

報告番号	※	第	号
------	---	---	---

主論文の要旨

論文題目 **Development and utilization of nested association mapping population in rice**
(イネ nested association mapping 集団の開発と利用)

氏名 KITONY Justine Kipruto

論文内容の要旨

高スループット DNA シーケンス技術の時代となり、量的形質遺伝子座(quantitative trait loci、QTL)マッピングにおける制限要因は、DNA マーカーではなく、遺伝解析材料となった。農業形質や環境適応などの遺伝的構造を研究するための遺伝資源は、作物の改良に必須である。Nested association mapping (NAM)は、複数の系統を親とする遺伝解析集団を利用する解析法で、従来の連鎖解析とアソシエーション解析の両方の利点を併せ持ち、複雑形質の解析に有用である。NAM 集団は、複数の組換え自殖系統 (recombinant inbred lines、RILs) の集まりであり、各 RIL は共通親と多様性供与親の交雑に由来する。NAM 集団を用いた遺伝解析では、連鎖解析をベースとする手法 (joint QTL mapping) および genome-wide association study (GWAS) をベースとする手法の両方を利用することができ、両者を照合してより良い精度を得ることが期待される。また、固定系統であるため、遺伝子型-環境間相互作用の解析にも利用可能である。一方、ゲノミックセレクション (genomic selection、GS) は、次世代の育種技術として注目されている。植物育種における GS の実装では、たとえば、DNA マーカー遺伝子型を説明変数とし、形質を従属変数とする予測モデルを作成するための比較的小規模な集団と、DNA マーカー遺伝子型のみを得るための大規模集団を組合せ、予測モデルと大規模集団の遺伝子型から、大規模集団からの選抜を DNA マーカーの情報のみで行う方法が想定されている。この予測モデルを評価するための集団としても NAM 集団は有用である。本研究では、イネ *aus* 品種を多様性供与親とし、*japonica* 品種である台中 65 号 (T65) を共通親とする NAM 集団 (aus-NAM) を作出した。初期の材料として、7 組合せの RILs からなる集団 (aus-NAM-I) を用いた解析を行い、遺伝率の高い形質である出穂期により、遺伝的振る舞いを確認した。次に、14 組合せからなる集団 (aus-NAM-II) に集団を拡大してさらに GWAS および形質予測モデルの作成と評価を行った。

まず、aus-NAM-Iの作出においては、名古屋大学で育成された5組合せからなる集団と、九州大学で育成された2組合せの計7組合せ(供与親はWRC2、WRC29、WRC31、WRC35、WRC39、DV85およびARC10313)、計895系統を得た。これらの系統の遺伝子型決定は *KpnI* と *MspI* の2種類の制限酵素を用いた genotyping-by-sequencing (GBS) 法により行った。得られた1塩基多型 (single nucleotide polymorphism, SNP) マーカーは交配組合せにより異なり、2868マーカーから4285マーカーが得られ、フィルタリングにより887系統を解析に使用した。集団構造 (population structure) を確認するために、確率的主成分分析を行った結果、WRC29由来の集団のみ隔離された集団を形成したが、十分にマッピングに利用可能と考えられた。形質として、2015年の東郷フィールドの通常期栽培における出穂期を用い、joint QTL 解析および GWAS 解析を試みた。joint QTL 解析においては7組合せのそれぞれのDNAマーカー連鎖地図において共通するSNPマーカーを抽出し、1786マーカーからなる共通連鎖地図を作成して解析を行った。joint QTL 解析に先立って行った組合せごとのQTL解析 (single family 解析) においては、染色体5、6、7および10に有意なQTLを検出した。joint QTL 解析においては染色体6、7および10にsingle family 解析と同じQTLが、染色体1、2および3に新たなQTLが検出され、染色体5のQTLは検出されなかった。染色体6および7のQTLは複数のピークとして検出された。また、GWASによる遺伝子マッピングでは、親系統のリシーケンスデータから41561個のSNPを抽出し、これをGBSによるマーカー遺伝子型にプロジェクションすることで遺伝子型データを作成し、これに系譜情報を加えた上で、集団構造と近縁度を考慮した混合線型モデル (MLM (Q+K)) により解析を行った。その結果、染色体6、7および10に有意なQTLを検出した。これらのQTLのうち、染色体6、7および10のものは既知の遺伝子座 (*RFT1*、*Hd3a*、*Hd1*、*Ghd7* および *Ehd1*) であると考えられたため、マッピング精度の評価を行った。*Ehd1*と思われるQTLのピークは約741kbの範囲に検出され、この範囲に*Ehd1*が含まれていた。*Hd1*と思われるjoint QTL 解析により検出されたQTLは、single family 解析においてはWRC39由来の集団のみで検出されており、その位置は実際の*Hd1*の位置とよく一致した。WRC39の塩基配列を確認したところ、WRC39は、解析に用いた親系統 (T65 および7系統の *aus*) の中で、唯一*Hd1*の機能型アレルを持っており、このQTLがsingle family 解析およびjoint QTL 解析により正しく検出されたと考えられた。しかしながら、GWAS 解析においては*Hd1*だけでなく、*Hd1*より6MB以上上流の*RFT1/Hd3a*を含む広い領域がピークとして検出され、single family 解析およびjoint QTL 解析の情報が無ければ*Hd1*をQTLをととして同定することは困難であった。*Ghd7*は、joint QTL 解析における染色体7の2箇所のピークのうちの1つとして検出されたが、こちらもGWASで得られた位置情報は不明確であった。これらのことから、NAM集団における遺伝子マッピングでは、位置情報の精度に優れるsingle family 解析およびjoint QTL 解析と、狭い範囲でピークを検出できるGWAS 解析、および親系統のリシーケンスデータを総合して利用することで、遺伝子マッピングおよび原因遺伝子の絞

り込みの迅速化が期待できることが明らかとなった。

次に、aus-NAM-I から親のリシーケンスデータがない 2 組合せ (DV85、ARC10313) を除き、新たに 9 組合せを育成して 14 組合せからなる aus-NAM-II 集団を確立した。aus-NAM-II の遺伝子型決定においては、aus-NAM-I で用いた手法に、イルミナの **index** を加えることで、従来のバーコード方式よりも多くの組合せ・多くの検体をひとまとめにして解析できるようになった。遺伝子型決定の結果、1818 系統を解析に使用した。形質評価は 2015 年および 2018 年の東郷フィールド・普通期栽培で行い、到穂日数、稈長、穂長、穂軸長、穂数、穂重、茎葉重、主茎の 1 次枝梗数、主茎の 1 穂粒数、バイオマス (穂重+茎葉重) および種子稔性を調査した。到穂日数に関する GWAS 解析により、既知の QTL (*Ehd1*、*Hd1*、*Hd3a/RFT1*、*Hd9*、*DTH2*、*Hd7*、*Se14* など) が正しく検出された。その他の形質に関しても、既知の QTL が検出されたほか、新規の QTL を検出した。到穂日数以外の形質では検出されるピークは少なく、また $-\log P$ 値も低い傾向で、これらの形質の遺伝率の低さが反映されていると考えられた。ゲノム予測 (genomic prediction) モデルの作成には、Bayesian B (BayesB)、Bayesian least absolute shrinkage regression (BL)、reproducing kernel Hilbert space regression (RKHS)、ridge regression best linear unbiased prediction (rrBLUP) の 4 つの手法を試みた。NAM 集団の GP においては RKHS が最も良い性能を示す事が明らかとなった。また、マーカー数は GBS によるマーカー数 (2006 マーカー) で十分であり、親系統のリシーケンスデータを用いた **projection** は不要であることが示された。これらの条件検討を行った結果、実測値と予測値の相関係数は最も高い形質で 0.87 (到穂日数、稈長)、最も低い形質で 0.56 (バイオマス) となった。実際の育種現場において到穂日数は優先順位が高いことから、イネにおける GS 育種の実装においては、到穂日数に着目すると、選抜を精度良く、かつ効率的に行なうことが示された。

以上、本研究においては、イネにおける NAM 集団の作出および高密度 DNA マーカーによる遺伝子型決定を行った。得られた材料は固定系統であり、複数回・複数箇所における形質評価が可能な優れた遺伝子マッピングのリソースである。また、遺伝子マッピングの精度および形質予測の精度は期待通りであることが示された。今後はさらなる新規遺伝子の探索や GP モデルの開発、遺伝子型-環境間相互作用のモデル化など幅広い分野での利用が期待される。