

## ESTADO DA ARTE E PERSPECTIVAS DAS PESQUISAS EM *Coffea arabica* E *Coffea canephora* NA ERA DA GENÔMICA

Taís Cristina Bastos Soares  
Ludymila Brandão Motta  
Maria Amélia Gava Ferrão

O genoma de *Coffea arabica* é composto por 44 cromossomos ( $2n = 4x$ ), o dobro do número de cromossomos *Coffea canephora*, e das demais espécies de *Coffea* (KRUG & MENDES, 1940; BOUHARMOUNT, 1959; KAMMACHER & CAPOT, 1972; CHARRIER, 1978). *C. arabica* é uma espécie alotetraplóide originada da hibridação natural de *C. canephora* (progenitor masculino) e *C. eugenioides* (progenitor feminino) (LASHERMES et al., 1999).

O cultivo de *C. arabica* e *C. canephora* faz da cafeicultura uma importante atividade econômica, com representativa produção na América, África e Ásia (FAO, 2003, CONAB, 2014).

### 1. PESQUISAS EM *Coffea* NA ERA DA GENÔMICA

A disponibilização de informações genômicas é de fundamental importância para que a produção agrícola continue a aumentar face à crescente população humana e das mudanças climáticas, e para reduzir o impacto ambiental das atividades agrícolas (MORREL et al., 2012).

Os primeiros dados de marcadores de DNA no cafeeiro foram obtidos utilizando Polimorfismos de Tamanho de Fragmentos de Restrição – RFLP, e permitiu a construção do primeiro mapa de ligação (PAILLARD et al., 1996). Estes foram utilizados por um curto período devido a grande quantidade de DNA requerida por tais marcadores, e à menor informatividade.

Durante os anos 1990, com o desenvolvimento da PCR surgiu um grande número de marcadores moleculares para avaliar a diversidade genética da espécie,

construir mapas genéticos e identificar QTLs, como os Polimorfismos de DNA amplificado ao acaso - RAPD e os Polimorfismos de comprimento de fragmentos amplificados - AFLP. Um grande salto foi conseguido com o avanço das técnicas de Sequências Simples Repetidas - SSR e Polimorfismo de Nucleotídeo Único - SNP.

Nos últimos anos, um grande avanço tem ocorrido com as metodologias de Sequenciamento de Nova Geração (NGS) (WILLIAMS et al., 1990; VOS et al., 1995; LITT & LUTY, 1989; BROOKES, 1999; KOCHKO et al., 2010). O enorme esforço técnico de pesquisas com marcadores moleculares e processos de sequenciamento de DNA ficou denominado como a Era da Genômica (RAMALHO & LAMBERT, 2004).

O sequenciamento de um genoma trata-se de desvendar as sequências de nucleotídeos do DNA de uma espécie. A primeira planta a ser sequenciada, *Arabidopsis thaliana*, levou aproximadamente 10 anos para ter o primeiro rascunho de seu genoma apresentado (*Arabidopsis* genome initiative, 2000). Hoje, com a utilização das segunda e terceira gerações de sequenciadores (e.g. 454, da Roche; SOLiD, da Applied Biosystems; *Genome Analyzer IIe*, da Illumina) e poderosos programas de bioinformática e modelagem computacional, os genomas podem ser sequenciados, montados e relacionados a características fenotípicas específicas de cada genótipo em poucas semanas (NEPOMUCENO et al., 2012).

O genoma de *C. canephora* é bastante extenso, apresentando aproximadamente 710 Mb (NOIROT et al., 2003), e como *C. arabica* é tetraplóide, este valor é ainda maior. O tamanho do genoma da espécie modelo vegetal *Arabidopsis thaliana* que teve todo genoma sequenciado de forma pioneira é estimado em 125 Mb (*Arabidopsis* genome initiative, 2000). Portanto, esta grande diferença de tamanho fez com que a maior complexidade e o alto custo para sequenciar o genoma inteiro do cafeeiro fosse um fator limitante quando o Projeto Genoma Café brasileiro foi proposto no ano de 2002.

Para reduzir gastos com o sequenciamento de genomas, uma estratégia que vem sendo bastante utilizada é o sequenciamento apenas do genoma estrutural (*Expressed Sequence Tags* - EST) (RAMALHO & LAMBERT, 2004).

### 1.1. Sequenciamento do genoma do cafeeiro

ESTs são sequências produzidas a partir de clones de cDNA. Estes clones de cDNA são geralmente organizados em grandes bibliotecas que fornecem um retrato da expressão do gene em um tecido específico ou em um órgão, em diferentes estádios de desenvolvimento e sob diferentes condições ambientais.

Diferentes grupos de pesquisa têm produzido amplos conjuntos de sequências EST em *Coffea*. No entanto, o número de ESTs disponíveis publicamente permanece baixo porque a maioria dessas sequências é de propriedade privada. Algumas instituições decidiram manter seus próprios recursos confidenciais por um tempo (O Projeto Genoma Café Brasileiro, CENICAFÉ—<https://alanine.cenargen.embrapa.br/cafEST>), enquanto outros (Nestlé, Institut de recherche pour le développement -IRD - <http://www.sgn.cornell.edu/content/coffee.pl>) o tornaram livremente disponíveis.

O Projeto Genoma Café Brasileiro gerou 130.792, 12.381 e 10.566 sequências EST de *C. arabica*, *C. canephora* e *C. racemosa*, respectivamente, reunidos em 33.000 unigenes (VIEIRA et al., 2006). O grupo de pesquisa CENICAFÉ produziu 32.961 ESTs de *C. arabica* cv. Caturra, originados dos tecidos de folhas, frutos e flores, montados em 10.799 unigenes (MONTROYA & VUONG, 2006).

Os estudos de desenvolvimento de sequências EST em *C. canephora* têm sido conduzidos por diferentes grupos de pesquisa. No IRD francês, 10.420 ESTs (montados em 5534 unigenes potenciais) foram produzidos a partir de bibliotecas de cDNA de *C. canephora* obtidas de frutos e folhas (PONCET et al., 2006). Incluindo os 47.000 ESTs, representando 13.175 unigenes, publicado pela Nestlé e da Universidade de Cornell (LIN et al., 2005), um total de 55.694 sequências estão atualmente disponíveis, compreendendo o principal recurso público para a comunidade científica. A partir de duas cultivares de *C. arabica* (Catuai vermelho e Bourbon vermelho), 1587 ESTs foram produzidos para desenvolver um microarray de cDNA contendo 1.506 ESTs de folhas e raízes embrionárias (De NARDI et al., 2006).

O GenBank oferece acesso a 187.715 ESTs de *C. arabica*, 70.407 ESTs de *C. canephora*, e *C. racemosa* com 10.838 ESTs (agosto, 2014). Este tipo de informação também é possível de se obter no banco de dados de EST da Universidade de Cornell (<http://www.sgn.cornell.edu/content/coffee.pl>), que

disponibilizou, em parceria com a Nestlé cerca de 47.000 ESTs provenientes de cinco bibliotecas cDNA de *C. canephora* organizados por tipo de tecido (Lin et al., 2005). Após o agrupamento e montagem das sequências, 13.175 unigenes foram identificados e usados para análise comparativa com os repertórios de genes de *Arabidopsis* e de tomate (*Solanum lycopersicum*). Na comparação computacional realizada, observou-se maior conservação das sequências entre *C. canephora* e tomate (ambos do clado Euasterid) do que entre *C. canephora* e *Arabidopsis* (clado Eurosids). Este número considerável de sequências representa um recurso valioso para estabelecer um catálogo exaustivo de genes para o gênero *Coffea*. No entanto, na ausência de softwares robustos de previsão de genes específicos para *Coffea*, recomenda-se que sejam realizados com algoritmos de previsão treinados com os genes de Eurosids (GUYOT et al., 2009).

O projeto de sequenciamento do genoma completo de *Coffea* só iniciou em 2010. Objetivou-se sequenciar 700 milhões de pares de bases (Mb), utilizando-se tecnologias de sequenciamento de nova geração associadas a informações já disponíveis para o gênero, como de marcadores moleculares, bancos de sequências parciais e mapas genéticos (BERTRAND, 2010). Pretendia-se, inicialmente, sequenciar um exemplar de *C. canephora*, uma vez que a espécie possui genoma diplóide, o que o torna menos complexo do que o tetraplóide *C. arabica*. Os resultados obtidos para *C. canephora*, um dos ancestrais de *C. arabica*, irão facilitar o subsequente sequenciamento dessa espécie tetraplóide. As sequências encontradas serão uma importante ferramenta na localização e identificação de genes controladores de caracteres de interesse agrônômico. Portanto, será possível desenvolver marcadores moleculares para selecionar variedades que apresentem tais características que permitam um incremento na qualidade da bebida, na sustentabilidade e na viabilidade econômica para a cultura (Bertrand, 2010).

O sequenciamento completo do genoma do café (*Coffea canephora*) foi publicado na revista Science no dia 5 de setembro de 2014 (DENOEUDE et al., 2014). Este projeto foi fruto de um consórcio internacional composto por 11 países – Brasil, França, Itália, Canadá, Alemanha, China, Espanha, Indonésia, Austrália, Índia e Estados Unidos. Os inéditos resultados encontrados possibilitarão prever o desenvolvimento de algumas características de interesse agrônômico e acelerar o melhoramento genético do cafeeiro de características relacionadas

à produtividade, precocidade, tolerância a estresses climáticos e resistência a doenças, por exemplo. O estudo comprovou, a partir de uma comparação entre os genomas do café, chá e cacau, que o surgimento da biossíntese de cafeína ocorreu independente e não oriunda de um ancestral comum, como se acreditava. Por enquanto, o banco de dados resultante do sequenciamento estrutural do café está na França, mas a ideia é trazê-lo para o Brasil, a exemplo do que foi gerado pelo genoma funcional, que desde 2004, está à disposição das instituições de pesquisa do Brasil e do exterior. Com o genoma de *C. canephora* sequenciado, se iniciam as buscas pelo sequenciamento completo de *C. arabica* (INFOCAFÉ, 2014).

## 1.2. Diversidade genética e DNA *fingerprint*

Ao longo do tempo os descritores morfológicos utilizados em espécies inseridas em programas de melhoramento genético se tornarão insuficientes para discriminar novas cultivares. Uma estratégia alternativa é a utilização de marcadores moleculares com a finalidade de identificação genética, o que é denominado *DNA fingerprint*. Esta técnica apresenta como principais vantagens em relação à fenotipagem: i) a análise do genótipo sem interferências do ambiente; ii) o alto polimorfismo presente no material genético; iii) o baixo custo da genotipagem; iv) o curto tempo requerido para as análises moleculares (RAMALHO & LAMBERT, 2004).

Marcadores de DNA já foram utilizados com a finalidade de i) estimar a diversidade genética e distinguir variedades crioulas de genótipos de populações naturais (TESFAYE et al., 2013); ii) auxiliar na determinação da origem geográfica no processo de rastreabilidade dos produtos alimentares no momento das transações comerciais; iii) caracterização de germoplasma e análise da diversidade genética (LASHERMES et al., 1996, ANTHONY et al., 2002, STEIGER et al., 2002).

Variação de DNA nuclear no café foi avaliado por meio de marcadores moleculares como RFLP, RAPD (LASHERMES et al., 1999; DINIZ et al., 2005; ANTHONY et al., 2002, SILVEIRA et al., 2003), AFLP (STEIGER et al., 2002; ANTHONY et al., 2002) e SSR (COMBES et al., 2000; ANTHONY et al., 2002; MONCADA & McCOUTH, 2004; MALUF et al., 2005; PONCET et

al., 2006; AGGARWAL et al., 2007; SILVESTRINI et al., 2007, VIEIRA et al., 2010, FERRÃO et al., 2014), em que foi demonstrado que a variação genética no gênero *Coffea* é baixo, especialmente entre variedades cultivadas de *C. arabica*.

Os dados de *fingerprinting* para já disponíveis para *C. canephora* e *C. arabica* podem ser usados para construir uma base de dados de DNA de referência que auxilie na identificação molecular de variedades, como já foi sugerido (AGGARWAL et al., 2004; HENDRE et al., 2008; VIEIRA et al., 2010).

Diversos marcadores moleculares estão disponíveis para estabelecer a origem das variedades de café para fins científicos, mas não tem sido utilizados e validados para fins comerciais. Com o avanço da ciência forense na área alimentícia, metodologias de análise de DNA começaram a ser requeridos para avaliar grãos de café verdes ou torrados, dos grupos Arabica, Robusta ou em blends. Este tipo de informação pode ser aplicado para verificar possíveis contaminações, permitir a rastreabilidade, e verificar a autenticidade do produto (TORNICASA et al., 2010). A pesquisa sobre café neste campo ainda é incipiente. Destacam-se a utilização de métodos baseados em PCR-RFLP (Spaniolas et al., 2006) e PCR em tempo real (TORNICASA et al., 2010).

### **1.3. Cafeeiro geneticamente modificado**

Os transgênicos normalmente apresentam ganhos de função, enquanto o método convencional em muitas vezes ocorre perda de função, uma vez que muitos alelos importantes são recessivos. Isto ocorre uma vez que no melhoramento clássico milhares de genes são manipulados simultaneamente, enquanto nos transgênicos trata-se de um ou poucos genes (GEPTS, 2002).

Os principais objetivos do uso da técnica de engenharia genética no cafeeiro são desenvolver novas cultivares tolerantes a estresses bióticos e abióticos, como pragas, doenças, herbicidas, seca, geada, além de materiais com alta qualidade da bebida e baixos teores de cafeína.

Os estudos do transcriptoma do cafeeiro, a disponibilidade de muitas ESTs de *C. canephora* e *C. arabica*, e o desenvolvimento de bibliotecas genômicas abriram novas possibilidades na área de genômica funcional de café. Isto vai ajudar no direcionamento da inserção da característica de interesse, utilizando

várias ferramentas de transformação, com o aumento de probabilidade de sucesso e redução dos custos (MISHRA & SLATER, 2012).

Caféeiros geneticamente modificados foram obtidos por diferentes grupos de pesquisa do mundo, destacando-se o Brasil (ex. ALBUQUERQUE et al., 2009), Índia (ex. KUMAR et al., 2006) e França (ex. RIBAS et al., 2011). Apesar dos avanços significativos nos últimos anos, a transformação do café ainda não é um procedimento de rotina nos programas de melhoramento. Mishra & Slater (2012) fizeram uma revisão detalhada sobre transformação genética no café.

#### **1.4. Técnicas de edição do genoma**

Abordagens das genéticas evolutiva e quantitativa podem ser utilizadas para identificar a localização genômica e o efeito de locos de importância agrônômica. No entanto, a validação dos efeitos genéticos e o uso de alelos individuais em programas de melhoramento de plantas demanda alto investimento de tempo e dinheiro (BERNARDO, 2008). Espera-se que a seleção genômica acelere a introgressão de alelos múltiplos favoráveis em populações reprodutoras. Em programa de introgressão assistida por marcadores, grandes segmentos cromossômicos são introduzidos, o que limita o uso de retrocruzamentos para testar o efeito genético de alelos individuais, e aumenta o risco de que uma introgressão de variação indesejada esteja associada (MORREL et al., 2012).

Tecnologias direcionadas de edição de genomas podem proporcionar oportunidades interessantes para alterar nucleotídeos individuais e pequenas regiões de genes nativos. O recente desenvolvimento destas tecnologias de edições específicas de genoma, como as nucleases de dedos de zinco (WEINTHAL et al., 2010) e nucleases TALE (transcription activator-like effector) (BOGDANOVE & VOYTAS, 2011) oferecem um grande potencial para resolver os problemas mencionados anteriormente. Essas tecnologias fazem uso de nucleases de sequências específicas que clivam os locos alvos, permitindo a criação de pequenas inserções e deleções (indels), a inserção de segmentos de DNA ou mesmo a substituição de alelos individuais. É até possível conseguir substituir sequências que causam mutações deletérias em linhagens elites (CHARLESWORTH & WILLIS, 2009). Ambos os métodos de nucleases de dedos de zinco e nucleases

TALE têm sido aplicados com sucesso para culturas vegetais, e seus impactos esperados para o melhoramento de plantas são enormes (SHUKLA et al., 2009; MORBITZER et al., 2010).

Uma vez que os ciclos de melhoramento do cafeeiro são longos, recomenda-se que estudos nesta área comecem a ser desenvolvidos. As técnicas de edição de genoma podem vir a contribuir na obtenção de características específicas com maior eficácia e rapidez, e menor custo.

### **1.5. Mapeamento genético**

A maioria dos mapas genéticos interespecíficos de café foram construídos para identificar QTLs envolvidos em características contrastantes existentes em várias espécies selvagens. Mapas genéticos de *C. canephora* e cruzamentos interespecíficos estão disponíveis (PAILLARD et al., 1996; KY et al., 2000; LASHERMES et al., 2001; HERRERA et al., 2002, AKAFFOU et al., 2003; COULIBALY et al., 2003; LEFEBVRE-PAUTIGNY et al., 2010, dentre outros). Essas informações disponíveis para *C. canephora* são importantes para que um mapa consenso seja criado para espécies de *Coffea*.

Mapas genéticos para *C. arabica* são mais escassos devido a algumas razões, como a complexidade de um genoma tetraplóide, o alto nível de homozigose e o pequeno número de polimorfismos (PEARL et al 2004; De OLIVEIRA, 2007).

### **1.6. Seleção assistida por marcadores e seleção genômica ampla**

A seleção assistida por marcadores é a correlação genética entre a marca e os genes envolvidos no controle dos caracteres expressos (RAMALHO & LAMBERT, 2004). A técnica pode ser utilizada em duas vertentes: i) para acelerar a recuperação de alelos de interesse; ou ii) auxiliar na eliminação de alelos indesejáveis aos programas de melhoramento, que estão ligados a locos de interesse (HOSPITAL & CHARCOSSET, 1997).

O desenvolvimento de genotipagem ‘high-throughput’ levou a uma mudança nas análises de mapeamento de QTLs para estudos tradicionais de associação. A nova proposta, ao invés de focar em duas linhagens parentais contrastantes



fenotipicamente, permite que no mapeamento associativo seja avaliada a correlação entre fenótipo e genótipo em conjuntos de indivíduos não aparentados. Com isto, amostra-se uma diversidade genética bem maior, resultante de muito mais eventos de recombinação, além de evitar as gerações de cruzamentos demorados que são necessárias para o mapeamento de QTLs (MYLES et al., 2009).

Marcadores genéticos de alta densidade estão sendo utilizados em estudos de associação de genoma (GWASs) e também podem ser explorados para seleção genômica (MORREL et al., 2012). A seleção genômica ampla é uma forma de seleção assistida por marcadores, em que um conjunto de dados moleculares é usado para fazer previsões fenotípicas (MEUWISSEN et al., 2001; HEFFNER et al., 2009). Seleção genômica e estudos de associação de genoma (GWASs) podem usar os mesmos dados genotípicos e fenotípicos. O que difere, principalmente é que os modelos de seleção genômica enfatizam a identificação de polimorfismos individuais controladores de características complexas por uma predição dos valores fenotípicos, que são baseados em um conjunto de dados de treinamento. Como os GWASs, a aplicação da seleção genômica estava sendo limitada pelo custo e disponibilidade de dados densos de marcadores do genoma. No entanto, com o surgimento metodologias de genotipagem em larga escala as informações moleculares passaram a apresentar baixo custo e tempo reduzido (ANDOLFATTO et al., 2011; ELSHIRE et al., 2011).

Hoje, a eficiência da seleção genômica está muito abaixo da meta sugerida na proposta inicial. A precisão está limitada, principalmente, pela ineficiência na predição do fenótipo. Apesar destes problemas, os métodos atuais de seleção genômica mostram-se de 2-3 vezes mais rápidos do que pelo ciclo de reprodução tradicional. Espera-se que a seleção genômica vai revolucionar reprodução na próxima década (MORREL et al., 2012).

O grande potencial apresentado pela Seleção Genômica Ampla (Genome-Wide Selection – GWS) tem estimulado alguns grupos de pesquisa cafeeira a investir esforços na utilização da técnica para auxiliar na predição de genótipos de interesse com grande eficácia e prazos bem mais curtos. Ressalta-se que o laboratório de Bioquímica e Biologia Molecular do CCA-UFES está inserido em um projeto de GWS para *C. canephora*, em parceria com o Incaper e Embrapa Café.

## 1.7. Bancos de dados do Cafeeiro disponíveis na web

Nome	Descrição
The CENICAFE coffee databases – Colombia	Banco de Dados de 32.000 ESTs de <i>C. arabica</i> e <i>C. liberica</i> . Disponível em: <a href="http://bioinformatics.cenicafe.org/">http://bioinformatics.cenicafe.org/</a> Cristancho et al. (2006)
TropGENE DB CIRAD – France	Biblioteca com 55.296 BACs e 253 marcadores SSR de <i>C. canephora</i> Disponível em: <a href="http://tropgenedb.cirad.fr/tropgene/JSP/interface.jsp?module=COFFEE">http://tropgenedb.cirad.fr/tropgene/JSP/interface.jsp?module=COFFEE</a> Leroy et al. (2005); Ruiz et al. (2004)
Ccmb coffee database – India	Banco de dados de caracterização molecular do germoplasma café disponível na Índia. Disponível em: <a href="http://www.ccmb.res.in/coffeegermplasm/index.htm">http://www.ccmb.res.in/coffeegermplasm/index.htm</a> (Em construção)
The Brazilian Coffee Genome EST Project CBP&D-Café – Brasil	130.792, 12.381 e 10.566 seqüências de <i>C. arabica</i> , <i>C. canephora</i> e <i>C. racemosa</i> , respectivamente, (37 bibliotecas de cDNA). Disponível em: <a href="http://www.lge.ibi.unicamp.br/cafe/">http://www.lge.ibi.unicamp.br/cafe/</a> Vieira et al. (2006)
Coffee DNA – University of Trieste, Italy	Banco de dados com 13,686 ESTs, 266 Microsatellites, 43 retrotransposon. Disponível em: <a href="http://www.coffeedna.net/">http://www.coffeedna.net/</a>
MoccaDB – IRD, France	Informações demarcadores RFLP, EST-SSR, SSR, SNP (to date, 638 markers) de aproximadamente 38 espécies de Rubiaceae. Disponível em: <a href="http://moccadb.mpl.ird.fr/">http://moccadb.mpl.ird.fr/</a> Plechakova et al. (2009)
The SOL Genomics Network (SGN) Cornell University – USA	Informações genômicas da família Solanaceae e da família relacionada no clade Euasterid. Apresenta 47.000 ESTs de café ( <i>C. canephora</i> var <i>robusta</i> ) divulgados pela Universidade de Cornell e Nestlé S.A. Disponível em: <a href="http://sgn.cornell.edu/">http://sgn.cornell.edu/</a> Mueller et al. (2005)

Tabela modificada e traduzida de Kochko et al., 2010.

## **2. FENOTIPAGEM**

Os estudos moleculares no melhoramento de plantas precisam estar associados à caracterização fenotípica. As culturas vegetais podem ser propagadas por clonagem e mantidas como linhas puras. Isso torna possível sequenciar um genótipo apenas uma vez, porém fenotipá-lo ao longo do tempo e em ambientes diferentes (NORDBORG et al., 2008).

Portanto, é necessário obter as informações em campo com a maior precisão experimental possível. As plantas obtidas por métodos biotecnológicos apresentam as mesmas necessidades de serem avaliadas intensivamente em campo para verificar se irão expressar as características desejadas. Portanto, a fenotipagem continua sendo de extrema importância para se avançar na caracterização e avaliação de materiais genéticos vegetais. A necessidade de grande quantidade de informações fenotípicas faz com que a importância da biometria na Era da Genômica seja ainda maior do que no passado (RAMALHO & LAMBERT, 2004).

## **3. PERSPECTIVAS FUTURAS**

As projeções divulgadas sobre o enorme crescimento populacional e a respeito dos impactos das mudanças climáticas globais evidenciam a necessidade da agricultura se desenvolver a passos largos para que seja possível suprir as demandas geradas por tais eventos.

Os bancos de informações genéticas que já estão disponíveis precisam ser analisados criticamente para que seja possível avançar nas avaliações com as novas tecnologias que estão surgindo, mas também é preciso aprender com as informações que já estão disponíveis na literatura. E para isto, é necessário buscar as várias informações geradas por diferentes grupos de pesquisas e integrá-las.

A compreensão dos mecanismos genéticos e o conhecimento da diversidade disponível permitirão obter germoplasmas de cafeeiros mais produtivos, tolerantes a estresses bióticos e abióticos, com maior qualidade de bebida e importância nutricional. As pesquisas multidisciplinares possuem o potencial de proporcionar um incremento significativo nos programas de melhoramento do cafeeiro, seja via

engenharia genética, seleção genômica ampla assistida por marcadores, edição de genomas, dentre outras metodologias. As técnicas moleculares precisam estar associadas com análises fenotípicas que gerem informações em volume e qualidade comparável com as informações genômicas que já se encontram disponíveis. Para lidar com este vultoso conjunto de informações precisa-se avançar, paralelamente, no desenvolvimento de softwares e na formação de profissionais na área de bioinformática.

#### 4. REFERÊNCIAS

AGGARWAL, R. K.; RAJKUMAR, R.; RAJENDRAKUMAR, P.; HENDRE, P. S.; BARUAH, A.; PHANINDRANATH, R.; ANNAPURNA, V.; PRAKASH, N. S.; SANTARAM, A.; SREENIVARSAN, C. S.; et al. Fingerprint of Indian coffee selections and development of reference DNA polymorphism panels for creating molecular IDs for variety identification. **Proceedings of 20th International Conference on Coffee Science (ASIC)**, Bangalore, Índia, p. 751-755, 2004.

AKAFFOU, D.S.; KY, C.L.; BARRE, P.; HAMON, S.; LOUARN, J.; NOIROT, M. Identification and mapping of a major gene (Ft1) involved in fructification time in the interspecific cross *Coffea pseudozanguebariae* x *C. liberica* var. Dewevrei: impact on caffeine content and seed weight. **Theoretical and Applied Genetics**, v. 106, p. 1486-1490, 2003.

ALBUQUERQUE E. V. S.; CUNHA, W. G.; BARBOSA, A. E. A. D.; et al. Transgenic coffee fruits from *Coffea arabica* genetically modified by bombardment, **In Vitro Cellular & Developmental Biology-Plant**, v. 45, n. 5, p. 532-539, 2009.

ANDOLFATTO, P.; et al. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. **Genome Res**, n.21, p. 610-617, 2011.

ARABIDOPSIS GENOME INITIATIVE. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature**, n. 408, p. 796-815, 2000.

BERNARDO, R. Molecular markers and selection for complex traits in plants: Learning from the last 20 years. **Crop Sci**, n. 48, p.1649-1664, 2008.

BERTRAND, B. **Decryption of the coffee genome has begun**. 2010. Disponível em: <<http://www.coffeegenome.org/communications/publications/CIRAD%20>

[Article.jpg](#)>. Acesso em: 27 jul. 2014.

BOGDANOVE, A. J.; VOYTAS, D. F. TAL effectors: customizable proteins for DNA targeting. **Science**, n. 333, p. 1843-1846, 2011.

BOUHARMOUNT, J. Recherche sur les affinités chromosomiques dans le genre *Coffea*. **INEAC Séries Sci**, 1959.

BROOKES, A. J. The essence of SNPs. **Gene**, n. 234, p.177-186, 1999.

CAPOT, J. L'amélioration du caféier robusta en Côte d'Ivoire: Les hybrides Arabusta. **Café Cacao Thé**, 1972.

CHARLESWORTH, D.; WILLIS, J. H. The genetics of inbreeding depression. **Nature Rev. Genet**, n. 10, p. 783-796, 2009.

CHARRIER, A. La structure génétique des caféiers spontanés de la région malgache (Mascaro-coffee). Mémoires ORSTOM 87, **Orstom**, Ed., Paris, 1978.

CONAB - Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira**. Café Safra 2014, primeira estimativa, janeiro/2014. Companhia Nacional de Abastecimento. Brasília, Conab, 2014.

COULIBALY, I.; LOUARN, J.; LORIEUX, M.; CHARRIER, A.; HAMON, S.; NOIROT, M. Pollen viability restoration in a *Coffea canephora* P. and C-heterocalyx stoffelen backcross: QTL identification for marker-assisted selection. **Theor. Appl. Genet**, n.106, p. 311-316, 2003.

CRISTANCHO, M.; RIVERA, L.; OROZCO, C.; CHALARCA, A.; MUELLER, L. Development of a Bioinformatics Platform at the Colombia National Coffee Research Center. **Association Scientifique Internationale du Café' (ASIC)**, Montpellier, France. 2006

DE NARDI, B.; DREOS, R.; DEL TERRA, L.; MARTELLOSSI, C.; ASQUINI, E.; TORNINCASA, P.; et al. Differential responses of *Coffea arabica* L. Leaves and roots to chemically induced systemic acquired resistance. **Genome**, n. 49, p. 1594-1605, 2006.

DENOEUDE, F., CARRETERO-PAULET, L., DEREPPER, A., DROC, G., GUYOT, R., PIETRELLE, M., et al. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, n. 345, p.1181-1184, 2014.

DE OLIVEIRA, A.C.B.; SAKIYAMA, N.S.; CAIXETA, E. T.; ZAMBOLIM, E.M.; RUFINO, R. J. N.; ZAMBOLIM, L. Partial map of *Coffea arabica* L. and recovery of the recurrent parent in backcross progenies. **Crop Breed. Appl. Biotechnol**, n. 7, p. 196-203, 2007.

ELSHIRE, R. J.; et al. A robust, simple genotyping-bysequencing (GBS) approach for high diversity species. **PLoS ONE**, v. 6, 2011.

FAO - Food and Agriculture Organization of the United Nations. Medium-term prospects for agricultural commodities. **Projections to the year 2010**, p. 89, 2003.

FERRÃO, L. F. V.; CAIXETA, E. T.; CRUZ, C. D.; DE SOUZA, F. F.; FERRÃO, M. A. G.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; SAKIYAMA, N. S. The effects of encoding data in diversity studies and the applicability of the weighting index approach for data analysis from different molecular markers. **Plant Systematics and Evolution**, p. 1-13, 2014.

GEPTS, P. A. A comparison between crop domestication, classical plant breeding, and genetic engineering. **Crop Science**, Madison, v. 42, p. 1780-1790, 2002.

GUYOT, R.; DE LA MARE, M.; VIADER, V.; HAMON, P.; CORITON, O.; BUSTAMANTE-PORRAS, J.; et al. Microcollinearity in an ethylene receptor coding gene region of the *Coffea canephora* genome is extensively conserved with *Vitis vinifera* and other distant dicotyledonous sequenced genomes. **BMC Plant Biol**, v. 9, n. 22, 2009.

HEFFNER, E. L.; SORRELLS, M. E.; JANNINK, J. L. Genomic selection for crop improvement. **Crop Sci**, n. 49, p. 1-12, 2009.

HENDRE, P. S.; PHANINDRANATH, R.; ANNAPURNA, V.; LALREMRUATA, A.; AGGARWAL, K. Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies. **BMC Plant Biol**, v. 8, n. 51, 2008.

HERRERA, J.C.; COMBES, M.C.; ANTHONY, F.; CHARRIER, A.; LASHERMES, P. Introgression into the allotetraploid coffee (*Coffea arabica* L.): segregation and recombination of the *C. canephora* genome in the tetraploid interspecific hybrid (*C. arabica* x *C. canephora*). **Theor. Appl. Genet**, 2002.

HOSPITAL, F.; CHARCOSSET, A. Marker-assisted introgression of quantitative trait loci. **Genetics**, Baltimore, v. 147, p. 1469-1485, 1997.

INFOCAFÉ. Disponível em [http://www.revistacafeicultura.com.br/index.php?tipo=ler&mat=55114&&utm\\_source=feedburner&utm\\_medium=email&utm\\_campaign=Feed%3A+RevistaCafeicultura+%28Revista+Cafeicultura%29](http://www.revistacafeicultura.com.br/index.php?tipo=ler&mat=55114&&utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+RevistaCafeicultura+%28Revista+Cafeicultura%29). Acesso em 29 de setembro de 2014.

KAMMACHER, P.; CAPOT, J. Sur les relations caryologiques entre *Coffea arabica* et *C. canephora*. **Café Cacao Thé**, 1972.

KOCHKO, A. D. E.; AKAFFOU, S.; ANDRADE, A. C.; CAMPA, C.; CROUZILLAT, D.; GUYOT, R.; HAMON, P.; MING, R.; MUELLER, L. A.; PONCET, V.; TRANCHANT-DUBREUIL, C.; HAMON, S. Advances in coffee genomics. **Adv Bot Res**, n.53, p. 23-63, 2010.

KRUG, C. A.; MENDES, A.J.T. Cytological observation in *Coffea*, J. **Genet**, 1940.

KUMAR, V.; SATYANARAYANA, K. V.; SARALA ITTY, S., et al. Stable transformation and direct regeneration in *Coffea canephora* P ex. Fr. by *Agrobacterium rhizogenes* mediated transformation without hairy-root phenotype. **Plant Cell Reports**, v. 25, n. 3, p. 214-222, 2006.

KY, C.L.; BARRE, P.; LORIEUX, M.; TROUSLOT, P.; AKAFFOU, S.; LOUARN, J.; CHARRIER, A.; HAMON, S.; NOIROT, M. Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). **Theor. Appl. Genet**, 2000.

LASHERMES, P.; COMBES, M.C.; PRAKASH, N.S.; TROUSLOT, P.; LORIEUX, M.; CHARRIER, A. Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meiosis. **Genome**, 2001.

LASHERMES, P.; COMBES, M.C.; ROBERT, J.; TROUSLOT, P.; D'HONT, A.; ANTHONY, F.; CHARRIER, A. Molecular characterization and origin of the *Coffea arabica*, L. **Genome. Mol. Gen. Genet**, 1999.

LASHERMES, P.; TROUSLOT, P.; COMBES, M. C.; CHARRIER, A. Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. **Euphatica**, 1996.

LEFEBVRE-PAUTIGNY, F.; WU, F.; PHILIPPOT, M.; RIGOREAU, M.; ZOUINE, M.; FRASSE, P.; BOUZAYEN, M., et al. "High resolution synteny maps allowing direct comparisons between the coffee and tomato genomes." **Tree**

**Genetics & Genomes**, v. 6, n. 4, p. 565-577, 2010.

LEROY, T.; MARRACCINI, P.; DUFOUR, M.; MONTAGNON, C.; LASHERMES, P.; SABAU, X.; et al. Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. **Theor. Appl. Genet**, n.111, p. 1032-104, 2005.

LIN, C.; MUELLER, L. A.; MC CARTHY, J.; CROUZILLAT, D.; PETIARD, V.; TANKSLEY, S.D. Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. **Theor. Appl. Genet**, n. 112, p. 114-130, 2005.

LITT, M.; JEFFREY, A. L.A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. **American journal of human genetics**, v. 44, n. 3, p. 397, 1989.

MEUWISSEN, T. H.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, n. 157, p. 1819-1829, 2001.

MISHRA, M. K., SLATER, A. Recent advances in the genetic transformation of coffee. **Biotechnology research international**, 2012.

MONTOYA, G.; VUONG, H.; CRISTANCHO, M.; MONCADA, P.; YEPES, M. Sequence analysis from leaves, flowers and fruits of *Coffea arabica* var. Caturra. **21<sup>st</sup> International Conference on Coffee Science (ASIC)**, Montpellier, France. 2006.

MORBITZER, R.; ROMER, P.; BOCH, J.; LAHAYE, T. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)- type transcription factors. **Proc. Natl Acad. Sci.**, USA, n. 107, p. 21617-21622, 2010.

MORRELL, P. L.; BUCKLER, E. S.; ROSS-IBARRA, J. Crop genomics: advances and applications. **Nature Reviews Genetics**, v. 13, n. 2, p. 85-96, 2011.

MUELLER, L.A.; SOLOW, T.H.; TAYLOR, N.; SKWARECKI, B.; BUELS, R.; BINNS, J.; et al. The SOL genomics network. A comparative resource for solanaceae biology and beyond. **Plant Physiol**, n. 138, p. 1310-1317, 2005.

MYLES, S. et al. Association mapping: critical considerations shift from genotyping to experimental design. **Plant Cell**, n. 21, p. 2194-2202, 2009.



NEPOMUCENO, A. L.; LABEX US BIOTECNOLOGIA.; ALVES, R. E. A Nova Era Genômica e a Biodiversidade Brasileira. **Embrapa Florestas, -Nota técnica (ALICE)**, 2012.

NGANOU, D. N.; DURAND, N.; TATSADJIEU, N. L.; MEILE, J. C.; EL-SHEIKHA, A. F.; MONTET, D.; MBUFUNG, C. M. Determination of coffee origin by using 28S rDNA fingerprinting of fungal communities by PCR-DGGE: application to the Cameroonian coffee. **International Journal of Biosciences (IJB)**, v. 2, n. 5, p.18-30, 2012.

NOIROT, M.; PONCET, V.; BARRE, P.; HAMON, P.; HAMON, S.; DE KOCHKO, A. Genome size variations in diploid African *Coffea* species. **Ann Bot London**, v. 92, n. 5, p. 709-714, 2003.

NORDBORG, M.; WEIGEL, D. Next-generation genetics in plants. **Nature**, n. 456, p. 720-723, 2008.

PAILLARD, M.; LASHERMES, P.; PÉTIARD, V. Construction of a molecular linkage map in coffee. **Theor. Appl. Genet.**, 1996.

PEARL, H.M.; NAGAI, C.; MOORE, P.H.; STEIGER, D.L.; OSGOOD, R.V.; MING, R. Construction of a genetic map for arabica coffee. **Theor. Appl. Genet.**, n. 108, p. 829-835, 2004.

PLECHAKOVA, O.; TRANCHANT-DUBREUIL, C.; BENEDET, F.; COUDERC, M.; TINAUT, A.; VIADER, V.; et al. MoccaDB – An integrative database for functional, comparative and diversity studies in the Rubiaceae family. **BMC Plant Biol**, v. 9, n. 123, 2009.

PONCET, V.; RONDEAU, M.; TRANCHANT, C.; CAYREL, A.; HAMON, S.; DE KOCHKO, A., et al. SSR mining in coffee tree EST databases: potential use of ESTSSRs as markers for the *Coffea* genus. **Mol. Genet. Genomics**, n. 276, p. 436-449, 2006.

RAMALHO, M. A. P.; LAMBERT, E. S. Biometria eo melhoramento de plantas na era da genômica. **Revista Brasileira de Milho e Sorgo**, v. 3, n. 2, 2010.

RIBAS, A. F.; DECHAMP, E.; CHAMPION, A.; et al. Agrobacterium mediated genetic transformation of *Coffea arabica* (L.) is greatly enhanced by using established embryogenic callus cultures. **BMC Plant Biology**, v. 11, n. 92, 2011.

RUIZ, M.; ROUARD, M.; RABOIN, L.M.; LARTAUD, M.; LAGODA, P.;

COURTOIS, B. TropGENE-DB, a multi-tropical crop information system. **Nucleic Acids Res**, n. 32, p. 364-367, 2004.

SPANIOLAS,S.; MAY, S.T.; BENNETT, M. J.; TUCKER,G. A. Authentication of Coffee by Means of PCR-RFLP Analysis and Lab-on-a-Chip Capillary Electrophoresis.**J. Agric. Food Chem**, v. 54, n. 20, p. 7466-7470, 2006.

SHUKLA, V. K.;et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. **Nature**, n. 459, p. 437-441, 2009.

TESFAYE, K.; GOVERS, K.; BEKELE, E.; BORSCH, T.ISSR fingerprinting of *Coffea arabica* throughout Ethiopia reveals high variability in wild populations and distinguishes them from landraces. **Plant Systematics and Evolution**, v. 300, n. 5, p. 881-897, 2014.

TORNINCASA,P.; FURLAN, M.; PALLAVICINI, A.; GRAZIOSI, G. Coffee species and varietal identification, 2010.

VIEIRA, E. S. N.; VON PINHO, E. V. R.; CARVALHO, M. G. G.; ESSELINK, D. G.; VOSMAN, B. Development of microsatellite markers for identifying Brazilian *Coffea arabica* varieties. **Genetics and molecular biology**, v. 33, n. 3, p. 507-514, 2010.

VIEIRA, L.G.E.; ANDRADE, A.C.; COLOMBO, C.A.;ARAÚJO, A.H.; METHA, A.; OLIVEIRA, A.C., et al.Brazilian coffee genome project: an EST-based genomic resource. **Braz. J. Plant Physiol**, v. 18, p. 95-108, 2006.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M., et al.AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Res.**, n. 23, p. 4407-4414,1995.

WEINTHAL, D.; TOVKACH, A.; ZEEVI, V.; TZFIRA, T. Genome editing in plant cells by zinc finger nucleases. **Trends Plant Sci.**, n. 15, p. 308-321, 2010.

WILLIAMS, J. G. K.; KUBELIK, A. R.; LIVAK, K. J.; RAFALSKI, J. A.; TINGEY, S. V.DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic acids research**, v. 18, n. 22, p. 6531-6535, 1990.