

Estrategia computacional para hacer predicciones de características de interés agronómico en papas nativas diploides y tetraploides, utilizando información genómica

Luis Garreta

Corporación colombiana de investigación agropecuaria AGROSAVIA

El propósito del proyecto fué estudiar estrategias bioinformáticas que aceleren los programas de mejoramiento genético en plantas, específicamente sobre papas, con el fin de identificar variantes en la secuencia de ADN asociadas a rasgos de interés agrónómico. Las variantes que se utilizaron fueron los polimorfismos de nucleótido simple (SNPs) que afectan a una sola base de la secuencia genómica, y las estrategias bioinformáticas fueron los Estudios de Asociación de Genoma Amplio (GWAS) y la de Selección Genómica (SG). Los GWAS comprenden pruebas estadísticas que identifican qué variantes en todo el genoma de un gran número de individuos se asocian con un rasgo específico, mientras que la SG tiene como objetivo crear modelos para analizar y asociar una gran cantidad de información genómica y fenotípica para predecir el desempeño de rasgos de interés agronómico en individuos ya sea de poblaciones de mejoramiento genético (generalmente progenies) o de los bancos de germoplasma, analizando las accesiones que componen esa colección.

Se desarrollaron dos herramientas de software para realizar GWAS y SG en organismos diploides y tetraploides y se dejó establecida una estrategia de mejoramiento que combina los resultados de GWAS para realizar SG. La herramienta de GWAS ejecuta e integra 4 software para realizar GWAS, dos para poliploides (GWASpoly y SHEsis) y dos para diploides (PLINK y TASSEL). La herramienta es fácil de usar, parametrizable, y tiene tanto interface de línea de comando como interface gráfica.

La segunda herramienta permite realizar análisis de SG a través de la ejecución de la implementación de un flujo de trabajo o pipeline bioinformático donde se pueden comparar hasta 14 algoritmos usados en SG. La herramienta de SG está basada en la librería de R BGWS [1], es simple de usar y configurar, y actualmente utiliza una interface en línea de comandos.

Finalmente la estrategia de mejoramiento toma los SNPs de las mejores asociaciones mostradas por nuestra herramienta de GWAS los cuales se toman como entrada para nuestra herramienta de SG, la cual dá como resultado un análisis de los algoritmos tanto en poder predictivo como costo computacional. De acuerdo a estos valores, la herramienta de SG selecciona el mejor algoritmo y lo aplica sobre los datos de una población objetivo para estimar su valor genético.

Referencias

[1] Charmet, G., Tran, L.-G., Auzanneau, J., Rincet, R., & Bouchet, S. (2020). BWGS: A R package for genomic selection and its application to a wheat breeding programme. PLOS ONE, 15(4), e0222733.