

NGS 기술 활용 돌연변이체 해석 및 연구현황

정유진 · 류호진 · 조용구 · 강권규

Current status and prospects to identify mutations responsible for mutant phenotypes by using NGS technology

Yu Jin Jung · Ho Jin Ryu · Yong-Gu Cho · Kwon Kyoo Kang

Received: 2 November 2016 / Revised: 4 November 2016 / Accepted: 4 November 2016
© Korean Society for Plant Biotechnology

Abstract Next-generation sequencing allows the identification of mutations responsible for mutant phenotypes by whole-genome resequencing and alignment to a reference genome. However, when the resequenced cultivar/line displays significant structural variation from the reference genome, mutations in the genome regions absent in the reference cannot be identified by simple alignment. In this review, we report the current status and prospects in identification of genes in mutant phenotypes, by using the methods MutMap, MutMap-Gap, and MutMap+. These methods delineate a candidate region harboring a mutation of interest, followed by *de novo* assembly, alignment, and identification of the mutation within genome gaps. These methods are likely to prove useful for cloning genes that exhibit significant structural variations, such as disease resistance genes of the nucleotide-binding site-leucine rich repeat (NBS-LRR) class.

Keywords Mutations, Next-generation sequencing, NGS, Genome wide association study, GWAS

서 언

유전학의 연구 목적 중 하나는 생물 종의 개체간 표현형 변이의 원인인 유전 변이를 탐색하는 것이다. 이 목적을 위해 유전적 연관(genetic association) 분석이 이용되어왔다. 즉, 생물 종의 집단을 표현형에 의해 2개 이상의 그룹으로 분류 할 때, 그 분류와 통계적으로 유의한 연관(association)을 나타내도록 계놈상의 변이를 찾는 분석이다. 계놈이 부모로부터 자손에게 유전될 때 물리적으로 서로 가까이 있는 DNA 영역은 더 자주 함께 후손에게 전달되기 때문에 강한 연관을 보인다(Yan et al. 2009). 한편, 서로 떨어진 위치에 있는 2개의 DNA 영역 사이에서는 영역 사이의 재조합에 의해 후손에게 전달이 독립되고, 연관이 약하다(Wallace et al. 2014). 따라서 표현형의 분류와 특정 계놈 변이의 연관성을 관찰함으로써 표현형을 결정하는 계놈 영역을 추정하는 것이 가능하다. 유전적 연관 분석 중에서 가장 많이 사용되는 방법이 연쇄 분석(linkage analysis)이다. 서로 다른 표현형을 나타내는 두 개체를 교배하여 자손을 얻은 개체 사이에서 주목하고 있는 표현형을 분리 할 때 분리하는 같은 계놈 영역을 찾는 데 매우 효과적인 방법으로 알려졌다(Zhu et al. 2008). 특히 전체 계놈 연관 분석 방법(Whole Genome Association Study; GWAS)은 집단에서 자연에 존재하는 개체나 계통을 대상으로 표현형으로 2군인 형질이 2군과 연관을 나타내는 변이 계놈 전체에서 검색하는 방법이다(Altshuler et al. 2008). GWAS에서는 인공적인 교배를 사용하지 않고, 생물 종의 진화 과정에서 자연스럽게 발생하여 누적

Y. J. Jung · K. K. Kang (✉)
국립한경대학교 원예생명과학과
(Department of Horticultural Life Science, Hankyong National University, Ansong, Gyeonggi-do 17579, Korea)
국립한경대학교 유전공학연구소
(Institute of Genetic Engineering, Hankyong National University, Ansong 17579, Korea)
e-mail: kykang@hknu.ac.kr

H. J. Ryu
충북대학교 자연과학대학 생물학과
(Department of Biology, Chungbuk National University, Cheongju 28644, Korea)

Y.-G. Cho
충북대학교 식물자원학과
(Department of Crop Science, Chungbuk National University, Cheongju 28644, Korea)

된 재조합의 연관 분석을 실시하여 농업적으로 중요한 형질 관련 유전자를 동정할 수 있다(Hunter and Crawford 2008). 계놈의 DNA 배열 해독(DNA 시퀀싱)은 1974년에 염기서열결정법의 발명 이후 30년 동안 주로 Sanger법이 이용되어왔다. 2005년 전후에 “차세대 시퀀싱(NGS)” 기술이 이용되기 시작했다. 454 기계 pyrosequencing에 이어 Solexa사의 Illumina 시퀀싱, ABI사의 SOLiD 시퀀싱이 가능하게 되었다. “차세대 시퀀싱의 발달로 생물의 전체 계놈 배열의 새 결정(de novo assembly)이 비약적으로 촉진되었다. 그 결과, 대략적인 전체 계놈 배열이 결정된 생물 종 수는 기하 급수적으로 증가하고 있다(Austin et al. 2011). De novo assembly는 주로 서열이 밝혀지지 않은 종의 계놈 서열을 알아내고자 할 때 수행하는 분석이며, 품종간 유전체 배열을 해독하고 reference sequence와 비교하는 작업을 re-sequencing 이라고 한다. Re-sequencing이 쉽고 간단하게 이루어지면, DNA 마커가 필요하지 않게 되며, 유전체 해석에 의해 선발되고 평가할 수 있기 때문이다. Re-sequencing을 통해 유전체 서열을 분석함으로써 표현형의 변이와 유전자 변이의 연관성을 해석하는 것이 가능하게 되었다(Altshuler et al. 2008; Hunter and Crawford, 2008). 이와 같이 GWAS은 이미 많은 생물에 적용되고 있으며, 교배 후 분리 집단에 있어서 re-sequencing으로 원인 유전자를 식별하는 기술이 개발되었다(Lupski et al. 2010).

본 리뷰에서는 유전체 계놈 서열 해독이 완료된 다양한 생물 종에서 차세대 시퀀서를 사용하여 신속하게 돌연변이 원인 유전자를 동정하는 기술에 대해 설명하고 고찰하고자 한다.

벼 돌연변이체에서 유발 SNPs의 동정을 위한 MutMap 적용

일본 Iwate 생물 공학 연구 센터에서는 Iwate 현에 적합한 벼 품종을 개발하기 위해서 Hitomebore의 수분수정 시 EMS 처리에 의해 육성한 15,000개의 돌연변이 계통과 Hitomebore와 20개 우수 품종과 교배 해 얻은 약 3,000개의 recombinant Inbred Lines (RILs) 계통 군을 육성하여 육종에 유용한 유전자 영역을 식별할 수 있는 DNA 마커 개발 및 선발에 따라 신속하게 품종을 육성하고 있다(Abe et al. 2012). 돌연변이 계통의 표현형의 원인 유전자를 전체 계놈 분석에 의해 신속하게 식별하는 기술로서 MutMap 법을 적용하여 보고하였다(Abe et al. 2012). 이 방법은 목적 표현형을 나타내는 돌연변이(hst1)와 돌연변이 처리에 이용한 부모 계통(Hitomebore)을 교배하여 얻어진 F1 개체를 자식하여 F2 개체를 100 개체 정도 얻었다. 변이 형질이 열성 돌연변이로 인한 경우 F2에서는 표현형이 야생형(wild type)과 돌연변이가 3 : 1의 비율로 분리되었다. F2 개체 중 변이형을 나타내는 개체 20 개체의 계

놈 DNA를 골고루 혼합(bulk화)하고, 차세대 시퀀서에 의해 전체 계놈의 염기배열을 분석하였다. 시퀀서는 75 ~ 100 bp의 짧은 염기서열 (short read)를 많이 얻을 수 있었으며, 이렇게 얻어진 short read를 Hitomebore 품종에서 미리 작성한 reference sequence에 맞게 정렬하였다. 이 때 계놈의 특정 부분을 몇 개의 short read가 있어 얼마나 많이 시퀀싱 할 것인지(depth)를 파악하여 SNP-index 값을 결정하였다. SNP-index는 정렬된 short read와 reference sequence 사이에 SNP가 있는 경우, 그 SNP 부위를 덮는 short read 전체 중 reference sequence와 다른 염기를 가진 short read의 비율이다. 예를 들어, depth 9 중 4 read가 기준 서열과 다른 염기를 보유하면 $SNP-index = 4 / 9 = 0.44$ 가 된다. 돌연변이 형질을 나타내는 F2 개체는 원인 SNP는 호모상태로 되어 있으며, 이는 돌연변이 F2 개체들의 DNA에서 얻어진 short read와 같은 SNP를 갖게 된다. 즉 원인 SNP는 SNP-index가 1을 나타내는 것으로 기대된다. 한편, 표현형과 관계없는 SNP에서 변이형 F2 bulk DNA에서 돌연변이 SNP를 포함한 short read와 부모형 (야생형) short read가 1 : 1로 기대되며, SNP-index는 0.5의 기대치를 갖는다. 또한 염색체에서 원인 SNP 근처에 있는 SNP는 연쇄에 의해 0.5와 1 사이의 SNP-index를 나타내는 것으로 예상된다. 따라서 돌연변이와 부모 계통 사이에 존재하는 모든 SNP(돌연변이 계통 군에서는 평균 약 2,000개)에 대한 SNP-index를 계산하였다. 내염성에 저항성을 보이는 돌연변이체의 원인 유전자를 동정하기 위해 MutMap 분석을 실시한 결과 가로축에 SNP의 염색체상의 위치, 세로축에 SNP-index의 값을 나타낸 그래프에서 SNP-index가 1이 되는 피크를 찾음으로써 내염성 돌연변이 유전자의 SNP를 식별할 수 있었다(Fig. 1). 염색체 6번상에서 SNP index의 2개의 피크를 찾아 이들 피크에 해당되는 영역이 내염성(hst1) 저항성에 관련이 있다고 생각할 수 있었다. 이들 영역의 부분을 상세히 조사한 결과 OsRR22 유전자 영역에서 TGG (Trp)가 TAG (Stop)로 변이가 일어나 저항성을 보였다(Abe et al. 2012).

따라서 MutMap 법은 돌연변이체와 돌연변이 처리에 이용한 부모 계통과 교배하여 육성한 F2 집단을 이용하여 내염성 저항성과 같은 양적 변이를 유발하는 유전자를 동정하는데 효율적으로 사용 가능할 것으로 생각한다(Takagi et al. 2015).

de novo 어셈블리를 이용한 MutMap-Gap 방법의 적용

주요 품종인 Hitomebore와 Koshihikari의 유전체 해독은 품종 특이적으로 같은 계놈 영역에서는 정확하게 Alignment 되지 않기 때문에 이러한 영역에 발생한 돌연변이는 검색할 수 없는 단점이 있다. 이런 단점을 보완하기 위해 Takagi 등(2013a)은 resequencing과 de novo assembly를 병용한 새

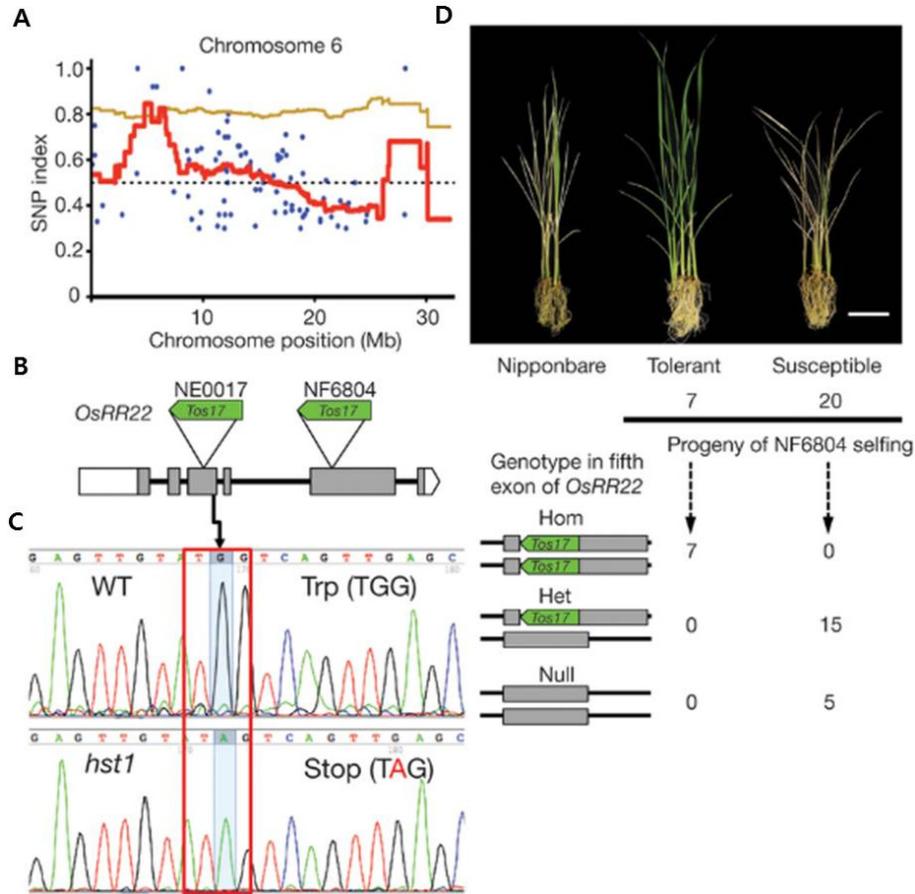


Fig. 1 Identification of the *hst1* mutation by MutMap. SNP-index plot of the rice chromosome 6 generated by MutMap analysis, showing a genomic region with the highest SNP-index peak harboring the candidate mutation. DNA bulked from 20 salt-tolerant F2 progeny obtained from the cross between WT Hitomebore and *hst1* was used for sequencing and MutMap analysis. Blue dots correspond to SNPs identified between *hst1* and Hitomebore WT genomes (Takagi et al. 2015)

로운 기법으로 MutMap-Gap 법을 확립하였다. Figure 2는 MutMap-Gap 법을 설명한 것으로 ① 돌연변이 원품종의 게놈을 시퀀싱하여 기준게놈 배열에 대해 정렬하고 (re-sequencing), 원 품종의 기준 배열을 생성한다. 이때 기준 게놈 배열에 정렬되지 않은 short read군을 mapping하여 paired-end reads로 따로 저장하였다. ② 목적으로 하는 돌연변이를 선발 후, 원품종의 기준 배열을 이용한 MutMap 법에 의해 원인 유전자 자리의 게놈 영역을 식별하였다. ③ MutMap법으로 특정된 후보 영역에 정렬된 short read 및 paired-end reads를 이용하여 de novo 어셈블리를 하고 contig를 구축하였다. 이 단계는 후보 영역의 게놈 배열을 de novo assembly에 따라 재구성하는 것을 의미한다. ④ 마지막으로, 구축된 contig에 대해 다시 MutMap 분석 한 후, 돌연변이 형질에 대한 원인 유전자 변이를 식별하였다. MutMap-Gap법을 적용함으로써 전체 게놈 서열 해독이 완료된 품종이면 해독에 이용한 품종과 다른 품종에서도 resequencing 기술에 의한 변이 분석이 가능하다. 이들 방법을 이용하여 도열병 저항성 Pi 유전자 기능이 손

실된 돌연변이 계통으로부터 Hitomebore 품종의 도열병 저항성 계통 육성에 이용하였다(Takagi et al. 2013a). 이와 같이 nucleotide-binding site-leucine rich repeat (NBS-LRR) 그룹들의 병 저항성 유전자와 같이 구조적 변이를 가진 유전자를 분리하는데 입증하였다.

MutMap+의 원리 및 육종에의 이용

교배 과정을 거치지 않고 돌연변이체의 원인유전자를 동정하기 위해 MutMap+ 법이 개발되었다(Abe et al. 2012; Fekih et al. 2013). 먼저 벼 수정 후 즉시 미성숙 배에 EMS를 처리한 다음 종자를 수확하여 M1 종자를 육성하였다. EMS 유래 돌연변이체들의 대부분은 이형접합체 상태의 M1 식물이다(Fig. 3A). 비록 M1 세대의 돌연변이 식물들에서 dominant와 semi-dominant는 구분되지만, 열성돌연변이는 구분이 되지 않아서, M2 세대를 진전시켜야 구별 가능하다. M2 자식세대에서 야생형과 돌연변이 표현형의 분리비가 3:1로 분리되는 것을 각각 선발하였다(Fig. 3 B).

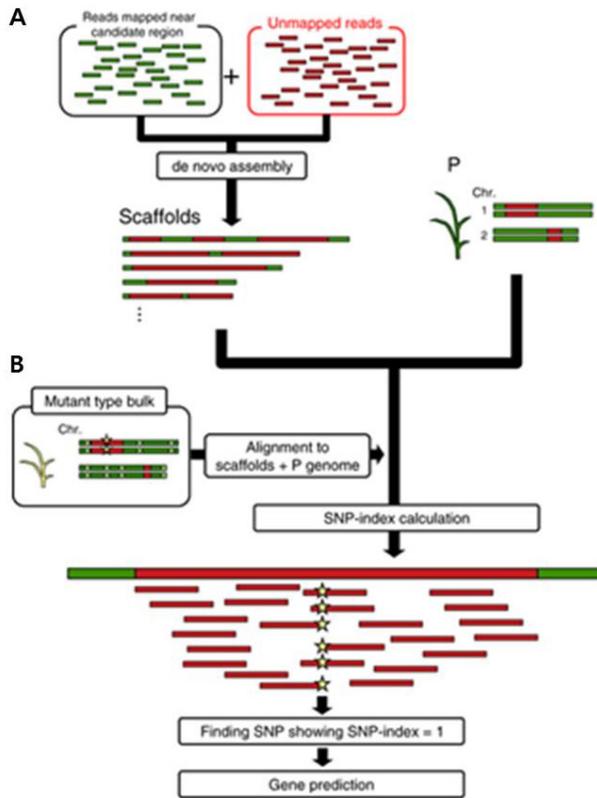


Fig. 2 MutMap-Gap fills the gap within a target genome region delineated by MutMap using *de novo* assembly. (a) Using a combination of the unmapped reads collected in the previous step (Fig. a) and the short reads aligned to the target region (Fig. c), *de novo* assembly is performed to reconstruct the sequence of the target interval. The scaffolds generated by the *de novo* assembly are combined with the P ‘reference sequence’ and serve as the ‘P+scaffolds’ reference sequence for alignment purposes in the following step. (b) Short reads derived from the bulk DNA of mutant F2 progeny are aligned to the ‘P+scaffolds’ reference. The scaffolds with single nucleotide polymorphisms (SNPs) showing an SNP index of 1, are likely to contain the causal nucleotide change responsible for the mutant phenotype. Chr., chromosome (TaKagi et al. 2013)

MutMap + 실험을 위해 포장에 M2 세대를 계통별로 10개 체 정도 파종하고 표현형의 분리를 개체 별로 체크하였다(Fig. 3B). 이들 중 야생형 중에서 heterozygote로 구성된 개체를 자식하여 M3 세대를 육성하였다. 이들 M3세대의 종자는 개체별로 채종하고, 80 개체 이상 포장에 심어 표현형의 분리를 조사하였다. 이렇게 얻어진 M3 식물들의 돌연변이형 bulk(20~40개체)와 야생형 bulk(20~40개체)를 구분하여 두 집단으로부터 DNA를 추출하였다(Fig. 3C). Bulk DNA는 각각 Illumina GAIIX sequencer에 의해 염기배열을 분석한 후, 양친 계통의 reference sequence와 비교 분석하였다. 각각의 bulk로부터 SNP-index와 SNP genomic position graphs를 그렸다(Fig. 3D). 돌연변이체로부터 얻어진 DNA pool (Fig. 3D)에 대해 SNP index plot에서 SNP index=1을 가진 유전체 영역의 발생에는 두 가지 이유가

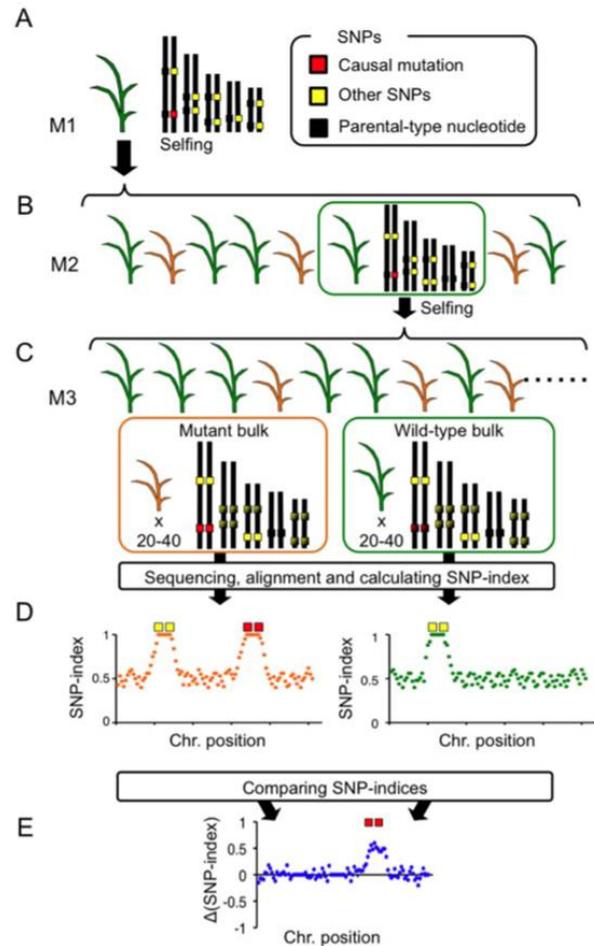


Fig. 3 A simplified scheme of MutMap+. (A) Seeds harvested following EMS mutagenesis of rice at immature embryo stage are used to establish the M1 generation, at which stage most of the mutations incorporated by EMS are in the heterozygous state. (B) M2 progeny obtained from a self-fertilized M1 plant segregate for wild-type (indicated by green color) and mutant (brown color) phenotypes. Here, we focus on the wild-type heterozygous individuals. (C) Heterozygous M2 plant are self-fertilized to obtain M3 progeny that segregate 3:1 for wild-type and mutant phenotypes. Genomic DNA from 20–40 M3 mutant and wild-type M3 progeny are separately bulked, and subjected to whole-genome sequencing. The resulting short reads are aligned to reference sequence of the cultivar used for mutagenesis. (D) SNP-index is calculated for each SNP, and plots relating the SNP-index and chromosome positions are obtained for both the mutant and wild-type M3 bulks separately. The two SNP-index plots are compared to identify the region with SNP-index=1 that is specific to the mutant bulk. (E) We can also evaluate D (SNP-index) plot, which is obtained by subtracting SNP-index value of wild-type bulk from that of mutant bulk. Genomic region harboring the causal mutation should have positive D (SNP-index) values (Fekih et al. 2013)

있다. 첫 번째는 실제로 이들 영역이 표현형의 원인돌연변이를 유발한 경우이고, 또 하나는 표현형과 무관한 SNP가 M2 세대에서 동형 접합체 상태로 고정되어 이들 모두가 M3 식물체에 나타나기 때문이다. 이것은 전체 계

놈 영역의 50%에서 일어날 것으로 예상되며, 돌연변이 bulk의 SNP index plot 에서 SNP를 살펴볼 때, 두 가지 유형을 구별하는 것은 불가능하다(Fekih et al. 2013). 그러나, 야생형과 돌연변이 bulk의 SNP index plot을 비교하여 임의로 동형 접합체 상태에 고정되어 SNP index = 1 가진 SNPs는 식별 할 수가 있다. M2 세대에서의 SNP의 임의의 고정에 의해 SNP-index= 1을 보여주는 영역들은 두 bulk 집단 사이에 공유되어야 하며, 반면에 유발 돌연변이를 가지는 영역은 돌연변이 bulk 집단에서만 특이적으로 나타난다(Fig. 3D). 이 데이터를 시각화하기 위해, Δ (SNP index) (Fig. 3E)를 얻기 위해 각각의 SNP 변이체 bulk와 야생형 대량의 SNP index를 빼서 Δ (SNP index) 값으로 표시했는데 대부분의 게놈에서는 값이 0 주변에 있지만, 돌연변이를 유발하는 영역에서는 상당히 높은 값으로 표시될 수 있다(Fekih et al. 2013; Boetzer et al. 2011; Clark et al. 2007; Li and Durbin 2009; Li et al. 2009). 이와 같이 NGS 기술의 발달로 인해 여러 종류의 생물 게놈의 sequencing 데이터가 방대한 양이 축적되었으며, 단시간에 유전체 분석이 저비용으로 가능할 수 있는 NGS platforms 이 개발되고 있다(Zhou et al. 2014; Lister et al. 2009). 최근의 연구에서는 강력한 NGS platforms을 이용하여 개체를 인식할 수 있는 GBS (genotype by sequencing), 품종을 구별할 수 있는 DNA marker 개발과 어떤 집단을 해석하는 QTL 분석 등에 이용되고 있다(Morishige et al. 2013). 최근 들어 NGS 기술은 GWAS, restriction-site-associated DNA (RAD) sequencing 및 MutMap 등에 광범위하게 활용되고 있다(Abe et al. 2012; Baxter et al. 2011; Elshire et al. 2011; Fekih et al. 2013; Harper et al. 2012; Huang et al. 2010; Morishige et al. 2013; Takagi et al. 2013b; Tian et al. 2011).

본 총설에서는 돌연변이의 원인 유전자 동정을 위해 MutMap법, MutMap-Gap법 및 MutMap+법 등에 대한 유용성을 살펴보았다. 본 논문의 저자들은 인위적으로 유발하는 돌연변이와 비교하여 긴 세월을 걸쳐 그 생물이 축적해 온 자연변이는 자연 및 인위적 선택을 받아 선발된 돌연변이도 포함되기 때문에 유용한 allele을 많아 관련 유전자들의 분리 동정에 사용될 것으로 기대한다. 또한 MutMap 방법들이 QTL-seq법과 함께 육종현장에서 신속하게 식별할 수 있는 수법으로 더 발전되길 기대해 본다

적 요

NGS 기술은 전체 게놈 시퀀싱 및 reference 게놈에 alignment에 의해 돌연변이 표현형에 관련된 돌연변이 식별에 이용한다. 그러나 품종 및 계통들을 resequence 하였을 경우 기존의 reference 게놈에 구조적 변이가 보이며, reference와 맞지 않는 게놈지역에서 돌연변이들은 단순한 alignment

로 찾을 수 없다. 본 리뷰에서는 NGS 기술을 이용하여 돌연변이체로부터 변이 관련 유전자를 식별하는 MutMap, MutMap-Gap 및 MutMap+ 방법을 기술하였고 지금까지의 연구현황에 대해 기술하였다. 아울러 이들 방법은 nucleotide-binding site-leucine rich repeat (NBS-LRR) 그룹들의 병 저항성 유전자와 같이 구조적 변이를 가진 유전자를 분리하는 등 유용성에 대해 고찰하였다.

사 사

본 성과물은 농촌진흥청 차세대바이오그린21, GM작물 개발사업단 연구사업(세부과제번호: PJ01185201)의 지원에 의해 이루어진 것임.

References

- Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, oshida K, Mitsuoka C, Tamiru M et al. (2012) Genome sequencing reveals genomically important loci in rice using MutMap. *Nature Biotechnology* 30:174-178
- Altshule D, Daly MJ, Lander ES (2008) Genetic mapping in human disease. *Science* 322 (5903):881-888
- Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW et al. (2011) Next-generation mapping of Arabidopsis genes. *Plant Journal* 67:715-725
- Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD, Blaxter ML (2011) Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS One* 6: e19315
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W (2011) Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578-579
- Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P, Warthmann N, Hu TT, Fu G, Hinds DA et al. (2007) Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* 317:338-342
- Elshir RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379
- Fekih R, Takagi H, Tamiru M, Abe A, Natsume S, Yaegashi H, Sharma S, Sharma S, Kanzaki H, Matsumura H et al. (2013) MutMap+: genetic mapping and mutant identification without crossing in rice. *PLoS One* 8:e68529
- Harper AL, Trick M, Higgins J, Fraser F, Clissold L, Wells R, Hattori C, Werner P, Bancroft I (2012) Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nature Biotechnol.* 30:798-802
- Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C,

- Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang QF, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genet.* 42:961-967
- Hunter KW, Crawford NPS (2008) The future of mouse QTL mapping to diagnose disease in mice in the age of whole-genome association studies. *Annu. Rev. Genet.* 42:131-141
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and samtools. *Bioinformatics* 25:2078-2079
- Lister R, Gregory BD, Ecker JR (2009) Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Curr. Opin. Plant Biol.* 12:107-118
- Morishige DT, Klein PE, Hilley JL, Sahraeian SME, Sharma A, Mullet JE (2013) Digital genotyping of sorghum—a diverse plant species with a large repeat-rich genome. *BMC Genomics* 14:448
- Lupski JR, Reid JG, Gonzaga-Jauregui C, Deiros DR, Chen DCY, Nazareth L, Bainbridge M, Dinh H, Jing C, Wheeler DA, Mcguire AL, Zhang F, Stankiewicz P, Halperin JJ, Yang C, Gehman C, Guo D, Irikat RK, Tom W, Fantin NJ, Muzny DM, Gibbs RA (2010) Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N. Engl. J. Med.* 362:1181-1191
- Takagi H, Uemura A, Yaegashi H, Tamiru M, Abe A, Mitsuoka C, Utsushi H, Natsume S, Kanzaki H, Matsumura H et al. (2013a) MutMap-Gap: whole-genome resequencing of mutant F2 progeny bulk combined with de novo assembly of gap regions identifies the rice blast resistance gene Pii. *New Phytol.* 200:276-283
- Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S et al. (2013b) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* 74:174-183
- Takagi H et al. (2015) MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nat. Biotechnol.* 33:445-449
- Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genet.* 43:6-11
- Wallace JG, Larsson SJ, Buckler ES (2014) Entering the second century of maize quantitative genetics. *Heredity* 112:30-38
- Yan J, Shah T, Warburton ML, Buckler ES, McMullen MD, Crouch J (2009) Genetic Characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *Plos One* 4(12):e8451
- Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, Shen J, Tian D (2004) Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Molecular Genetics and Genomics* 271:402-405
- Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. *Plant Genome* 1:5-20