

Descoberta de SNPs no pool gênico Mesoamericano através de re-sequenciamento genômico da cultivar Pérola

Guilherme Graciano Levindo¹, Rosana Pereira Vianello²

O cultivo de feijão do grupo comercial carioca representa quase 70% da área nacional plantada. Dentre as cultivares recomendadas, a BRS Pérola, lançada há 18 anos, ainda figura como uma das mais plantada no Brasil, resultante de seus atributos de qualidade agrônômica, tais como produtividade, resistência a enfermidades e tipo de grão. O genoma recém-sequenciado da espécie *Phaseolus vulgaris* abre novas perspectivas para a realização de experimentos de genômica e genética molecular de feijão. Esse estudo teve como objetivo analisar as informações de genoma da cultivar BRS Pérola geradas através do re-sequenciamento, seguido pela identificação de variações nucleotídicas através do alinhamento com o genomas de outras leguminosas. O re-sequenciamento foi baseado em tecnologia Illumina GAI. As sequências foram submetidas à filtragem para remoção do genoma de organelas. Utilizando o programa Bowtie2 foram alinhados os genomas de Pérola com os das variedades Mesoamericana (BAT93) e Andina (G19833) de feijoeiro comum. Os alinhamentos foram submetidos à identificação de polimorfismos do tipo SNPs utilizando o SAMtools e tendo como critérios o *Phred mínimo* de 200pb com qualidade > 20, cobertura mínima de 10x e uma região com cinco bases adjacentes sem *gaps*. O re-sequenciamento resultou em 114,6 milhões de *reads* com tamanho médio de 76 pb, equivalente a 14,5X de cobertura do genoma do feijoeiro comum. O conteúdo de GC foi estimado em 36%. Um total de 12% dos *reads* alinhou com genoma de organelas e foi removido da análise. Ao todo, 81% dos *reads* se alinharam ao genoma da variedade Mesoamericana, a partir do qual foram identificados 376.409 SNPs. Conforme o esperado, uma menor cobertura foi observada no alinhamento com a variedade Andina, equivalente a 76% dos *reads*, porém resultando em um número significativo maior de SNPs (1.207.964). Do total de SNPs identificados, 48% foram localizados em regiões não repetitivas do genoma, 7% em transcritos primários e 4% em regiões codantes. Os resultados obtidos comprovaram grande potencial na detecção e seleção de um grande número de SNPs a partir de re-sequenciamento via tecnologia NGS, principalmente no pool gênico Mesoamericano, onde se encontra representado a maior parte do material genético melhorado do Brasil.

¹ Estudante de graduação em Biotecnologia da Universidade Federal de Goiás, Bolsista PIBIC da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, graciano.gui@gmail.com

² Bióloga, Ph.D. em Genética Molecular, pesquisadora da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO, rosana.vianello@embrapa.br