $F_2$ progeny, six plants showed the ICP phenotype, whereas 28 showed a WT $Na^+$ accumulation level, suggesting that a single recessive locus is responsible for the ion profile phenotype ($P = 0.322102$, $\chi^2 = 0.980392$).

**Identification of a 523-bp Genomic Deletion in *AtHKT1* Gene by Microarray-Based Mapping.** To test the hypothesis that oligonucleotide-based microarrays can be used to identify deletions, we performed two different types of chip hybridization experiments. Genomic DNA extracted directly from the original $M_3$ FN1148 mutant without backcrossing was used in one experiment. In the other experiment, pooled genomic DNA of five individual backcross-derived $Na^+$ overaccumulation $F_2$ mutant plants was used for MABS analyses. Genome-wide mutant DNA hybridization intensities were compared to two control chips in which WT genomic DNA was hybridized to ATH1 chips. As expected, the experiment using the $M_3$ FN1148 mutant DNA pool showed many putative deletions as judged by lack of or poor hybridization intensity compared to the WT control. Significant numbers of putative deletions were also identified in the experiment using MABS analysis of the $Na^+$ overaccumulation $F_2$ plant DNA pool.

Reduced hybridization to individual oligonucleotide probes is likely due to experiment-to-experiment variability and variation in the hybridization efficiency between individual probes. To identify potential deletions in the presence of significant noise resulting from individual oligonucleotide hybridizations, the two experiments were compared to identify common deletions of ≥2 adjacent probes in both experiments. Among the >250,000 individual probes analyzed, only four potential common deletions were identified in the following ATH1 probe sets: 245156_at, 255812_at, 266145_at, and 266146_at (Fig. 3*A*). These four potential deletions showed reduced hybridizations to two neighboring oligonucleotide probes in the two mutant arrays compared to the two WT arrays (Fig. 3*A*, grouped red signals of two adjacent probes in two experiments).

Further analyses of the absolute hybridization reduction

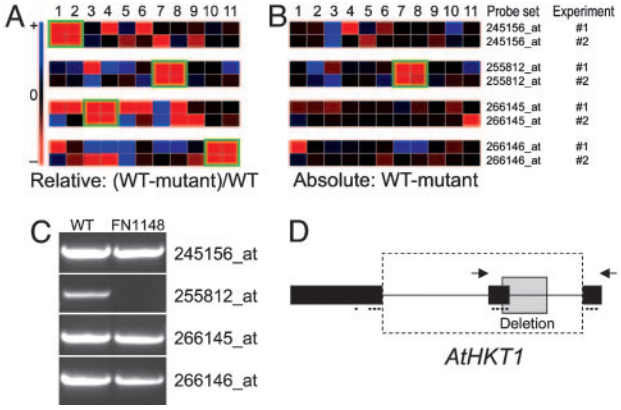


**Fig. 3.** Chip-based mapping of FN1148 mutant. (*A*) Four potential common deletions were predicted in two experiments by chip-based mapping (green boxes). Relative hybridization intensity changes [(WT − mutant)/WT] of all oligonucleotide probes for the four genes are shown. (*B*) Absolute hybridization intensity change (WT − mutant) suggested the most promising deletion (green box). Red represents mutant hybridization intensities that are lower than WT, whereas blue represents mutant hybridization intensities higher than WT. Black represents the same hybridization intensity in WT and in the FN1148 mutant. Probe set ID numbers are shown at the far right in *B*. The hybridization differences for experiment 1 (MABS pool) and experiment 2 (FN1148 $M_3$ pool) are illustrated for the four loci. Probe numbers (1–11) are shown above each hybridization signal in *A* and *B*. Note that all four loci indicate a putative deletion in two neighboring probes in both experiments (four grouped red signals in *A*). (*C*) PCR analyses confirmed one deletion in probe set 255812_at. PCR primers were located within the deleted regions as indicated by chip hybridization and amplified products were confirmed by sequencing. Pooled DNA samples for chip hybridization were used for PCR. (*D*) PCR cloning located a 523-bp genomic deletion within the second exon and intron of the *AtHKT1* gene (gray box). RT-PCR cloning revealed an incorrect splicing pattern of the *AtHKT1* gene in FN1148 (dotted box). Black boxes represent exons, and black lines represent introns. Arrows represent the primers used for PCR cloning. Short lines under exons show approximate locations of the 11 ATH1 probe positions. Dotted box represents the sequence removed in FN1148 by incorrect splicing.

(WT − mutant) of the four potential deletions showed that only the hybridization reduction in probe set 255812_at showed large absolute reductions in two adjacent probes in both experiments (Fig. 3*B*). The other three putative loci were found to include exon probes that are separated by an intron, which may have decreased the hybridization signal, thus decreasing the hybridization signal-to-noise ratio. PCR analyses showed that, of the four putative mutations, the deletion within probe set 255812_at is the only true deletion (Fig. 3*C*). This probe set corresponds to the gene *AtHKT1* (At4g10310). Two PCR primers located outside the potential deletion region were designed for PCR cloning of the genomic fragment carrying the deletion (Fig. 3*D*, black arrows). DNA sequencing identified a 523-bp genomic deletion within the *AtHKT1* gene, which includes the last 72 bp of the second exon and the first 451 bp of the second intron (Fig. 3*D*, gray box). The genomic deletion includes a loss of the last 5 bp of probe 7 and the whole probe 8 (Fig. 3 *A* and *B*, probe set 255812_at, and *D*). Further RT-PCR analyses showed that the genomic deletion within the *AtHKT1* gene in FN1148 resulted in an incorrectly spliced *athkt1* mRNA (Fig. 3*D*, dotted box). This incorrectly spliced *athkt1* mRNA lost the whole second exon, which caused a point mutation from methionine (ATG) to isoleucine (ATT) and the loss of 78 amino acid residues. No ORF shift was observed in the third exon (Fig. 3*D*). *AtHKT1* encodes a sodium transporter (21) with eight transmembrane domains (22). The deletion is located within transmembrane domains 6 and 7 and the following P-loop.

The findings that only one deletion was identified by comparing the experiment using FN1148 $M_3$ plants and the experiments using MABS analyses suggests that (*i*) comparative analyses of two independent experiments will aid considerably in identifying target deletions in the presence of hybridization noise of individual oligonucleotide probes. (*ii*) MABS analyses mask unrelated deletions, thus allowing rapid identification of ORF deletions responsible for a phenotype by using oligonucleotide microarrays.

**$Na^+$ Overaccumulation Phenotype Cosegregates with the *AtHKT1* Deletion in FN1148.** To further analyze whether the deletion within *AtHKT1* is responsible for the $Na^+$ overaccumulation phenotype of the FN1148 mutant, cosegregation analyses were performed by using newly selected plants from the $F_2$ segregating population derived from a backcross of FN1148 to WT Col-0. ICP profiling showed that 5 of 31 $F_2$ plants showed a dramatic $Na^+$ overaccumulation phenotype in leaves (Fig. 4*A Upper*). PCR experiments were then performed to determine which of the individual plants showed a homozygous deletion in the *AtHKT1* gene. The homozygous deletion was only identified in the same five individuals that showed the leaf $Na^+$ overaccumulation phenotype (Fig. 4*A Lower*, PCR/A). These results support the hypothesis that the deletion within the *AtHKT1* gene is responsible for the ICP phenotype of the FN1148 mutant.

**Transgenic Expression of *AtHKT1* Gene Complements Salt Sensitivity of FN1148 Mutant.** An *AtHKT1* cDNA construct driven by a native *AtHKT1* promoter fragment was transformed into the FN1148

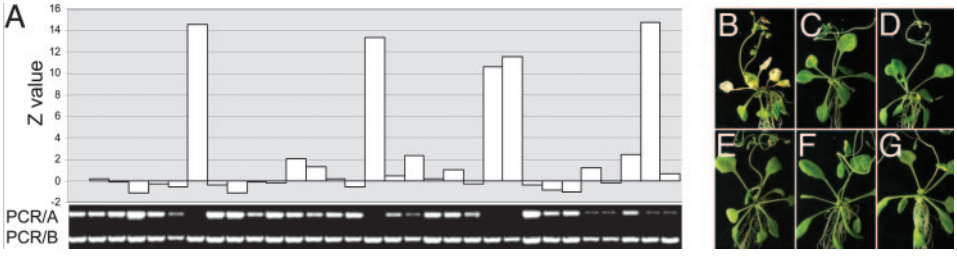


**Fig. 4.** Deletion in *AtHKT1* is responsible for FN1148 phenotypes. (*A*) Cosegregation of $Na^+$ overaccumulation in leaves phenotype (*Upper*) and the *homozygous* deletion within the *AtHKT1* gene (*Lower*, PCR/A). $Na^+$ accumulation was analyzed in individual $F_2$ plant leaves by ICP optical emission spectroscopy and plotted against *Z* values (standard deviations from the mean value of control WT). Homozygous *AtHKT1* deletions were identified by PCR with primers located within the deletion (PCR/A). PCR amplification (32 cycles) of an undeleted AtHKT1 fragment was used as a control to ensure functional amplification of genomic DNA (PCR/B). Leaf tissues of 31 $F_2$ plants were subjected to combined ICP sampling and DNA extraction. (*B–G*) The *AtHKT1* cDNA driven by the AtHKT1 promoter complemented the salt sensitivity of the FN1148 mutant. The FN1148 mutant (*B*), *AtHKT1* complemented FN1148 transgenic line (*C*), and WT (*D*) were transferred to minimal medium plates (see *Materials and Methods*) supplemented with 40 mM NaCl. For controls, in the absence of NaCl, the FN1148 mutant (*E*), *AtHKT1* complemented FN1148 transgenic line (*F*), and WT (*G*) were grown on minimal medium plates. Phenotype was scored on three independent transgenic lines, the FN1148 mutant, and WT seedlings 18 days after transfer to the minimal medium plates. One representative transgenic line is displayed in *C* and *F*.

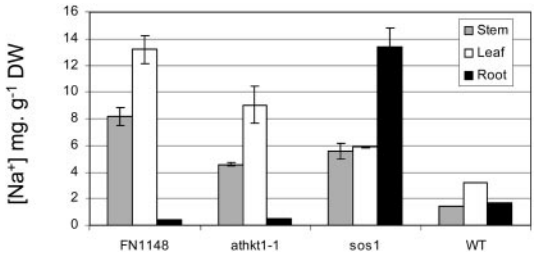


**Fig. 5.** FN1148 mutant shows a similar $Na^+$ accumulation phenotype to other *athkt1* allele. Three-week-old FN1148, *athkt1–1*, *sos1*, and WT plants grown in hydroponic medium were exposed to 2.5 mM NaCl for 6 days. Stems, rosette leaves, and roots were harvested, and the $Na^+$ content was determined by ICP optical emission spectroscopy. Data show mean values ± SE, $n = 3$ plants.

mutant. Three independent transgenic lines expressing the WT *AtHKT1* mRNA were subjected to whole plant salt sensitivity analyses. In the presence of 40 mM $Na^+$, the FN1148 mutant showed a significant salt sensitive phenotype (Fig. 4*B*): wilted, chlorotic, or brown rosette leaves. In contrast, all of the three transgenic lines (e.g., Fig. 4*C*) grew similar to WT plants (Fig. 4*D*). In the absence of $Na^+$ stress, no difference was observed between the FN1148 mutants (Fig. 4*E*), the complementation lines (e.g., Fig. 4*F*), and WT plants (Fig. 4*G*). These results showed that *AtHKT1* complements the salt sensitivity of the FN1148 mutant.

**Comparative Analyses Between Different Salt-Sensitive Mutants.** Comparative ICP analyses between FN1148, the salt overly sensitive mutant *sos1* (23), and a T-DNA insertion allele *athkt1–1* (24) (kindly provided by P. M. Hasegawa and A. Rus, Purdue University, West Lafayette, IN) were pursued. As shown in Fig. 5, $Na^+$ overaccumulation in stems and rosette leaves was observed in both the FN1148 mutant and the *athkt1–1* T-DNA insertion mutant. Moreover, in roots of FN1148 and *athkt1–1*, less $Na^+$ was detected compared to the WT (Fig. 5; $P < 0.0008$ FN1148, $P < 0.014$ *athkt1–1*). In the salt overly sensitive *sos1* mutant, which lies in a different transporter gene (25), overall $Na^+$ accumulation in stems, leaves, and roots was increased compared to the WT (Fig. 5). These results showed that the FN1148 mutant shows similar $Na^+$ overaccumulation in shoot and $Na^+$ underaccumulation in root phenotypes to the *athkt1–1* allele and other *athkt1* alleles (26, 27).

## Discussion

**Microarray-Based Mapping Is an Efficient and Rapid Mutant Gene Isolation Method.** Here, we describe the development of a microarray-based rapid cloning strategy and the successful application of this method to clone ion profile mutant genes (2). For traditional map-based cloning of recessive ion profile mutants, which generally have quantitative phenotypes (2), ≈2,000 cross-originated $F_2$ plants would have to be ICP profiled to obtain 500 homozygous $F_2$ mutants for fine mapping. Moreover, ≈2,500 $F_3$ plants might need to be ICP profiled to confirm the 500 $F_2$ mutants due to the quantitative nature of ICP phenotypes (see Introduction). In contrast, 34 $F_2$ backcrossed plants were ICP profiled here to obtain approximately six homozygous mutants, five of which were pooled for MABS analysis and the microarray-based mutant gene identification (see *Materials and Methods*). By hybridizing genomic DNA of both pooled mutant and WT plants and comparative analyses of two mutant and two control oligonucleotides microarrays, we have successfully identified a 523-bp genomic deletion in the *AtHKT1* gene of the $Na^+$ overaccumulation mutant FN1148. A mutant $Ca^{2+}$/calmodulin-dependent protein kinase gene was identified by using a microarray approach based on mRNA abundance (28). The general applicability of this elegant mRNA hybridization-based study may be limited because some mutations may affect the transcript abundance of other genes and because of the low mRNA levels of many expressed genes.

Recent studies of other *athkt1* alleles showed $Na^+$ accumulation phenotypes similar to that of the FN1148 *athkt1* allele (26, 27). The presented findings further confirm the predicted function of the AtHKT1 $Na^+$ transporter in controlling $Na^+$ distribution between shoots and roots (26, 27) under diverse growth conditions (Figs. 2*A*, 4*A*, and 5). The microarray-based cloning strategy requires only backcrossing to the parental ecotype, thus circumventing natural variation related phenotypes resulting from crosses between two ecotypes, which is required for map-based cloning. Importantly, the rapid cloning method described here can be expanded to cloning of deletion mutant genes in other organisms with extensive microarray coverage, including human disease genes.

**MABS Analysis Greatly Enhances Microarray-Based Cloning.** The fast neutron mutagenized seeds used here have an average of 60–70 deletions per genome (www.arabidopsis.com/main/cat/seeds/M2/FN/!2fb.html). After one backcross, 50% of the mutagenized genome is replaced by WT genome, which suggests that an average of 30–35 deletions are still present in each $F_2$ genome. To identify a mutation by microarray analyses of individual mutant lines, additional backcrosses would be necessary to remove additional deletions. MABS analysis can enable mutant locus identification after a single backcross as demonstrated here, or after two backcrosses.

Conventional bulked segregant analysis usually is conducted by crossing a mutant to a different ecotype to generate two DNA pools from $F_1$ and $F_2$ mutant plants, between which only the target mutation region would show a genetic polymorphism, thus allowing identification of a molecular marker that is linked to a target mutation region (4). MABS analysis allows direct identification of deleted genes (Fig. 3). In a selected $F_2$ mutant pool, only the target mutation region and the neighboring linked regions are homozygous, whereas other mutations are arbitrary and would be heterozygous in a pool of sufficiently large size. When hybridized to microarrays, the heterozygous deletions will be masked by the corresponding WT genome. Thus, with MABS analysis, only the homozygous target deletion can be detected by comparing the hybridization intensity of both the WT pool and backcrossed $F_2$ mutant pool independent of how many unrelated heterozygous deletions are still present after one backcross.

Genomic DNA was pooled from five selected $F_2$ mutants for MABS analysis in the present study. With a pool of 10 $F_2$ mutants, the probability of linking an unlinked molecular marker to a target mutation is theoretically $2 \times 10^{-6}$ $(2(1 - [1/4]^n)[1/4]^n)$, where $n$ is number of pooled individuals (4), assuming that a single WT locus copy is sufficient to mask the corresponding unlinked deletions. However, our data show that, because of hybridization noise, independent replicate experiments are needed to identify the target mutation locus. Furthermore, there is a possibility that a deletion that is not responsible for the phenotype might be genetically closely linked to a target deletion. Such closely linked deletions will also be identified by microarray-based MABS analyses. Complementation analyses (Fig. 5 *B–D*) will be necessary to rule out unrelated deletions.

**Hybridization Noise in Oligonucleotide Microarray-Based Mutant Cloning.** For microarray analyses of mRNA expression using the ATH1 chips, 11 probes contribute to the expression value of each gene, and therefore variation in individual probe hybridization intensities is "averaged out" by statistical analyses (www.affymetrix.com). However, in the present analysis, genomic DNA, which has more nucleotide sequence complexity compared to mRNA, was used for chip hybridization, and we

searched >250,000 perfect match probes for intensity changes at the individual probe level. Therefore, noise of genomic DNA hybridization to individual probes is a more significant problem in MABS analysis. To identify deletions, we initially searched for putative deletions that spanned at least two adjacent probes. No potential deletions were identified when we searched both experiments for a ≥40% decreased signal. However, when using a less stringent criterion (at least two adjacent probes with ≥30% decreased signal), significant numbers of potential deletions were identified in each individual experiment, which are in part due to noise in the hybridization signal of individual probes. Then, the two experiments were compared to analyze for common probes with lower hybridization signals compared to WT controls. Only four common potential deletions were identified (Fig. 3*A*). These findings suggest that replicates using independent DNA samples are necessary for rapidly identifying deletions in the presence of significant hybridization noise. Note that typically larger fast neutron deletions spanning more than two probes would be easier to identify than the FN1148 mutant, which deletes only 5 bp in one probe and an additional entire probe (Fig. 3 *A*, *B*, and *D*). If a deletion covers only one probe, more than two independent microarray experiments and possibly increased hybridization stringency would greatly aid in overcoming hybridization noise problems. In addition, rough mapping of a mutant locus to reduce the number of probes analyzed, if natural variation does not occur for the analyzed elements, could be used in conjunction with chip mapping.

**Improved Design of Oligonucleotide Microarrays for GeneChip-Based Cloning.** The ATH1 probe coverage is sufficient for mapping a substantial number of fast neutron-generated mutants (Fig. 1) based on the large size of fast neutron deletions (15, 16, 18–20). However, there is a significant probability that some small and 5′ ORF deletions may not be detectable by ATH1 microarray-based hybridizations if these small deletions do not fall within the probe sets. Furthermore, even a large deletion may not be detectable if most of the deletion falls within an intron. Improved design of oligonucleotide arrays could address these limitations. For example, exons could be tiled contiguously with oligonucleotides to cover all of the possible deletions within exons. In addition, two intron probes could be included for each intron/exon junction to cover deletions that fall mainly within introns. In addition, longer oligonucleotide probes would greatly reduce the above discussed hybridization noise by increasing hybridization specificity, particularly when applying this method to isolate mutant genes in organisms with a large genome size. Oligonucleotide microarrays with both improved probe coverage and longer oligonucleotide probes are becoming publicly available (www.chem.agilent.com, www.nimblegen.com, and www.affymetrix.com).

In summary, a practical microarray-based masking bulk segregant cloning strategy was developed and successfully used to rapidly clone a shoot $Na^+$ overaccumulation mutant gene. We also analyzed important sources of noise and demonstrated and discussed methods to circumvent these problems. The presented method is applicable to large-scale ion profile mutant cloning and also to isolation of genes that confer mutant phenotypes requiring intricate analyses of each individual, such as metabolic and cellular signaling mutants.

We thank Annette Deyhle (The Scripps Institution of Oceanography, University of California at San Diego, La Jolla) for use of the ICP facility, Dr. Justin Borevitz (The Salk Institute) for helpful discussions, Lutfunnessa Shireen (GeneChip Core Facility, University of California at San Diego) for chip hybridizations, Dr. Jian-Kang Zhu for providing *sos1* seeds (University of California, Riverside), and Jared Young for reading of the manuscript. This research was supported by National Science Foundation Grants 0077378-DBI and IBN-0419695 (to J.I.S. and the Ionome Group), Department of Energy Grant DOE-DE-FG02–03ER15449, and National Institute of Environmental Health Sciences Grant 1 P42 ESI0337 (to J.I.S.).

1. Lukowitz, W., Gillmor, C. S. & Scheible, W.-R. (2000) *Plant Physiol.* **123,** 795–806.
2. Lahner, B., Gong, J. M., Mahmoudian, M., Smith, E. L., Abid, K. B., Rogers, E. E., Guerinot, M. L., Harper, J. F., Ward, J. M., McIntyre, L., *et al.* (2003) *Nat. Biotechnol.* **21,** 1215–1221.
3. Bentsink, L., Yuan, K., Koornneef, M. & Vreugdenhil, D. (2003) *Theor. Appl. Genet.* **106,** 1234–1243.
4. Michelmore, R., Paran, I. & Kesseli, R. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 9828–9832.
5. Borevitz, J. O., Liang, D., Plouffe, D., Chang, H.-S., Zhu, T., Weigel, D., Berry, C. C., Winzeler, E. & Chory, J. (2003) *Genome Res.* **13,** 513–523.
6. Wolyn D, Borevitz J, Loudet O, Schwartz C, Maloof J, Ecker J, Berry, C. & Chory, J. (2004) *Genetics* **167,** 907–917.
7. Arteca, R. N & Arteca, J. M. (2000) *Physiol. Plant.* **108,** 188–193.
8. Gong, J.-M., Lee, D. A. & Schroeder, J. I. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 10118–10123.
9. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. (2003) *Bioinformatics* **19,** 185–193.
10. Ihaka, R. & Gentleman, R. (1996) *J. Comput. Graph. Stat.* **5,** 299–314.
11. Lukowitz, W., Mayer, U. & Jurgens, G. (1996) *Cell* **84,** 61–71.
12. Aukerman, M. J., Hirschfeld, M., Wester, L., Weaver, M., Clack, T., Amasino, R. M. & Sharrock, R. A. (1997) *Plant Cell* **9,** 1317–1326.
13. Clough, S. J & Bent, A. F. (1998) *Plant J.* **16,** 735–743.
14. Redman, J. C., Haas, B. J., Tanimoto, G. & Town, C. D. (2004) *Plant J.* **38,** 545–561.
15. Li, X., Song, Y., Century, K., Straight, S., Ronald, P., Dong, X., Lassner, M. & Zhang, Y. (2001) *Plant J.* **27,** 235–242.
16. Li, X. & Zhang, Y. (2002) *Funct. Integr. Genomics* **2,** 254–258.
17. *Arabidopsis* Genome Initiative (2000) *Nature* **408,** 796–815.
18. Tor, M., Gordon, P., Cuzick, A., Eulgem, T., Sinapidou, E., Mert-Turk, F., Can, C., Dangl, J. L. & Holub, E. B. (2002) *Plant Cell* **14,** 993–1003.
19. Silverstone, A. L., Ciampaglio, C. N. & Sun, T.-P. (1998) *Plant Cell* **10,** 155–170.
20. Michaels, S. D. & Amasino, R. M. (1999) *Plant Cell* **11,** 949–956.
21. Uozumi, N., Kim, E. J., Rubio, F., Yamaguchi, T., Muto, S., Tsuboi, A., Bakker, E. P., Nakamura, T. & Schroeder, J. I. (2000) *Plant Physiol.* **122,** 1249–1260.
22. Kato, Y., Sakaguchi, M., Mori, Y., Saito, K., Nakamura, T., Bakker, E. P., Sato, Y., Goshima, S. & Uozumi, N. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 6488–6493.
23. Wu, S. J., Ding, L. & Zhu, J. K. (1996) *Plant Cell* **8,** 617–627.
24. Rus, A., Yokoi, S., Sharkhuu, A., Reddy, M., Lee, B.-H., Matsumoto, T. K., Koiwa, H., Zhu, J.-K., Bressan, R. A. & Hasegawa, P. M. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 14150–14155.
25. Shi, H., Quintero, F. J., Pardo, J. M. & Zhu, J.-K. (2002) *Plant Cell* **14,** 465–477.
26. Mäser, P., Eckelman, B., Vaidyanathan, R., Horie, T., Fairbairn, D. J., Kubo, M., Yamagami, M., Yamaguchi, K., Nishimura, M., Uozumi, N., *et al.* (2002) *FEBS Lett.* **531,** 157–161.
27. Berthomieu, P., Conejero, G., Nublat, A., Brackenbury, W. J., Lambert, C., Savio, C., Uozumi, N., Oiki, S., Yamada, K., Cellier, F., Gosti, F., *et al.* (2003) *EMBO J.* **22,** 2004–2014.
28. Mitra, R. M., Gleason, C. A., Edwards, A., Hadfield, J., Downie, J. A., Oldroyd, G. E. D. & Long, S. R. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 4701–4705.