Genomic resources for pigeonpea

by Damaris A. ODENY[1]*, Trushar SHAH[2] and Rachit K. SAXENA[3]

Abstract: Pigeonpea is increasingly playing a critical role in the lives of many farmers in the semi-arid tropics, where it is grown for both subsistence and commercial use. The recent availability of genomic resources opened a new chapter in pigeonpea breeding and has led to a lot of progress in a relatively short period of time. Molecular markers have been developed and used for germplasm characterization. Both whole genome and transcriptome sequences have further improved our understanding of the genome. This review provides a summary of genomic resources available for pigeonpea breeding and also briefly discusses the way forward for pigeonpea improvement.

Key words: DArT, EST, NGS, QTL, whole genome sequence

Introduction

Advanced tools for the breeding and manipulation of pigeonpea [*Cajanus cajan* (L.) Millsp.] have been limited in the past but are developing quite rapidly. From just a few microsatellite markers in 2001, pigeonpea currently has a draft genome sequence (8), transcriptome assemblies, several on-going whole genome re-sequencing (WGRS) efforts to catalogue maximum possible variations and genomics-assisted breeding activities. To some extent, the research investment in advanced tools has also been matched with the development of relevant populations that have facilitated the identification of genomic regions contributing to traits of interest. The current major interest in hybrid pigeonpea production has further benefited from better characterization of male sterility through the sequencing of mitochondrial genomes.

[1]International Crops Research Institute for the Semi-Arid Tropics, Eastern and Southern Africa, Nairobi, Kenya (d.odeny@cgiar.org)
[2]International Institute for Tropical Agriculture, Nairobi , Kenya
[3]International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India

Molecular marker availability and utilization

Simple sequence repeat markers (SSRs) have been the most widely utilized molecular markers in pigeonpea. More than 3,000 SSRs have been reported and Diversity arrays technology (DArT) markers have also been developed. These markers have been used in characterization of germplasm, development of low density linkage maps, quantitative trait loci (QTL) studies and hybrid purity testing. Additionally, single nucleotide polymorphism (SNP) markers have been recently developed, converted into competitive allele-specific polymerase chain reaction (KASPar) (6) and BeadXpress (Illumina, San Diego, CA) (5) assays (Table 1). The SNP markers have been integrated into a consensus pigeonpea map. The availability of a draft reference genome further makes WGRS and genotyping-by-sequencing (GBS) feasible for routine genotyping in pigeonpea.

Linkage maps and quantitative trait loci (QTLs) identified

Due to the low polymorphism levels within the cultivated species experienced in pigeonpea, earlier linkage maps were developed using F_2 mapping populations between inter-specific crosses (3, 9). With the increasing numbers of polymorphic markers, intra-specific linkage maps have been generated based on genic SSR markers (3) and also on Bacterial Artificial Chromosomes (BAC)-end SSRs (1). The linkage maps have been further used to identify QTLs for several agronomic traits (1, 3). Both QTL analysis and comparative genomics have led to the identification of *CcTFL1* (4), a likely candidate gene for determinate growth habit. High density linkage and association mapping studies are in progress, especially in combination with GBS and WGRS in order to identify novel alleles for traits of agronomic importance such as disease resistance and yield and yield related traits.

Table 1. Genomic resources in terms of molecular markers, genotyping assays and sequence reads available in public domain

Genomic resource	Number of markers/assays and reads
Simple sequence repeat (SSR) markers	3,200
Single nucleotide polymorphisms (SNPs)	10,000
GoldenGate assays	768 SNPs
KASPar assays	1,616 SNPs
Sanger expressed sequence tags	~20,000
454 /FLX reads	496,705
Tentative unique sequences (TUSs)	21,432
Illumina reads(million reads)	> 160 (14 parents)

Genome and transcriptome sequences

The draft genome sequence for pigeonpea has recently become available (8). Scaffolds representing about 72.7% of the genome (estimated total genome size of 833 Mb) have been assembled for the genotype ICPL 87119 (popularly known as Asha) using the Illumina platform. The draft genome, which has 48,680 protein coding genes, has also provided a large resource of molecular markers, e.g. 309,052 SSRs and 28,104 SNPs. Mitochondrial genome sequence of three pigeonpea lines (ICPA 2039 - male sterile line, ICPB 2039 - the maintainer line and ICPH 2433 - a hybrid line) and a wild relative (ICPW 29) have also been sequenced in an effort to characterize the molecular mechanisms of cytoplasmic male sterility (CMS) (7). A transcriptome assembly comprising of 21,434 transcript assembly contigs (TACs) (2) has been generated, which combined a number of earlier transcriptome studies by analyzing 9,888 Sanger ESTs, 43,324 contigs from 1.696 million FLX/454 reads and 127,754 tentative unique sequences (TUSs). The genome and transcriptome data, together with their associated gene annotations are available on the Legume Information System (LIS) at http://www.legumeinfo.org. The LIS provides a community resource that integrates genetic, genomic and trait data across legume species. The raw data is available on the webpage of the International Initiative on Pigeonpea Genomics (IIPG) website at www.icrisat.org/gt-bt/iipg/Home.html.

Gaps

There is need to undertake more functional characterization of interesting genes identified from the genome and transcriptome sequencing processes in order to improve the annotation of the available reference genome. TILLING (Targeted induced local lesions in genome) can further reveal rare mutations resulting in novel phenotypes as has been done in other major crops. There are currently no known reports of genomic selection in pigeonpea. The possibility of undertaking WGRS/GBS together with precise trait phenotyping in pigeonpea makes genomic selection a very attractive process for varietal and hybrid improvement.

Conclusion and outlook

The genomic resources made available in pigeonpea have positioned pigeonpea well for advanced and facilitated breeding process. The relevant mapping populations will be useful in addressing some of the most serious breeding challenges including pests and diseases. Other than the utilization of wild species accessions in the understanding of cytoplasmic male sterility, the wild relatives present an alternative source of novel genes that can now be exploited with relative ease given the availability of genomic resources. Innovation in pigeonpea will come from translating the available advanced tools into practical solutions for the farmer, consumer and processors.

References

(1) Bohra A, Saxena RK, Gnanesh BN, Saxena KB, Byregowda M, Rathore A, Kavikishor PB, Cook DR, Varshney RK (2012) An intra-specific consensus genetic map of pigeonpea [*Cajanus cajan* (L.) Millsp.] derived from six mapping populations. Theor Appl Genet 25:1325-1338

(2) Kudapa H, Bharti AK, Cannon SB, Farmer AD, Mulaosmanovic B, Kramer R, Bohra A, Weeks NT, Crow JA, Tuteja R, Shah T, Dutta S, Gupta DK, Singh A, Gaikwad K, Sharma TR, May GD, Singh NK, Varshney RK (2012) A comprehensive transcriptome assembly of pigeonpea (*Cajanus cajan* L.) using Sanger and second-generation sequencing platforms. Mol Plant 5:1020-1028

(3) Kumawat G, Raje RS, Bhutani S, Pal JK, Mithra SVCR, Gaikwad K, Sharma TR, Singh NK (2012) Molecular mapping of QTLs for plant type and earliness traits in pigeonpea (*Cajanus cajan* L. Millsp.). BMC Genet13:84-100

(4) Mir RR, Kudapa H, Srikanth S, Saxena RK, Sharma A, Azam S, Saxena K, Penmetsa RV, Varshney RK (2014) Candidate gene analysis for determinacy in pigeonpea (*Cajanus* spp.). Theor Appl Genet 127:2663-2678

(5) Roorkiwal M, Sawargaonkar SL, Chitikineni A, Thudi M, Saxena RK, Upadhyaya HD, Vales MI, Riera-Lizarazu O, Varshney RK (2013) Single nucleotide polymorphism genotyping for breeding and genetics applications in chickpea and pigeonpea using the BeadXpress Platform. Plant Genome 6:1-10

(6) Saxena RK, Penmetsa RV, Upadhyaya HD, Kumar A, Carrasquilla-Garcia N, Schlueter JA, Farmer A, Whaley AM, Sarma BK, May GD, Cook DR, Varshney RK (2012) Large-scale development of cost-effective single-nucleotide polymorphism marker assays for genetic mapping in pigeonpea and comparative mapping in legumes. DNA Res19:449-461

(7) Tuteja, R. Saxena RK, Davila J, Shah T, Chen W, Xiao Y-L, Fan G, Saxena KB, Alverson AJ, Spillane C, Town C, Varshney RK (2013) Cytoplasmic male sterility-associated chimeric open reading frames identified by mitochondrial genome sequencing of four *Cajanus* genotypes. DNA Res 20:485-495

(8) Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MTA, Azam S, Fan G, Whaley AM, Farmer AD, Sheridan J, Iwata A, Tuteja R, Penmetsa RV, Wu W, Upadhyaya HD, Yang S-P, Shah T, Saxena KB, Michael T, McCombie WR, Yang B, Zhang G, Yang H, Wang J, Spillane C, Cook DR, May GD, Xu X, Jackson SA (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat Biotechnol 30:83-89

(9) Yang SY, Saxena RK, Kulwal PL, Ash GJ, Dubey A, Harper JDI, Upadhyaya HD, Gothalwal R, Kilian A, Varshney RK (2011) The first genetic map of pigeonpea based on diversity arrays technology (DArT) markers. J Genet 90:103-109