# Exom capture assay to detect SNPs in wheat mutated population through EMS

**Momina Hussain,Atif Iqbal,Bradley J Till,Mehboob ur Rahman**

## Abstract

Wheat—a staple food crop of many countries, has always been a target for improving its resilience to biotic and abiotic stresses. Among these, rust diseases are the most detrimental in depressing wheat production. In the present study, chemical mutagen was used to induce mutations in a wheat variety NN-Gandum-1 (moderately resistant to leaf and yellow rust) for improving resistance to the disease as well as studying function of genes conferring resistance to the disease. In the present investigation, 0.8% EMS dose was found optimum for supporting 45-55% germination of NN-Gandum-1. A total of 3,634 $M_2$ fertile plants were produced from each of the $M_1$ plant. Out of these, 33 (0.91 %) and 20 plants (0.55%) showed absolute resistance to leaf and yellow rust, respectively. While 126 (3.46%) and 127 plants (3.49%) exhibited high susceptibility to the leaf and yellow rust, respectively. In $M_4$ generation, a total of 11 $M_4$ lines (nine absolute resistant and two highly susceptible) and one wild type were selected for NGS-based exome capture assay. A total of 104,779 SNPs were identified that were randomly distributed throughout the wheat sub genomes (A, B and D). Maximum number of mutations were found in introns. Highest number of SNPs were detected in chr.2B (14,273 SNPs). Mutation density was one mutation per 20.91 Mb. Highest mutation frequency was found in chr.7D (1/2.83 Mb) and the lowest in chr.2D (1/85.17 Mb). Out of the detected mutations, 101 SNPs were selected using priority analysis criteria. Out of these, SNPs detected in *Lr21* were selected for further analysis. The SNP identified in chimeric allele (*Lr21*) of a resistant mutant (N1-252) was located in NBS domain of chr.1BS at 3.4 Mb that possibly activated the function of chimeric allele. Through computational analysis, it was demonstrated that identified SNP substituted the glutamic acid with alanine—resulted in altered protein structure. It was concluded that induced mutations are extremely helpful not only in generating novel genetic resource but can also exploited for getting insight into the important biological circuits of different traits of complex genomes like wheat.

**Citation:** Momina Hussain,Atif Iqbal,Bradley J Till,Mehboob ur Rahman Exom capture assay to detect SNPs in wheat mutated population through EMS. **protocols.io**

dx.doi.org/10.17504/protocols.io.pj5dkq6

**Published:** 19 Apr 2018

## Protocol

### Step 1.

### Plant Material

The hexaploid ‘NN-Gandum-1’ (*Triticum aestivum* L) is a spring wheat variety that was bred by making a cross (Chirya-3/Opata//2x parula/3/ Rohtas-90) at the PGMB Labs, NIBGE, Faisalabad Pakistan. ‘NN-Gandum-1’ (NN-1) has been evaluated for studying its adaptability throughout the

Punjab province in Pakistan. Because of its high yield potential and limited tolerance to rust diseases (moderately resistant) especially leaf rust, this variety was selected for exposing to a mutagen. In this regard, seed were exposed to chemical mutagen ethyl methane sulphonate (EMS).

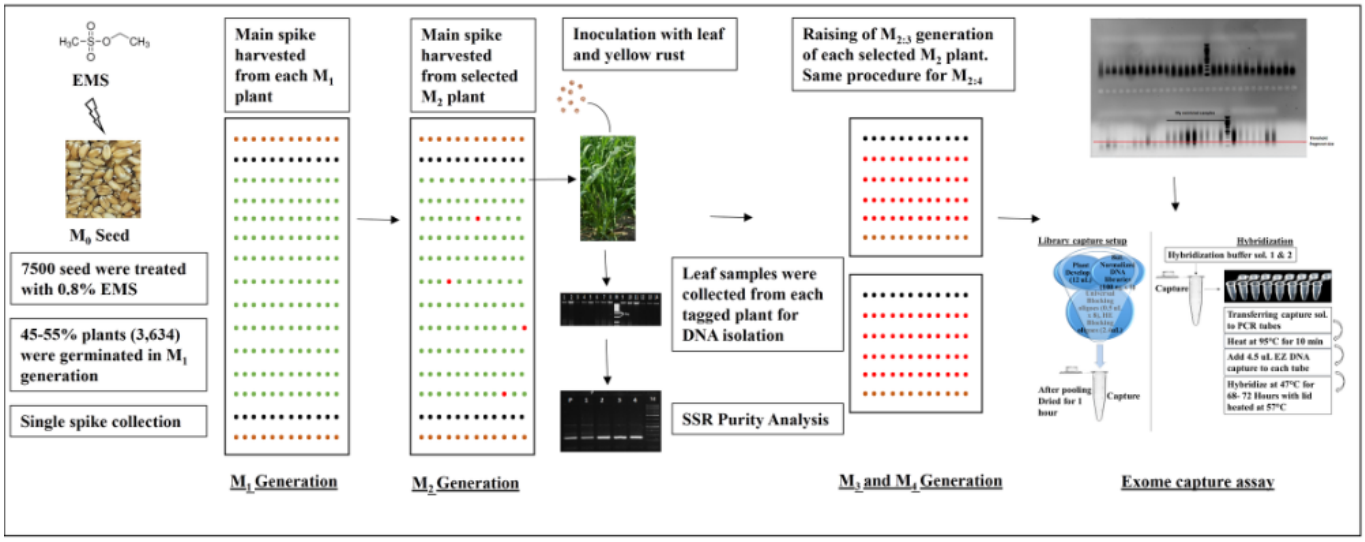
**Step 2.**

**Mutagenesis and mutant population development**

To optimize the dose of EMS, 24 batches of 50 seed each were treated with different EMS concentrations (ranging from 0.4 – 0.9 %); each concentration was also tested for two different time exposure and temperatures i.e. 1 and 2 hours and 33°C and 35°C, respectively. After treatment, seed were washed under running water for 3 hours to remove complete residuals of EMS. Before the treatment, seed were sterilized with 50% and 70% sodium hypochlorite and ethanol, respectively, each with three washings. On the basis of germination percentage (45-55%), 0.8% EMS for 2 hour at 35°C was chosen to mutagenize a batch of 7,500 wheat seed.

After an overnight drying, $M_1$ seed were manually sown in experimental field in NIBGE (Faisalabad, Pakistan) with a cultivation distance of 30 cm between each row. Seed were sown in the form of beds, each bed contains 100 rows and each row contains eight plants. Approx. 3,700 $M_1$ (out of 7,500 seed) plants were germinated. At physiological maturity, the main spike of each $M_1$ plant was bagged to ensure self-pollination. At maturity, the $M_2$ seed from each main spike was harvested to sow $M_2$ generation. It is worth mentioning that standard agronomic practices were applied from sowing till harvesting in each generation.  Eight $M_2$ plants (of each $M_1$ plant) in each row were sown in bed and were thinned to leave a single $M_2$ plant per row. At key development stage, from germination to maturity, based on a visual characterization of plant, a systemic phenotyping scoring of mutant population was carried out. Response to rust diseases (yellow and brown rust) was recorded by adopting universally used rating scale (Peterson et al., 1948; Loegering, 1959). The coefficient of infection (CI) was calculated by multiplying the disease response value with the disease infection intensity. Average coefficient (ACI) was also worked out by summing CI values of each genotype divided by the total number of locations. Leaf tissues were collected from each labeled plant per row for DNA isolation (CTAB method) for conducting SSR analysis (DNA uniformity analysis). The plant showing natural variations were rejected. Generation was advanced to $M_3$ by adopting the same procedure as did for $M_2$. The main selfed spike was sown to develop $M_4$ population. Data for morphological traits as well as response to rust diseases were recorded. Out of these, 11 mutant lines showing strong resistance (nine mutant lines) and high susceptibility (two mutant lines) were selected for exome capture assay (Fig 1).

**Fig 1. Development procedure of mutant population.** Brown color dots represent the genotype ‘Morocco’, Black color dot represent ‘NN-Gandum-1’ (wild type), Green dots represent mutant plants, Red represents the selected mutant plants




**Step 3.**

**SSR uniformity analysis**

In total, 20 SSRs were selected based on their high PIC values—calculated by surveying a set of wheat genotypes (96) with 650 SSR (unpublished results), and were surveyed on $M_2$ plants. Out of these, WMS-46, WMS-249 and WMS-311 amplified polymorphic alleles.

Preparation of wheat exome capture libraries:

**Step 4.**

**Isolation of genomic DNA and initial preparation for exome capture**

Total genomic DNA was extracted from young seedlings of $M_4$ plants of 12 samples (11 mutants and one wild) by adopting large-scale extraction protocol (32) with little modifications. Initially, the nuclei were purified followed by treatment with proteinase K and purification with phenol-chloroform. The extracted genomic DNA concentration was normalized to a final concentration of 200 ng/µL by adding a 0.1 X TE buffer (0.1 mM EDTA, 10mM Tris-HCl, pH 8) (32). In total 130 µl final volume of each sample was made.

The genomic DNA from each sample was chopped down to an average size of 250-450 bp with the help of a Covaris E220 machine (Duty cycle = 20 %, intensity = 175 W, cycles per burst = 200 and time = 90 sec). Ends of the resultant fragments were repaired by adding 2.5 µL end repair enzyme followed by the ligation of A-tails and adaptors using KAPA's kit and Bioo Adapters (Sciclone G3 robot). Genomic libraries for conducting exome capture experiment were made using high-throughput library preparation kits from KAPA Biosystems, Inc. (Wilmington, MA, USA, catalog number KK2612). This procedure was described by Kraileva and co-workers (12).

**Step 5.**

**Capture setup**

In total, 1.2 µg of genomic DNA libraries were made by adding 150 ng of each of the eight samples (150 ng x 8). In this tube, 12 µL Plant 'Developer' reagent from Nimblegen (catalog number 6684335001) was added. In this solution, 0.5 µL barcode HE blocking oligoes (Bioo, catalog number

520999) and 2.4 µL HE universal blocking oligo (Bioo) were also added. After pooling all these components, the capture was dried (drying of DNA) in a vacuum centrifuge with heat for one hour. To facilitate this process, a hole was made in the lid of the reaction tube. After drying, the hole was sealed with a piece of tape. Now the capture is ready to proceed further (12).

**Step 6.**

**DNA preparation for hybridization**

After capture preparation, each capture was dissolved in 7.5 µL hybridization solution-5 and 3 µL hybridization solution-6 (Nimblegen hybridization kit (5634253001, Roche)). It was vortexed vigorously for 10 seconds and centrifuged shortly to collect the solution. The solution was transferred to a pre-labeled 0.2 µL PCR tubes. These tubes were heated at 95°C for 10 minutes in a thermo cycler. Then a total of 4.5 µL of Nimblegen EZ DNA capture (custom, Roche) was added to each tube. The solution was mixed thoroughly using vortex followed by short spinning to collect the solution (12).

**Step 7.**

**Hybridization**

This step was performed in a thermal cycler for 68-72 hours at 47°C while the lid temperature was kept at 57°C. Captured hybrids were washed using three washing buffer (wash buffer I, II and III) and one bead wash buffer (12).

**Step 8.**

**PCR amplification and final clean-up**

Captures were amplified by PCR using a primer pairs, F: AATGATACGGCGACCACCGATCTACAC and R: CAAGCAGAAGACGGCATACGAGAT (custom synthesized by Sigma). Master stock solution was prepared with 10 µM Tris (pH 8.0) at 100 µM, and the working solution was 10 µM (1:10 dilution of master stock) with 1:1 primer mix (5 uM of each primer). The PCR cocktail for two reactions was prepared by adding 20 µL captured DNA, 25 µL 2x KAPA mix and 5 µL primer stock (10 µM). PCR amplification was conducted in duplicates using KAPA's amplification kit, and programmed for one cycle for 45 sec at 98 ºC followed by 10 cycles each for 15 sec at 98 ºC, 30 sec at 60 ºC and 30 sec at 72 ºC followed by a final extension cycle for 1 min at 72 ºC. Library was left on beads. Remaining 10 µL of the captured beads was kept for quality control checks at -20°C (Henry et al. 2014; (12, 25). The PCR products were purified by extensive washing (five times) using 80% ethanol (Sigma, catalog number ET0005) by adopting a published protocol (33).

**Step 9.**

**DNA quantification**

The enrichment of the targeted exons was tested using qRT-PCR. In this regard, the purified captured DNA was quantified, and primers designed for the amplification of two wheat housekeeping marker genes (F: ATCGGATTCGACAACATGC, R: ATATGGCCTGTCGTGAGTGA for Nuclear-encoded Rubisco and F: AAAGGCGTCAAGATGGAGTT, R: GGAATCCACCAACCATAACC for Malate Dehydrogenase) were custom synthesized by Sigma (12).

**Step 10.**

**Illumina Sequencing**

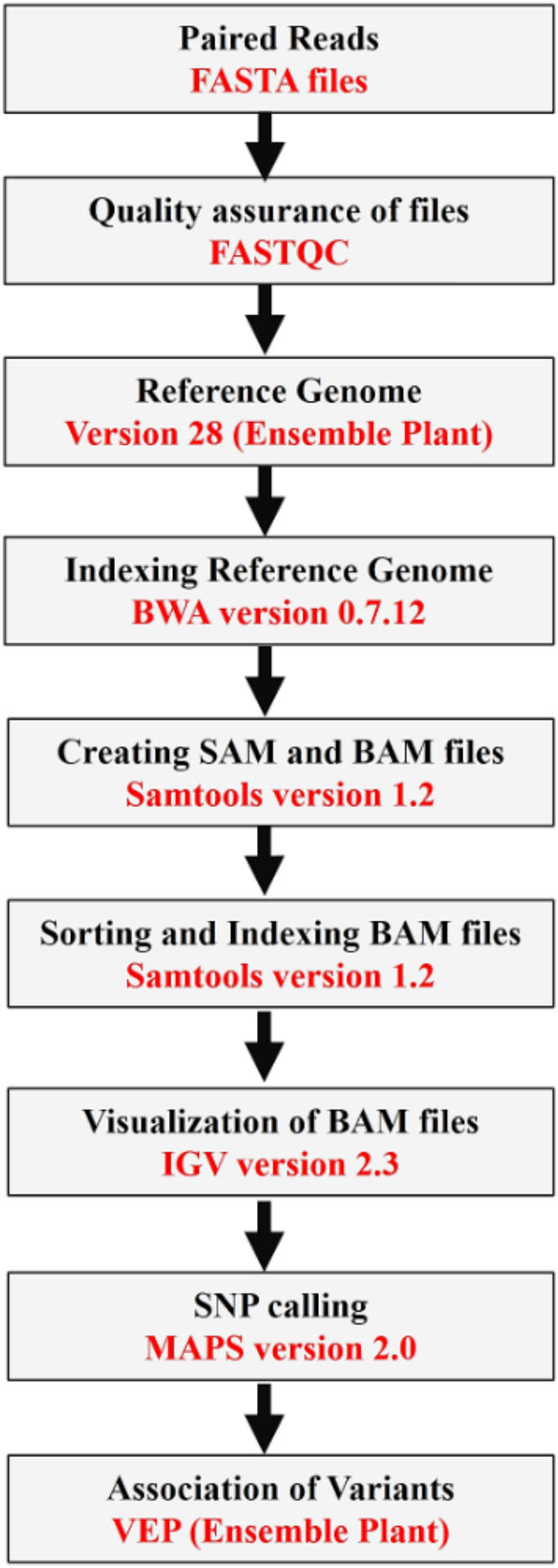
The final captures were submitted to Beijing Genome Institute (BGI) for sequencing with Illumina Hisequation 2000 (Illumina).

**Alignment of exome sequencing reads against wheat draft genome**

Bioinformatics analysis of sequence of each fragment was undertaken from start to stop codon. Illumina generated-sequences were processed using Fastq software for testing the quality of reads. The generated reads were assembled and aligned against wheat reference genome using ‘bwa aln’ and ‘bwa sampe’ programs (34). Then samtools and bamtools were used to generate sequence alignment map (SAM) and binary version of SAM files called as BAM file (35) (Fig 2). In total, 104,779 mutations (heterozygous and homozygous) were identified in all 11 mutant samples (N1-32, N1-61, N1-127, N1-236, N1-252, N1-506, N1-700, N1-701, N1-827, N1-910 and N1-1621) when aligned against the wheat draft genome (TGACv1) (http://plants.ensembl.org/Triticum_aestivum/Info/Index).

**Fig 2. Schematic diagram showing various steps deployed for the identification of**

**mutations**



**Step 11.**

**Calling SNPs**

Mutation and polymorphism survey (MAPS) (http://comailab.genomecenter.ucdavis.edu/index.php/MAPS) and 'mpileup' pipeline (http://comailab.genomecenter.ucdavis.edu/index.php/Mpileup) were deployed to select bases in the reference covered by at least one read at quality > 20 in a minimum number of samples. To differentiate the real SNPs from false mutations and or sequencing errors, an additional MAPS parameter was used. The threshold was setup independently for homozygous and heterozygous (12).

**Step 12.**

**Effect of SNPs on gene**

Mutations affecting gene function were identified by running a Variant Effect Predictor (VEP) from Ensembl tools release 78 in offline mode (36). This software estimates the impact of SNPs including deletions, insertions, or structural variants on various genes, transcripts, and protein sequence, as well as on regulatory regions. The effects predicted by VEP were extracted and counted for determining the number of genes interrupted by stop/splice or missense mutations. Then mutations affecting each gene were also counted. Sometime, mutation affected more than one gene, and then both effects were considered. If the same mutation found in more than one mutant line then it was counted multiple times. When this experiment was done, wheat genome annotation was in its initial stages of development, thus few mutations identified by VEP as 'intergenic' would be in genes which were incomplete or yet to be annotated (37).

**Step 13.**