

Nom et Prénom : NYIRAGUHIRWA SOLANGE

Date de soutenance : 13/07/2022

Directeur de Thèse : IBRIZ MOHAMMED

Sujet de Thèse :

**GENETIC DIVERSITY AND GENOME - WIDE ASSOCIATION STUDY OF IMPORTANT TRAITS
IN BARLEY (*Hordeum vulgare* L.)**

Abstract :

The cultivated barley (*Hordeum vulgare* L.) ranks the fourth most important cereal worldwide. It is used to feed animals, produce malt, and the human diet. Yield increase and yield stability are the top barley breeding goal. However, diseases such as the Net form of Net blotch (NFB) and powdery mildew (PM) reduce yield and grain quality. For barley destined for human consumption, an increase in micronutrients in grains, especially zinc and iron, is essential to alleviate malnutrition. Thus, breeders need to select new alleles and use them to develop higher-yielding, nutritious, and disease-resistant cultivars. Within the overall objective of enhancing barley productivity and providing valuable information to the barley breeders community, our study is divided into two parts. The first part of this study evaluated genetic diversity and population structure in 113 barley accessions predominantly from North Africa using 14 simple sequence repeat (SSR) markers. These accessions showed a higher degree of genetic variation. They were clustered into three subpopulations according to their geographic origin. The results of this study are crucial for conservation and their utilization in plant breeding. The second part of this study aims at locating genomic regions associated with important barley traits such as yield and yield components trait; disease resistance to NFB and PM, and micronutrients (Zn and Fe) content in grains via genome-wide association study (GWAS). A diverse collection of 496 barley was assembled and phenotyped for these traits in a multilocation trial. The collection was further genotyped with a 50K iselect SNP array of Illumina. Association mapping was conducted in TASSEL software using a mixed linear (MLM) model. Bonferroni correction ($p < 0.005$) was applied to reduce false SNPs resulting in 241 SNP associated with all studied traits. Most of these SNP were identified in previous studies, and some are colocalized with known genes that affect these traits. SNPs were annotated for putative candidate genes and showed that most of them correspond to proteins associated with the corresponding traits. Once the significant SNP identified in this study are validated in other experiments or populations, they can be converted into user-friendly markers and used to develop barley cultivars with improved yield, quality, and disease resistance via marker-assisted selection.

Keywords: barley, GWAS, SNP, disease resistance, QTL, yield, micronutrient

Résumé:

L'orge cultivée (*Hordeum vulgare* L.) se classe au quatrième rang mondial des céréales. Ses utilisations comprennent l'alimentation du bétail, le maltage et l'alimentation humaine. L'objectif de la sélection d'orge est de développer des variétés à haut rendement, nutritives et tolérantes au stress. Ainsi, les sélectionneurs doivent identifier de nouveaux gènes et les utiliser pour développer des variétés améliorées. Dans la première partie de cette étude, nous avons évalué la diversité génétique et la structure de la population de 113 accessions d'orge provenant principalement d'Afrique du Nord à l'aide de 14 marqueurs de répétition de séquence simple (SSR). Ces accessions ont montré un degré plus élevé de variation génétique. Structure de la population, analyse en composantes principales et regroupement par voisinage divisé en 3 sous-populations regroupées principalement selon leur origine géographique. Les résultats de cette étude sont cruciaux pour la conservation et leur utilisation dans la sélection végétale. La deuxième partie de cette étude vise à localiser les régions génomiques associées à des caractères importants de l'orge tels que le rendement et le caractère des composantes du rendement ; la résistance aux maladies NFB et PM ; teneur en micronutriments (Zn et Fe) dans les grains via une étude d'association à l'échelle du génome (GWAS). Pour cela, une collection diversifiée de 496 orges a été assemblée et phénotypée pour ces caractères dans un essai multilocal. La collection a ensuite été génotypée avec un tableau iselect SNP 50K d'Illumina. La cartographie des associations a été réalisée dans le logiciel TASSEL à l'aide d'un modèle linéaire mixte (MLM). La correction de Bonferroni ($p < 0,005$) a été appliquée pour réduire les faux SNP résultant en 241 SNP associés à tous les traits étudiés. La plupart de ces SNP ont été identifiés dans des études antérieures,

et certains sont colocalisés avec des gènes connus qui affectent ces traits. Les SNP ont été annotés pour les gènes candidats putatifs et ont montré que la plupart d'entre eux correspondent à des protéines associées aux traits correspondants. Les SNP identifiés dans cette étude peuvent être convertis en marqueurs conviviaux et utilisés pour développer des lignées d'orge avec un rendement, une qualité et une résistance aux maladies améliorés via la sélection assistée par marqueurs, après leur validation dans d'autres matériels génétiques.

Mots clés : Orge, GWAS, SNP, résistance aux maladies, QTL, rendement, micronutriment