



Hinc patriam sustinet

Instituto Superior de Agronomia
Universidade Técnica de Lisboa

**Caracterização da diversidade genómica de
sequências repetitivas no género *Avena***

Projecto
Licenciatura em Biologia

Joana Rita António Rodrigues
Orientadora: Investigadora Manuela Gomes da Silva

Lisboa, 2012

Agradecimentos

Começo por agradecer à professora Manuela Silva, por me ter aceite e integrado na sua equipa, por toda a simpatia e dedicação. Pelos ensinamentos transmitidos. Pelo desafio que me lançou ao propor este trabalho que me fez crescer como pessoa e como bióloga. O meu muito obrigada.

À professora Wanda Viegas, pelos conhecimentos que me transmitiu ao longo dos anos como aluna do ISA e por me ter recebido tão bem no laboratório.

À Diana Tomás, por todos os conhecimentos que me transmitiu, mas também pela amizade, paciência e boa disposição. Por me “localizar geograficamente” o material, quando a memória me falhava. Pelo tempo partilhado em “estufa *time*”.

À Augusta Barão, por conseguir pôr-me sempre a sorrir, ao mesmo tempo que me transmitia conhecimentos valiosos. Pelo tempo despendido comigo ao microscópio.

Aos colegas do departamento de genética, pelo bom ambiente de trabalho e por não deixarem passar um dia sem o tornar imprevisível e inesquecível.

Aos amigos e á família por terem sempre uma palavra de apoio, um abraço no momento certo e nunca duvidarem que eu seria capaz de completar esta etapa. Muito obrigada.

Sumário

O género *Avena* L. contém espécies com diferentes níveis de ploidia, sendo a sua evolução ainda não totalmente compreendida. O presente trabalho tem como objectivo contribuir para a caracterização genómica de diferentes espécies e variedades de *Avena* através de análise molecular e citogenética de distintas sequências repetitivas.

A análise molecular por PCR mostra que a região ITS1 de rDNA (Internal Transcribed Spacer) é mais variável que a ITS2 e que sequências repetitivas características dos genomas A e C de *Avena* permitem uma melhor distinção intraespecífica. Por outro lado, a determinação do número de NORs (Nucleolar Organizer Regions) por hibridação *in situ* fluorescente revelou que *A. strigosa* apresenta 4 NORs e *A. sativa* apresenta 6, como anteriormente descrito para outras variedades. Em *A. sterilis* o número de NORs difere entre linhas, variando entre 6 e 8. Nas espécies hexaplóides estudadas foi possível confirmar ainda a existência de translocações em células metafásicas por FISH utilizando uma sonda específica de genoma C. A organização interfásica do rDNA assim como do genoma C foi também caracterizada pela primeira vez em *A. sativa*, revelando um padrão de organização distinto do anteriormente descrito para outras espécies de cereais.

Índice

Agradecimentos	ii
Sumário	iii
Índice	iv
1. Introdução	1
1.1. Género <i>Avena</i>	1
1.2. Sequências repetitivas.....	2
1.3. Objectivo	2
2. Materiais e métodos	3
2.1. Material Vegetal	3
2.2. Métodos	3
2.2.1. Análise molecular por PCR	3
2.2.2. Análise citogenética por FISH	7
3. Resultados	10
3.1. Análise molecular	10
3.1.1. rDNA 45s – ITS1 e ITS2	10
3.1.2. Sequências específicas de <i>Avena</i>	12
3.2. Análise citogenética.....	14
4. Discussão e perspectivas futuras	16
5. Referências bibliográficas	18
Anexos	20

Caracterização da diversidade genómica de sequências repetitivas no género *Avena*

1. Introdução

1.1. Género *Avena*

O género *Avena* L. pertence á família *Poaceae*, subfamília *Pooideae*, tribo *Aveneae*. Sendo originário da região oeste do mediterrâneo, encontram-se membros deste género dispersos por Portugal, Espanha e Grã-Bretanha (centros de diversidade da *A. strigosa*); Egipto, Etiópia e sudoeste Asiático (centros de diversidade da *A. sterilis*); Rússia e China (centros de diversidade da *A. sativa*). (Tavares et al., 1993; Loskutov, 2008).

Neste género existem espécies com três níveis de ploidia, sendo o número básico de cromossomas $n = 7$: diplóides ($2n=14$) com genoma AA ou CC, tetraplóides ($2n=28$) com genoma AABB ou AACC, e hexaplóides ($2n=42$) com genoma AACCCD (Linares et al., 1998; Katsiotis et al., 2000). Vários estudos demonstram uma grande afinidade entre os genomas A e B assim como entre os genomas A e D, pelo que se pensa que tanto o genoma B como o D tenham evoluído de um genoma A ancestral. Esta hipótese é também apoiada pelo facto de, até à data, não haver registo de espécies diplóides com genoma B ou D (Katsiotis and Heslop-Harrison, 1997, Katsiotis et al., 2000). Apesar de a evolução deste género ainda não ser totalmente compreendida, é consensual a hipótese que sugere que duas espécies diplóides, uma com genoma A e outra com C, tenham hibridado originando um híbrido AC que sofreu seguidamente uma duplicação cromossómica que originou uma espécie tetraplóide. Esta espécie terá posteriormente hibridado com uma outra espécie diplóide de genoma A' precursora do genoma D, a que se seguiu uma nova duplicação cromossómica, dando assim origem a uma espécie hexaplóide, possivelmente *A. sterilis* devido á sua grande dispersão. Pensa-se que terá sido a *A. sterilis* a espécie hexaplóide ancestral na origem das várias espécies com genoma AACCCD (Loskutov, 2008).

Dada a grande proximidade genómica de todas as espécies hexaplóides de *Avena*, estas são frequentemente interférteis, característica com um grande potencial uma vez que permite a produção de genótipos híbridos entre espécies comerciais com elevada produtividade e espécies selvagens com resistência a doenças e com poucas exigências ecológicas. Estes híbridos poderão ser essenciais não só para a produção de variedades melhoradas mas também como modelos para o entendimento da evolução do género *Avena* (Serenó et al., 1998).

1.2. Sequências repetitivas

O genoma das plantas é maioritariamente constituído por sequências repetitivas presentes em elevado número de cópias, que podem ser altamente conservadas entre a maioria dos organismos ou específicas de uma espécie. Podem encontrar-se dispersas por todo o genoma ou concentradas numa localização específica e esta distribuição é característica para todas as espécies de um género (Katsiotis and Heslop-Harrison, 2000). Estas particularidades tornam as sequências repetitivas excelentes ferramentas para estudos filogenéticos.

O DNA ribossomal é constituído pelas unidades de rDNA que incluem o conjunto de genes que codificam para o RNA ribossomal 45S e a 5S e apresentam localização cromossómica específica e organizadas. As subunidades 45S são constituídas pelas seguintes regiões: *spacer* - 18S – ITS1 – 5,8S – ITS2 – 25S, estando localizadas nos NORs (Nucleolar Organizer Regions) (Richard et al., 2008). As sequências codificantes 18S, 5,8S e 25S estão altamente conservadas entre espécies, ao contrário do que se verifica relativamente à região intergénica (*spacer*) e aos ITSs (Internal Transcribed Spacers) que sendo mais variáveis são muito utilizadas em estudos filogenéticos intraespecíficos (Nikoloudakis et al., 2007). Os ITSs foram utilizados neste estudo para distinguir espécies muito próximas e diferentes variedades da mesma espécie. No entanto, uma vez que a capacidade destas sequências retratarem a verdadeira relação filogenética entre variedades tem sido recentemente questionada (Álvarez and Wendel, 2003; Peng et al., 2010), foram também analisadas sequências específicas do género *Avena* para complementar a sua caracterização.

A sequência pAs120a consiste num fragmento de 389 bp isolado de *A. strigosa* clonada no plasmídeo pUc19. Foi identificada por Linares et al. (1998) que, utilizando hibridação *in situ* fluorescente, demonstraram que esta sequência se encontra dispersa exclusivamente no genoma A de *A. sativa*. Em 1992, Solano et al. obtiveram uma sequência de 313 bp de *A. murphy* introduzida no plasmídeo pUc19 denominada pAm1 que demonstraram por estudos citogenéticos ser específica de genoma C estando dispersa por todos os cromossomas deste genoma.

1.3. Objectivo

O objectivo deste trabalho foi realizar uma caracterização genómica comparativa de diferentes espécies do género *Avena*, com diferentes níveis de ploidia, relativamente a sequências de rDNA 45S e sequências específicas de genomas de *Avena* por análise molecular e citogenética. A análise molecular teve como objectivo distinguir diferentes espécies e variedades de *Avena* e realizar a sua análise filogenética enquanto que o estudo citogenético permitiu identificar o número de NORs e caracterizar a organização interfásica do rDNA e do genoma C.

2. Materiais e métodos

2.1. Material Vegetal

As sementes utilizadas tiveram diferentes proveniências: *A. strigosa* (2n=14, AA) oriunda de Oeiras (Estação Agronómica Nacional, EAN, Oeiras) (strO); *A. sativa* (2n=42, AACDD): linha Kyto originária da Finlândia (USDA – National Small Grains Germplasm Research Facility) (satK), linha 2424 (recolhida de semente de agricultor na ilha da Madeira, gentilmente cedidas pela Eng^a Manuela Veloso, Estação Agronómica Nacional, EAN, Oeiras) (satM), cultivar comercial certificada S. Eulália (gentilmente cedida pela Eng^a Isabel Fernandes, Direcção-Geral de Agricultura e Desenvolvimento Rural, Tapada da Ajuda) (satE), cultivares comerciais S. Mamede e S. José (gentilmente cedidas pelo Eng^o Benvindo Maçãs, Estação Nacional de Melhoramento de Plantas de Elvas); *A. sterilis* (2n=42, AACDD): linha PI267989-1 (steO) proveniente de Oeiras (USDA – National Small Grains Germplasm Research Facility), linha colhida em Elvas (steE) (gentilmente cedida pelo Eng^o José Carlos Costa, ISA-UTL).

2.2. Métodos

As sementes foram colocadas em caixas de Petri com papel de filtro humedecido com água destilada, envoltas em alumínio para evitar a entrada de luz. Para sincronizar a germinação foram submetidas a um tratamento a 4°C durante cerca de 3 dias. Posteriormente, foram colocadas numa câmara de crescimento com as seguintes condições de crescimento: 16h/dia a 22°C e 8h/noite a 15°C.

Ao fim de uma semana e depois de terem sido recolhidos ápices radiculares, as plântulas foram repicadas para *Jiffy Pots*, onde se mantiveram durante um período semelhante tendo sido por fim transferidas para vasos e mantidas em estufa.

2.2.1. Análise molecular por PCR

Extracção de DNA

Folhas novas e saudáveis de plantas adultas com cerca de um mês foram colhidas, envolvidas individualmente em alumínio e colocadas em azoto líquido para um congelamento instantâneo. Seguidamente, todas as amostras foram guardadas a -80°C para posterior utilização.

Procedeu-se à extracção de DNA a partir do material foliar utilizando o kit Citogene DNA Purification Kit da Citomed¹. Utilizando o seguinte protocolo: juntar num almofariz azoto líquido e 10-30mg do material congelado, mexer cuidadosamente e triturar rapidamente as folhas até ficarem em pó. Adicionar 300µl de *cell lysis* ao tecido foliar e incubar a 65°C durante 60 minutos. Adicionar 1,5µl de RNaseA ao lisado celular, agitar no vortex e incubar a 37°C durante 15 minutos. Colocar as amostras à temperatura ambiente, adicionar 100µl de *Protein Precipitation*, homogeneizar as amostras no vortex durante 20 segundos e centrifugar a 13,000-16,000 x g durante 3 minutos. Recuperar o sobrenadante contendo o DNA, adicionar 300µl de isopropanol 100%, misturar as amostras invertendo cuidadosamente os tubos para evitar danificar o DNA e centrifugar a 13,000-16,000 x g durante 1 minuto. Verter o sobrenadante e secar o tubo com o *pellet* contendo o DNA em papel absorvente. Adicionar 300µl de etanol 70% e inverter os tubos várias vezes para lavar o *pellet* de DNA. Centrifugar mais uma vez a 13,000-16,000 x g durante 1 minuto e com cuidado remover o etanol. Inverter os tubos e deixar secar o etanol ao ar durante 10 a 15 minutos. Por fim, ressuspender as amostras de DNA em 50µl de *DNA Hydration* durante a noite à temperatura ambiente.

PCR (Polymerase Chain Reaction)

As reacções de PCR foram realizadas da seguinte forma: para cada reacção de 50µl foram utilizados 5 a 10µl de DNA a 10ng/µl, 1,5µl de MgCl₂ (50mM), 5µl de *buffer* 10xPCR, 0,5µl de dNTP's (0,25mM), 0,5µl de *Taq Polymerase* (5U/µl), 0,5µl de *primer forward* (100pmol/µl) e 0,5µl de *primer reverse* (100pmol/µl) e H₂O até perfazer os 50µl. Os *primers* utilizados para amplificar as diferentes sequências encontram-se listados no Quadro 1. A amplificação decorreu em tubos de 0,2ml em termociclador. As sequências ITS1, ITS2 e pAs120a foram amplificadas com o seguinte programa: 94°C durante 4 min; 34 ciclos de 94°C, 30s; 60°C, 30s e 72°C, 45s; 72°C, 5min e 14°C. O programa utilizado para amplificar a sequência pAm1 foi: 94°C durante 4 min; 29 ciclos de 94°C, 45s; 55°C, 45s e 72°C, 3min; 72°C, 10min e 14°C.

Quadro 1. *Primers* utilizados na análise molecular.

Sequência	<i>Primer Forward</i>	<i>Primer Reverse</i>	Tamanho Esperado	Referência
ITS1 ²	ITS1 TCCGTAGGTGAACCTGCGG	ITS2 GCTGCGTTCATCGATGC	302bp	White et al., 1990
ITS2 ²	ITS3 GCATCGATGAAGAACGCAGC	ITS4 TCCTCCGCTTATTGATATGC	~380bp	White et al., 1990
pAs120a	As120a fow CAACAAGTCTTGCTTGGTT	As120a rev AGGAAGTCGACCGAGGAACT	285 bp	Desenhados no programa primer ³ com base na sequência pAs120 (accession: AJ001922)

¹ Citogene® DNA Cell&Tissue Kit. Refª: CCT-0150; Lot.: 6k780.

² *Primers* desenhados para as sequências codificantes do rDNA

³ Rozen and Skaletsky, 2000

pAm1	(GT)9C GTGTGTGTGTGTGTTGTGTC	Am1 rev TGGATCCACCACTGGTCATT	Am1 Rev desenhado no programa primer3 com base na sequência pAm1 (accession: X83958).
------	---------------------------------------	--	---

Electroforese, captação e tratamento de imagem

Os produtos de PCR obtidos foram analisados por electroforese em gel de agarose 1% a 2%, durante 1-4h e detectados por coloração com brometo de etídio. Seguidamente, os géis foram fotografados usando o BioRad GEL DOC 2000 e as imagens tratadas com o programa Adobe Photoshop CS 5 Extended.

As bandas com o tamanho esperado foram isoladas do gel de agarose, com um bisturi e colocadas em tubos de 1,5ml sendo de seguida purificadas utilizando o High Pure PCR Product Purification Kit da Roche⁴. O protocolo consistiu nos seguintes passos: adicionar 300µl de *binding buffer*, incubar 10 minutos a 56°C vortexando a cada 2 ou 3 minutos; quando a agarose estiver totalmente dissolvida, adicionar 150µl de isopropanol e agitar cuidadosamente; pipetar o conteúdo do tubo de 1,5ml para uma coluna combinada num tubo colector; centrifugar 1 minuto à velocidade máxima da centrífuga de bancada (~15800g); descartar o eluído; fazer duas lavagens com 500µl e 200µl de *wash buffer*, centrifugar 1 minuto à velocidade máxima e descartar o eluído após cada lavagem; para assegurar que a coluna se encontra completamente seca, para que o DNA fique o mais puro possível centrifugar 1 minuto à velocidade máxima sem qualquer tampão; recombinar a coluna com um tubo novo de 1,5ml; adicionar 30µl de *elution buffer* à coluna e centrifugar 1 minuto à velocidade máxima; guardar o DNA purificado a -20°C.

Clonagem

As amostras de DNA purificadas foram clonadas utilizando o TA Cloning® Kit da Invitrogen⁵ segundo o seguinte protocolo: para cada reacção de 10µl de ligação foram utilizados 1µl de 10x *ligation buffer*, 2µl de *pCR® 2.1 vector* (25 ng/µl), 1µl T4 DNA ligase, Xµl de produto de PCR⁶ fresco (menos de 24h após a amplificação por PCR) e H₂O esterilizada até perfazer os 10µl. A ligação decorreu em tubos de 0,2ml em termociclador a 14°C durante a noite. Para cada transformação de células o protocolo foi o seguinte: descongelar em gelo um tubo de 50µl de *One Shot® Competent Cells*⁷; pipetar 2µl da ligação para o tubo de células competentes e misturar cuidadosamente com a ponta da pipeta; incubar em gelo durante 30 minutos; colocar os tubos 30 segundos a 42°C e

⁴ Ref. 11732676001

⁵ Cat. No. K2040-40

⁶
$$X \text{ ng produto de PCR} = \frac{(Y \text{ pares de bases de produto de PCR}) \times (50 \text{ ng pCR}® 2.1 \text{ vector})}{(\text{pares de bases do pCR}® 2.1 \text{ vector} \sim 3900)}$$

⁷ One Shot® TOP10 Chemically Competent E. coli Cat. No. C4040-06.

transferi-los imediatamente para gelo e adicionar 250µl de meio S.O.C.⁸; agitar horizontalmente os tubos a 225rpm durante uma hora a 37°C. Espalhar 20 e 40µl de cada transformação em placas com meio LB⁹, ampicilina (100µg/ml) para selecção de transformantes e X-gal¹⁰ (20µg/ml) para selecção de recombinantes; incubar as placas durante a noite a 37°C; colocar 2-3 horas a 4°C.

Para selecção de transformantes fazer riscado de 5 colónias brancas de cada transformação e verificar por PCR se o fragmento está presente na colónia. As reacções de PCR foram realizadas da seguinte forma: para cada reacção de 20µl foram utilizados 0,6µl de MgCl₂ (50mM), 2µl de *buffer* 10xPCR, 0,2µl de dNTP's (0,25mM), 0,2µl de *Taq Polymerase* (5U/µl), 0,2µl de *primer forward* (100pmol/µl) M13 (5'-TGTAACGAGCCAGT-3') e 0,2µl de *primer reverse* (100pmol/µl) M13 (5'-CAGGAAACAGATGACC-3') 16,6µl de H₂O e como template tocar com um palito autoclavado numa colónia branca e colocar no mix. A amplificação decorreu em tubos de 0,2ml em termociclador com o seguinte programa: 94°C durante 3 minutos; 29 ciclos de 94°C, 30s; 54°C, 30s e 72°C, 90s; 72°C, 3 minutos e 14°C.

Após electroforese, seleccionar colónias que apresentam a banda com o tamanho esperado e colocar a crescer em meio líquido LB¹¹ com ampicilina (100µg/ml).

Para isolar e purificar o vector foi utilizado o Kit NZYMiniprep da Nzytech¹². O protocolo consistiu nos seguintes passos: centrifugar 1-5ml da cultura em meio líquido LB a 8000 x g durante 2 minutos; descartar o sobrenadante e ressuspender o pellet em 250µl de *buffer A1* vortexando vigorosamente; adicionar 250µl de *buffer A2* e inverter o tubo algumas vezes; incubar a temperatura ambiente durante 4 minutos; adicionar 300µl de *buffer A3* e mexer cuidadosamente; centrifugar 10 minutos a 11000 x g; colocar o sobrenadante numa coluna NZYTech, num tubo de 2ml; centrifugar 1 minuto a 11000 x g e descartar a eluição; adicionar 500µl de *buffer AY* à coluna e centrifugar 1 minuto a 11000 x g, descartar o eluído; adicionar 600µl de *buffer A4* à coluna; centrifugar 1 minuto a 11000 x g e descartar o eluído; centrifugar 2 minutos a 11000 x g para secar a coluna; colocar a coluna num tubo de 1,5ml e adicionar 30 a 50µl de *buffer AE*; incubar 1 minuto a temperatura ambiente; centrifugar 1 minuto a 11000 x g.

Fazer uma restrição ao plasmídeo purificado com a *Restriction Endonuclease Eco RI* da Roche¹³ para confirmar os resultados. Visto que a sequência que a enzima reconhece (GAATTC) se

⁸ 2% Triptona; 0,5% Extrato de levedura; 10 mM NaCl; 2,5 mM KCl; 10 mM MgCl₂; 10 mM MgSO₄; 20 mM glucose (dextrose)

⁹ 1,0% Triptona; 0,5% Extrato de levedura; 1,0% NaCl; 15g/L Agar; pH 7,0 autoclavado

¹⁰ Galactose ligada a um composto orgânico aromático

¹¹ 1,0% Triptona; 0,5% Extrato de levedura; 1,0% NaCl; pH 7,0 autoclavado

¹² Cat. No. MB01002

¹³ Cat. No. 10703737001

encontra nas zonas adjacentes ao local de inserção do fragmento no vector, esperou-se a produção do fragmento com o tamanho expectável. A reacção realizou-se da seguinte forma: num tubo de 0,2ml colocar 1µg de DNA; 2,5µl de *10xSuRE/Cut buffer H*; 1 unidade de Eco RI e H₂O até completar 25µl; em termociclador 1h a 37°C e 15 minutos a 65°C

Sequenciação e análise de dados

O DNA purificado foi enviado para a Stabvida¹⁴ para sequenciação. Os dados foram analisados com o programa Bioedit¹⁵. As árvores filogenéticas obtidas com o programa MEGA¹⁶ com o método Neighbor-Joining, análise com 1000 repetições bootstrap. Para análise do padrão de bandas foi utilizado o software NTSYSpc 2.0¹⁷, com o índice de Jaccard¹⁸ no método Neighbor-Joining.

2.2.2. Análise citogenética por FISH

Recolha, tratamento e fixação das raízes

Foram recolhidas raízes de plântulas com cerca de uma semana antes de serem repicadas para *Jiffy Pots* e colocadas em gelo fundente durante 24 horas para obter células c-metafásicas. Seguidamente, as raízes foram fixadas em etanol acético (solução 3:1 (v/v) etanol absoluto/ácido acético) com agitação e renovando o fixador 4 a 5 vezes durante cerca de 24h e posteriormente guardadas a -20°C. Foram também colhidas raízes de plantas adultas com cerca de 4 semanas. Para isso, retiraram-se cuidadosamente as plantas dos vasos para recolher as raízes que foram lavadas para retirar os vestígios de terra e igualmente submetidas ao mesmo tratamento em gelo e posterior fixação.

Esfregaços de ápices radiculares

Para fazer esfregaços de ápices radiculares lavaram-se as raízes em solução tampão 1xEB¹⁹ com agitação durante 2 x 5 minutos com renovação da solução tampão para garantir que o fixador

¹⁴ Campus FCT UNL Edifício Departamental de Química 2825-516 Caparica

¹⁵ Tom Hall Ibis Biosciences

¹⁶ Tamura et al., 2011

¹⁷ Exeter Software

¹⁸ $J_{(A,B)} = \frac{|A \cap B|}{|A \cup B|}$

¹⁹ 1xEB: A – Ácido cítrico monohidratado 0,1 M; B – Citrato trissódico desidratado 0,1 M; 10xEB: adicionar na proporção 2:3 (A:B); diluir 1:10 em água destilada para obter 1xEB

foi todo removido. A solução tampão foi seguidamente substituída por solução enzimática²⁰, seguindo-se uma incubação a 37°C durante 1 hora. Após este período, retirou-se a solução enzimática e adicionou-se nova solução tampão 1xEB.

À lupa, cada ápice radicular foi colocado numa lâmina com uma gota de ácido acético a 60% dispersando o material com o auxílio de agulhas. De seguida, colocou-se uma lamela sobre o material e retirou-se o excesso de ácido com papel de filtro, sendo aplicadas pancadas secas verticalmente sobre a zona do esfregaço com uma agulha para separar as células e desfazer o citoplasma. Finalmente os esfregaços obtidos foram observados ao microscópio de contraste de fase. As lâminas que apresentavam núcleos interfásicos e metafases separadas e com pouco citoplasma foram congeladas utilizando CO₂ para retirar as lamelas com uma lâmina de barbear, deixadas a secar ao ar e guardadas a 4°C até posterior utilização.

Marcação das sondas

A sonda pTa71 (Gerlach and Bedbrook, 1979) foi marcada com digoxigenina por “Nick Translation”. Recorreu-se ao “DIG-Nick Translation Mix” da Roche²¹ que consiste nos seguintes passos: para reacção de 20µl colocar 1µg de DNA, H₂O até perfazer 16µl e 4µl de *DIG-Nick translation Mix*; Incubar durante 90 minutos a 15°C, e para parar a reacção adicionou-se 1µl de EDTA²² 0,5M, seguido de 10 minutos a 60°C.

A sonda pAm1 foi marcada com biotina por PCR. Para cada reacção de 50µl foram utilizados 1µl de DNA a 10ng/µl, 1,5µl de MgCl₂ (50mM), 5µl de *buffer* 10xPCR, 0,5µl de dNTP's (0,25mM), 1µl de dUTP²³, 0,5µl de *Taq Polymerase* (5U/µl), 0,5µl de *primer* universal T7 (5'-TAATACGACTATAGGG-3') e 0,5µl de *primer* universal U19 (5'-GTTTTCCCAGCGACGT-3'), ambos a 100pmol/µl, e H₂O até perfazer os 50µl. A amplificação decorreu em tubos de 0,2ml em termociclador com o seguinte programa: 94°C durante 3 minutos; 29 ciclos de 94°C, 30s; 54°C, 30s e 72°C, 90s; 72°C, 3min e 14°C.

Hibridação *In Situ* Fluorescente (FISH)

Na hibridação foram utilizadas as sondas: pTa71 para rDNA 45S, e pAm1, sequência específica de genoma C.

²⁰ Solução enzimática: 0,3% (p/v) celulase R10, 0,3% pectoliase Y-23 e 0,25% citohelicase (v/v), 0,03% EDTA em 2 x SSC, pH 4.2

²¹ Cat. No. 11745816910

²² ácido etilenodiamino tetra-acético

²³ Biotin – 16 – dUTP Cat. No. 11093070910 Roche

Para reduzir possíveis interações entre as sondas e componentes celulares e facilitar a acessibilidade das sondas e reagentes de detecção, antes da hibridação as lâminas foram sujeitas a pré-tratamentos. Estes consistiram numa fixação com formaldeído 2% em H₂O (100µl/lâmina) durante 20 minutos, para garantir a fixação do material celular, seguida de uma lavagem com 2xSSC²⁴; e incubação em câmara húmida a 37°C com solução de pepsina²⁵ durante 10 minutos, para a digestão do citoplasma e fracção proteica do DNA seguida de lavagem com 2xSSC (2 x 5 minutos) para remoção da pepsina e, por fim, tratamento com solução de RNase²⁶ a 37°C durante 1 hora para degradar o RNA, visto que se pretende efectuar hibridação DNA:DNA, seguida de lavagens com 2xSSC (3x5 minutos). No final dos pré-tratamentos, realizou-se nova fixação com formaldeído 2% em H₂O durante 10 minutos seguida de lavagem com 2xSSC e desidratação com etanol a 70% (v/v) durante 3 minutos e etanol 100% por mais 3 minutos, deixando secar ao ar durante pelo menos 30 minutos.

A mistura de hibridação foi preparada da seguinte forma: 20µl de formamida a 100%; 2µl de 20xSSC; 0,5µl de SDS²⁷ 10%, 1µl de DNA de esperma de salmão (5 µg/µl); 5µl de pTa71 marcado com digoxigenina; 1µl de pAm1 marcado com biotina, 8µl de sulfato de dextrano 50%; 2,5µl H₂O para 40 µl/lâmina.

Para promover a desnaturação da mistura de hibridação, esta foi aquecida a 95°C durante 10 minutos, sendo colocada de seguida em gelo durante 5 minutos, para evitar renaturação. Foram adicionados 40µl da mistura a cada lâmina, cobrindo a região da preparação com uma lamela de plástico. As lâminas foram então colocadas no termociclador e submetidas a um programa que promove a desnaturação a 78°C, durante 10 minutos e posterior descida gradual da temperatura até 37°C, temperatura a que as lâminas foram mantidas durante a noite em câmara húmida.

No dia seguinte, as lâminas foram sujeitas a uma série de lavagens pós-hibridação da seguinte forma: 2xSSC, 3 minutos a 42°C com agitação; formamida a 20% (10ml de formamida para 40ml de 0,1xSSC), 10 minutos a 42°C (lavagem de estringência a 85%); 2xSSC, 2x5 minutos a 42°C; 2xSSC, 2x5 minutos a temperatura ambiente: 4xSSC/Tween²⁸, 2x5 minutos à temperatura ambiente. Posteriormente foi usada uma solução de BSA 5%²⁹, durante 5 minutos à temperatura ambiente para bloquear interferências com o anticorpo utilizado na detecção das sondas. Em seguida, colocaram-se

²⁴ Saline Sodium Citrate: 0,3 M NaCl, 0,03 M citrato de sódio em água destilada, ajustar a pH 7,0

²⁵ Pepsina (actividade 3.200-4.500 unidades/mg de proteína) – Solução “stock” de 500 µg/ml em 0,01 M HCl. Usar uma diluição 1:75 em 0,01 M HCl

²⁶ RNase – Solução “stock” 10 mg/ml em 10 m M Tris.HCl, pH 8, preparada a partir de ribonuclease A, solução em 10 mM Tris.HCl, pH 8 e 50% glicerol. Usar 100mg/ml diluindo 1/100 em 2xSSC

²⁷ dodecil sulfato de sódio

²⁸ 4xSSC 0,2% (v/v) Tween 20 - Preparação de 500 ml: fazer uma diluição 1:5 de 20xSSC em água, adicionar 1 ml de Tween 20 e misturar bem devido à sua viscosidade

²⁹ Solução de BSA 5% - Albumina de soro de bovino (BSA, B-8894, fracção V isenta de globulina, Sigma) em 4xSSC/Tween

50µl/lâmina da mistura de detecção, que contém BSA 5% e anti-digoxigenina conjugada com fluoresceína (1:20, excitação no azul, emissão no verde amarelado, detecção de sondas marcadas com digoxigenina) e estreptavidina conjugada com fluorocromo Cy3 (1:200, excitação no verde, emissão no vermelho, detecção de sondas marcadas com biotina), 1 hora em câmara húmida a 37°C. No final, as lâminas foram lavadas em 4xSSC/Tween, 3x5 minutos. Por último, as lâminas foram sujeitas à coloração do DNA com DAPI³⁰, composto com emissão no azul. Para tal, foram colocados 7µl de DAPI em citifluor³¹ na zona do esfregaço, colocando por fim uma lamela de vidro e removendo o excesso com o auxílio de papel de filtro.

As lâminas foram observadas no microscópio de epifluorescência Zeiss AxioImager Z1 e a captação de imagens foi realizada com o programa axiovision 4.5 e com a câmara digital Zeiss, modelo Axio Cam MRc5. O tratamento de imagens foi efectuado com o programa Adobe Photoshop CS 5 Extended.

3. Resultados

3.1. Análise molecular

3.1.1. rDNA 45s – ITS1 e ITS2

Na Figura 1A está representado o gel de agarose com os produtos de amplificação da região ITS1. A banda com o tamanho esperado (302bp) foi isolada, purificada e reamplificada para posterior clonagem e sequenciação. As sequências obtidas encontram-se alinhadas no anexo 1. A árvore filogenética apresentada na Figura 1B corresponde às sequências ITS1 das diferentes variedades de *Avena* e à sua comparação com outras espécies vegetais, sendo que a informação referente a estas se encontra no anexo 2.

Com esta análise é possível verificar que entre as diferentes variedades de *Avena* estudadas existem variações em apenas 6 nucleótidos nas sequências ITS1, correspondendo a um nível de homologia entre os 98% e os 100% (anexo 3), enquanto que entre as espécies de *Avena* e a sequência obtida a partir do pTa71 (unidade de rDNA isolada de *Triticum aestivum*, Gerlach and Bedbrook, 1979) utilizado como controlo existem 32 nucleótidos variáveis e 7 gaps, correspondendo a um alinhamento de 85% entre *Avena* e *T. aestivum*. A sequência ITS1 obtida a partir do plasmídeo pTa71 apresenta uma homologia de 99% com a sequência de *T. aestivum* publicada (número de acesso: FJ609737.1) o que constitui um excelente controlo interno do estudo efectuado. A análise

³⁰ DAPI (dihidroclorato de 4', 6-diamino-2-fenilindole), solução de 2 µg/ml

³¹ Citifluor AF1 – Solução glicerol/PBS redutora da dissipação de fluorescência

filogenética demonstra que todas as espécies de aveia utilizadas neste estudo se encontram no mesmo grupo sendo que a sequência de ITS1 permite distinguir entre a *A. strigosa* e a *A. sterilis* proveniente de Elvas. Verifica-se ainda a diferença entre as espécies de *Avena* e o controlo referido, sendo a sequência obtida a partir do pTa71 agrupada com *T. aestivum*.

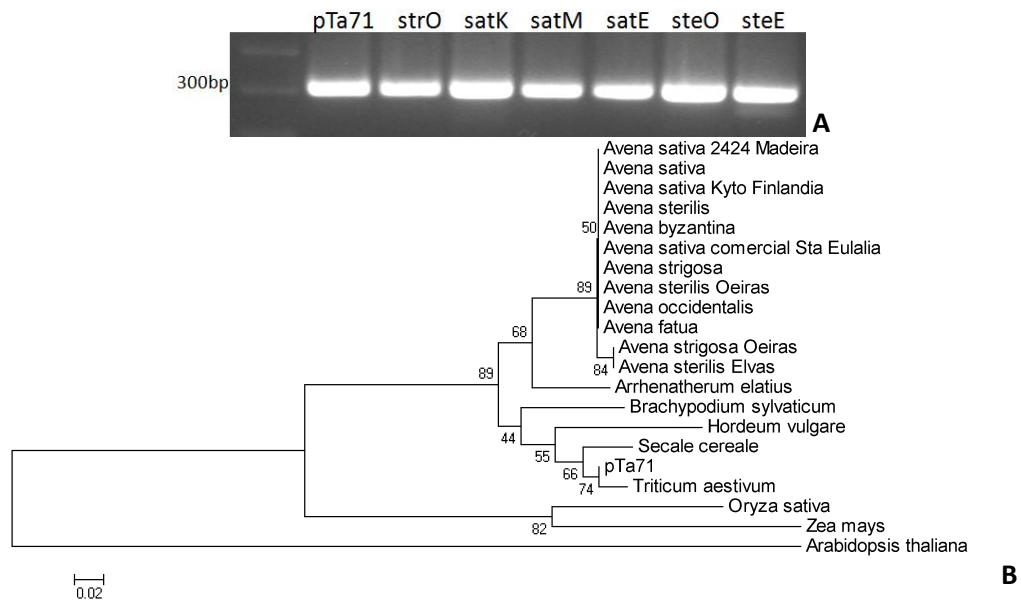


Figura 1. Produto de amplificação (A) e respectiva árvore filogenética (B) da sequência ITS1 obtida de pTa71 (accession number: FJ609737.1) e de diferentes espécies de *Avena*: *A. strigosa* (strO), *A. sativa* (satK, satM e satE) e *A. sterilis* (steO e steE).

Na Figura 2A está representado o gel de agarose com os produtos de amplificação da região ITS2. A banda com o tamanho esperado (380bp) foi isolada, purificada e sequenciada. No anexo 4 é apresentado o alinhamento das sequências obtidas das diferentes variedades de *Avena* comparadas na árvore filogenética da Figura 2B.

A análise das sequências obtidas mostra que existe apenas um nucleótido de diferença entre as sequências do ITS2 obtidas nas diversas variedades de *Avena* estudadas. Após o alinhamento das sequências de *Avena* com o controlo verificou-se que 48 nucleótidos variavam, existindo também 6 gaps, o que corresponde a um alinhamento entre aveias e *T. aestivum* de 78% (anexo 5). No que diz respeito ao ITS2 a homologia obtida entre a sequência amplificada do plasmídeo pTa71 e a sequência publicada é de 98%.

A análise da árvore filogenética obtida para as sequências ITS2 revela que todas as espécies de *Avena* em estudo se mantem agrupadas não sendo no entanto possível distingui-las. A sequência ITS2 obtida do pTa71 controlo encontra-se também claramente separada do grupo das aveias estudadas, estando agrupado com *T. aestivum* assim como se verificou para a ITS1.

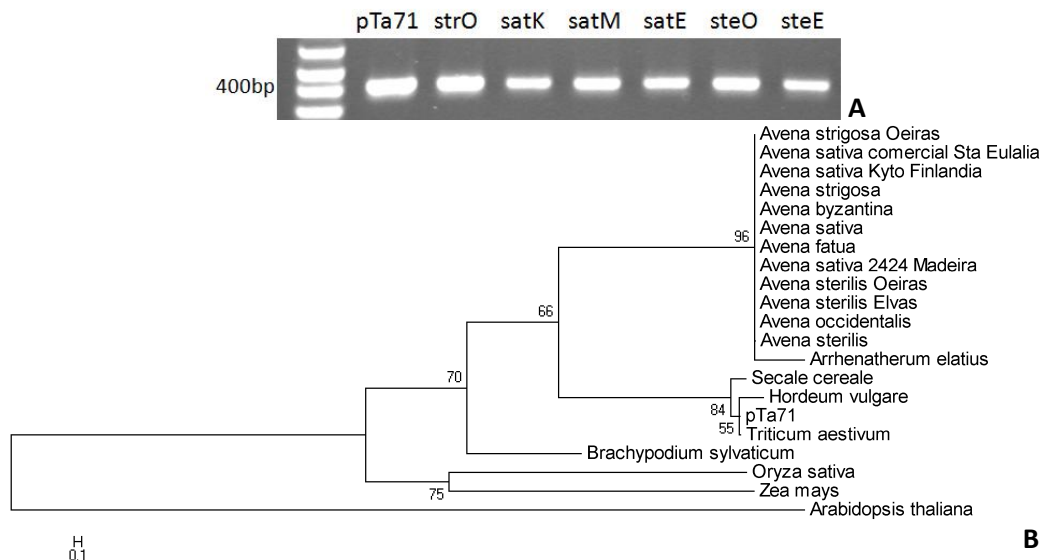


Figura 2. Produto de amplificação (A) e respectiva árvore filogenética (B) da sequência ITS2 obtida de pTa71 (accession number: FJ609737.1) e de diferentes espécies de *Avena*: *A. strigosa* (strO), *A. sativa* (satK, satM e satE) e *A. sterilis* (steO e steE).

3.1.2. Sequências específicas de *Avena*

Sequência específica do genoma A – pAs120a

Na Figura 3A está a amplificação da sequência pAs120a específica do genoma A. A banda com o tamanho esperado (285 bp) foi isolada, purificada e clonada para posterior sequenciação. No anexo 6 encontra-se o alinhamento das sequências pAs120a que permitiram obter a árvore filogenética da Figura 3B.

O alinhamento revelou um nível de homologia entre as diferentes variedades de aveia estudadas entre 91% a 96% (anexo 7) verificando-se existirem alterações em 45 posições. A sequência controlo, amplificação do fragmento de 389 bp isolado de *A. strigosa* clonada no plasmídeo pUc19 (gentilmente cedido por A. Fominaya, Departamento de Biologia Celular e Genética, Universidad de Alcalá (UAH) – Madrid, Espanha), foi também alinhada com a sequência pAs120 (accession: AJ001922) da qual foi obtida, e verificou-se um alinhamento de 100%.

A análise da árvore filogenética permite verificar que as variedades de espécie *A. sterilis* se encontram num grupo distinto de *A. strigosa* e que as variedades de *A. sativa* se apresentam mais próximas da sequência controlo pAs120a.

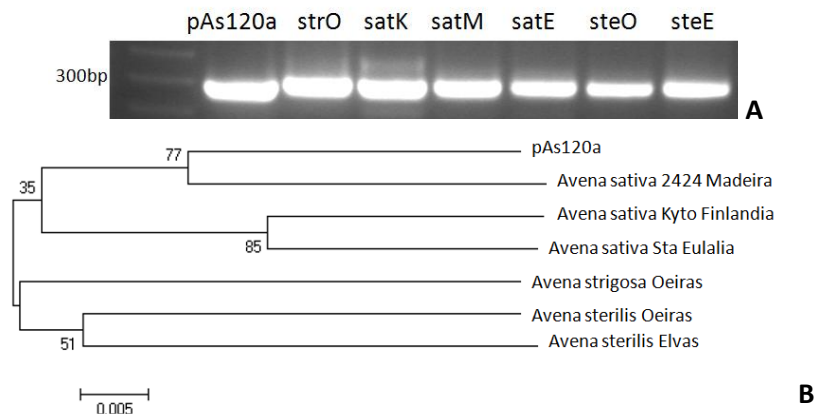


Figura 3. Produto de amplificação (A) e respectiva árvore filogenética (B) da sequência pAs120a obtida de pAs120 (accession number: AJ001922) e de diferentes espécies de Avena: *A. strigosa* (strO), *A. sativa* (satK, satM e satE) e *A. sterilis* (steO e steE).

Sequência específica do genoma C – pAm1

Ao amplificar a sequência pAm1 com os *primers* desenhados *Am1 forward* e *reverse* não se obteve uma banda única com o tamanho espectável (167bp), como se pode verificar no gel presente no anexo 8. Assim, para analisar esta sequência utilizou-se o *primer Am1 reverse* desenhado conjuntamente com o *primer* para o micro satélite (GT)⁹C, obtendo-se desta forma o padrão de bandas presente na Figura 4A. A presença de bandas com tamanho entre os 100bp e os 3000bp foi utilizada para a construção de uma matriz binária onde 1 corresponde à presença e 0 á ausência de bandas. Essa informação foi analisada utilizando o *software* NTSYS-pc que permitiu a obtenção da árvore filogenética apresentada na Figura 4B.

Analisando a árvore filogenética verifica-se que existe uma distinção clara entre a *A. strigosa* e as restantes espécies da *Avena*, que a linha de aveia proveniente de Elvas está também isolada, estando as várias linhas de *A. sativa* estudadas e a linha de *A. sterilis* originária de Oeiras mais próximas.

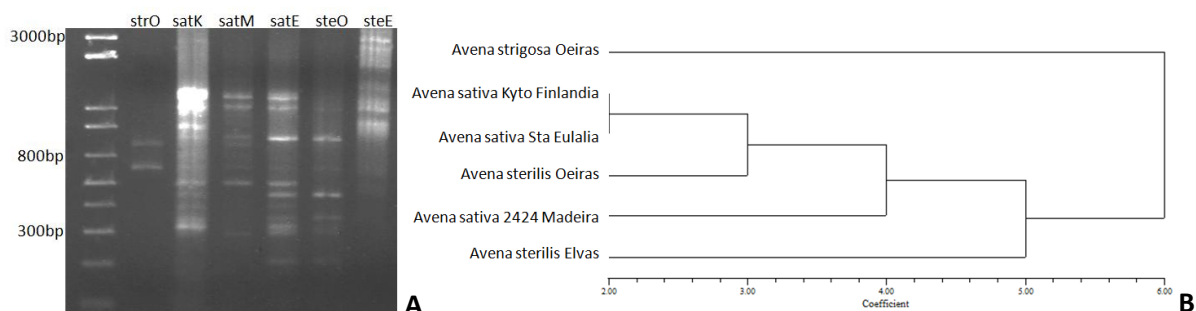


Figura 4. Padrão de bandas obtido com *primers* (GT)⁹C e Am1 rev (A) e respectiva árvore filogenética (B) de diferentes espécies de Avena: *A. strigosa* (strO), *A. sativa* (satK, satM e satE) e *A. sterilis* (steO, steE).

3.2. Análise citogenética

Esta análise teve como objectivos caracterizar o número de NORs em células c-metafásicas de *A. strigosa*, *A. sativa* e *A. sterilis* assim como estudar a organização interfásica do rDNA 45S e do genoma C.

Avena strigosa

Na figura 5 encontra-se um núcleo interfásico de *A. strigosa* hibridado com a sonda pTa71 onde se verifica que esta espécie possui quatro marcações de pTa71. Quando a sonda pAm1 específica de genoma C foi utilizada nesta espécie não houve hibridação.

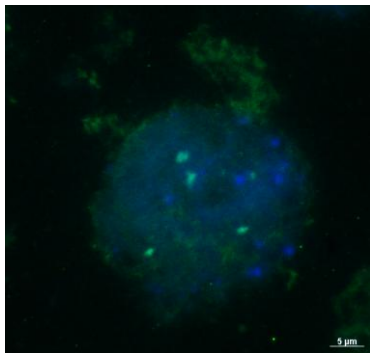


Figura 5. Célula interfásica de *A. strigosa* com marcação da sonda pTa71 (verde) e coloração DAPI (azul).

Avena sativa

Em células c-metafásicas da cultivar comercial de *A. sativa* cv. São Mamede, como se pode observar na célula apresentada na Figura 6A, foi possível observar que 14 cromossomas hibridaram integralmente com a sonda pAm1, correspondente ao genoma C. Além disso, verifica-se que existem igualmente regiões terminais dos restantes cromossomas marcadas com essa sonda. Relativamente à hibridação utilizando como sonda pTa71, verifica-se que existem 6 marcações distribuídas por cromossomas que não apresentam hibridação com a sonda pAm1.

Utilizando a sonda pAm1 foi ainda possível estudar a organização do genoma C em núcleos interfásicos da cultivar *A. sativa* São José como se verifica na Figura 6B. Assim, foi possível observar que os cromossomas do genoma C apresentam uma distribuição radial, estando os telómeros localizados na periferia do núcleo, enquanto as regiões centroméricas se parecem posicionar no centro do núcleo.

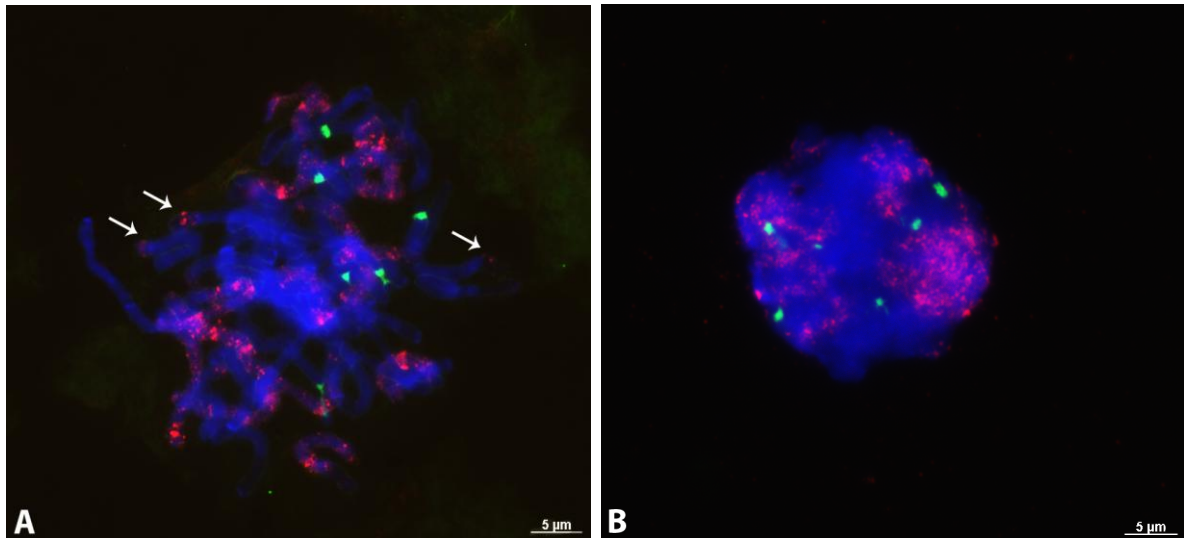


Figura 6. Célula metafásica de *A. sativa* cv. São Mamede (A) e célula interfásica de *A. sativa* cv. São José (B) com marcação da sonda pTa71 (verde), pAm1 (vermelho) e coloração DAPI (azul). Setas indicam cromossomas parcialmente hibridados com pAm1.

Avena sterilis

A Figura 7A permite verificar que existem 6 marcações de pTa71 na linha proveniente de Oeiras de *A. sterilis*, tendo sido possível diferenciar os 3 pares de cromossomas apresentando marcações com pTa71 com base na diferente organização da cromatina observada: um par de NORs apresenta um único bloco heterocromático e dois pares apresentam dois blocos de cromatina condensada separados por uma fracção mais descondensada de cromatina, sendo que num dos pares esses blocos apresentam dimensões semelhantes e o outro apresenta blocos com diferentes dimensões. No entanto, na *A. sterilis* proveniente de Elvas o número de sinais de hibridação *in situ* com a sonda pTa71 é 8 (Figura 7B e 7C), não tendo sido possível distinguir diferenças morfológicas entre os sinais de hibridação obtidos. Quanto à distribuição de pAm1 verificou-se que este marca igualmente 14 cromossomas na linha de *A. sterilis* proveniente de Oeiras (Figura 7A.)

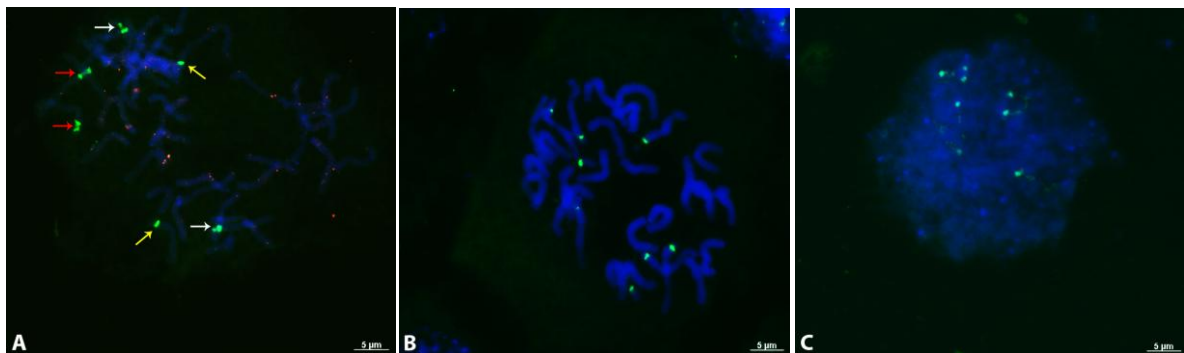


Figura 7. (A) Célula metafásica da linha de *A. sterilis* proveniente de Oeiras com marcação das sondas pTa71 (verde), pAm1 (vermelho) e coloração com DAPI (azul), diferentes pares de NORs indicados por setas com diferentes cores. Células metafásica e interfásica da linha de *A. sterilis* proveniente de Elvas (B e C respectivamente), com coloração DAPI (azul) e marcação da sonda pTa71 (verde).

4. Discussão e perspectivas futuras

A análise molecular desenvolvida neste trabalho permitiu identificar polimorfismos em diferentes sequências repetitivas entre as diferentes espécies / variedades de *Avena* analisadas e ainda efectuar a sua análise filogenética. Em relação às sequências de rDNA 45S estudadas foi possível verificar que a sequência do ITS1 não é tão conservada como a do ITS2, uma vez que entre as diferentes espécies da *Avena* estudadas, o ITS1 varia em 5 posições e o ITS2 apenas em uma. Apesar de as sequências de ITS1 serem muito semelhantes, o seu estudo permitiu distinguir algumas variedades analisadas, ao contrário do que se verifica em relação ao ITS2. A análise das árvores filogenéticas obtidas para estas sequências revela que as diversas variedades da mesma espécie nem sempre se encontram no mesmo grupo. Os estudos de evolução do género *Avena* realizados até à data com as sequências ITS mostram que estas sequências são úteis na diferenciação de genomas podendo, no entanto, não ser as melhores ferramentas para avaliar a diversidade intraespecífica, pelo que se torna necessário a confirmação dos resultados obtidos através do estudo de outras sequências (Nikoloudakis et al., 2007; Peng et al., 2010).

Relativamente às sequências repetitivas características de *Avena* - pAs120a e pAm1 - verificou-se que a sequência específica de genoma A (pAs120a) é a que permite fazer uma melhor distinção das espécies na análise filogenética uma vez que é a mais variável. Assim, com base nesta sequência foi possível agrupar as duas *A. sterilis* em estudo, diferenciar *A. strigosa* e verificar que as variedades analisadas de *A. sativa* são muito semelhantes entre si e estão também mais próximas da sequência utilizada como controlo, pAs120a isolada de *A. strigosa* (Linares et al., 1998). A amplificação da sequência pAm1 específica de genoma C com *primers* desenhados com base na sequência publicada (número de acesso: X83958) não produziu uma banda única tendo sido efectuado o estudo comparativo do padrão de bandas obtido com o *primer reverse* conjuntamente com um *primer* ancorado para micro satélite (GT)⁹C através da construção de matrizes binárias. Esta análise revelou igualmente que *A. strigosa* se encontra mais distante das outras espécies de *Avena*, o que seria de esperar visto que esta sequência não foi descrita no genoma A. Verificou-se ainda que a linha de *A. sterilis* proveniente de Oeiras é mais próxima das linhas de *A. sativa* estudadas comparativamente com a outra linha de *A. sterilis*, que se encontra também isolada. Estas sequências são muito utilizadas para identificar genomas do género *Avena* com a técnica de FISH (Linares et al., 1998; Irigoyen et al., 2002) mas até à data não existem estudos filogenéticos publicados que as utilizem numa análise molecular, pelo que os dados obtidos neste trabalho são bastante promissores e potenciam a sua utilização futura como marcadores moleculares úteis na caracterização inter e intraespecífica no género *Avena*.

A análise citogenética realizada permitiu identificar 4 marcações de pTa71 em *A. strigosa* como descrito por Katsiotis and Heslop-Harrison, 1997. Nesta espécie verificou-se que a sonda pAm1 não hibridou, o que era esperado visto que a sonda utilizada é específica de genoma C (Solano et al., 1992) e a *A. strigosa* tem genoma A.

Em *A. sativa* o número de marcações observadas nas variedades estudadas neste trabalho revelou a presença de 6 marcações de pTa71. Este resultado está de acordo com descrições anteriores relativamente à caracterização de distintas cultivares de *A. sativa* (Fominaya et al., 1995, Linares et al., 1996, 1998 e 2000, Irigoyen et al., 2002). A hibridação com a sonda pAm1 permitiu identificar os cromossomas do genoma C, revelando que as marcações de pTa71 não estão presentes neste genoma, como foi proposto por Irigoyen et al., 2002. Em núcleos interfásicos de *A. sativa*, para além dos blocos heterocromáticos que parecem apresentar uma localização perinucleolar, foi possível observar descondensações do rDNA coincidentes com a região nucleolar com menor intensidade da coloração DAPI. A organização do rDNA 45S nunca tinha sido descrita em espécies de *Avena*, assemelhando-se a padrões de organização anteriormente descritos em centeio (Caperta et al., 2002).

O estudo citogenético realizado em *A. sterilis* permitiu verificar que, nesta espécie, o número de NORs varia de acordo com a variedade estudada. Na *A. sterilis* proveniente de Oeiras identificaram-se 6 NORs, enquanto que na linha proveniente de Elvas se observaram 8 NORs dos quais 6 apresentam comparativamente maiores dimensões, como descrito por Fominaya et al., 1995. Adicionalmente foi possível diferenciar os pares de NORs na linha *A. sterilis* proveniente de Oeiras com base nos seus padrões de organização (um dos pares apenas um bloco heterocromático, outro par com dois blocos heterocromáticos do mesmo tamanho e um par de NORs que apresenta dois blocos de diferentes dimensões). A caracterização dos blocos heterocromáticos desta linha de *A. sterilis* ainda não tinha sido descrita, parecendo assemelhar-se à observada por Linares et al., 1992 em *A. byzantina*, uma espécie hexaplóide de genoma AACDD, contrastando com a organização dos NORs de *A. sativa* que dificilmente possibilita a sua distinção (Irigoyen et al., 2002). No entanto, Varela 2011 descreveu em células prometafásicas um padrão de organização semelhante para os 3 pares de *Nor loci* de maiores dimensões descritos numa linha de *A. sterilis* proveniente de Elvas.

Utilizando a sonda pAm1 foi possível observar 14 cromossomas totalmente marcados em *A. sativa*, existindo ainda cromossomas em que a sonda híbrida exclusivamente nas regiões teloméricas, como se pode ver na Figura 6B. Irigoyen et al. 2002 propôs que durante a evolução das espécies hexaplóides existiram translocações entre os vários genomas o que explicaria a presença de cromossomas com zonas terminais hibridadas com pAm1. Fominaya et al., 1995 observou que *A. sterilis* hibridada com pAm1 marcava 14 cromossomas quase totalmente e três pares dos genomas

A/D também hibridavam nas zonas finais do braço longo. Neste estudo foi possível confirmar a presença de 14 cromossomas do genoma C utilizando a sonda específica pAm1 apenas na linha de *A. sterilis* proveniente de Oeiras. Por outro lado, a análise dos resultados obtidos com esta sonda em núcleos interfásicos de *A. sativa* permitiu caracterizar pela primeira vez a topologia nuclear em espécies do género *Avena*, parecendo indicar que os cromossomas do genoma C apresentam uma distribuição radial, estando os telómeros localizados na periferia do núcleo, enquanto as regiões centroméricas se parecem posicionar no centro. Esta organização é bastante distinta da normalmente observada em outros cereais em que os telómeros e os centrómeros se localizam em regiões opostas no núcleo (Aragon-Alcaide et al., 1996; Dong and Jiang, 1998, Delgado et al., 2010).

Com este trabalho foi possível verificar que sequências repetitivas são extremamente úteis para a caracterização de diferentes espécies do género *Avena*, sobretudo as sequências características de *Avena*. Assim, a caracterização genómica das diferentes variedades de *Avena* efectuada neste trabalho constitui uma excelente base para a definição de marcadores moleculares que permitam caracterizar inequivocamente diferentes variedades portuguesas de aveia, úteis na pesquisa de variedades com características agronómicas mais adaptadas a condições ambientais adversas, assim como na confirmação da constituição genotípica de híbridos entre diferentes espécies.

A análise citogenética permitiu confirmar os resultados publicados em relação ao número de NORs das espécies estudadas e à dispersão da sequência específica de genoma C, pAm1 (Fominaya et al., 1995; Linares et al., 1996, 1998 e 2000, Irigoyen et al., 2002), tendo possibilitado igualmente obter importantes pistas relativamente aos padrões de organização da cromatina ribossomal assim como à topologia nuclear em diferentes espécies de *Avena*. No entanto, no futuro pretende-se confirmar os diferentes parâmetros da organização interfásica descrita através da análise simultânea por hibridação *in situ* e imunolocalização ou coloração com nitrato de prata de nucléolos em análise em núcleos interfásicos com a estrutura tridimensional preservada.

5. Referências bibliográficas:

- **Álvarez, I., Wendel, J.** (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29: 417-434;
- **Aragon-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S., Moore, G.** (1996). A cereal centromeric sequence. *Chromosoma* 105:261-268.
- **Caperta, A.D., Neves, N., Morais-Cecilio, L., Malho, R., Viegas, W.** (2002) Genome restructuring in rye affects the expression, organization and disposition of homologous rdna loci. *Journal of Cell Science* 115:2839-2846.
- **Delgado, M., Frescatada, M., Ribeiro-Varandas, E., Viegas, W., Jones, R.** (2010). Dynamics of functional heterochromatic domains in the plant interphase nucleus. *Cytogenet Genome Res* 129:64-71.

- **Dong, F., Jiang, J.** (1998). Non-rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. *Chromosome Research* 6:551-558.
- **Fominaya, A., Hueros, G., Loarce, Y., Ferrer, E.** (1995). Chromosomal distribution of a repeated DNA sequence from C-genome heterochromatin and the identification of a new ribosomal DNA locus in the *Avena* genus. *Genome*, 38: 548-557;
- **Gerlach, W.L., Bedbrook, J.R.** (1979). Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Research* 7: 1869-1885;
- **Irigoyen, M.L., Linares, C., Ferrer, E., Fominaya, A.** (2002). Fluorescence *in situ* hybridization mapping of *Avena sativa* L. cv. SunII and its monosomic lines using cloned repetitive DNA sequences. *Genome* 45: 1230–1237;
- **Katsiotis, A., Heslop-Harrison, J.** (1997). The close relationship between the A and B genomes in *Avena* L. (Poaceae) determined by molecular cytogenetic analysis of total genomic, tandemly and dispersed repetitive DNA sequences. *Annals of Botany* 79: 103-109;
- **Katsiotis, A., Loukas, M., Heslop-Harrison, J.** (2000). Repetitive DNA, Genome and Species Relationships in *Avena* and *Arrhenatherum* (Poaceae). *Annals of Botany* 86: 1135-1142;
- **Linares, C., Vega, C., Ferrer, E., Fominaya, A.** (1992). Identification of C-banded chromosomes in meiosis and the analysis of nucleolar activity in *Avena byzantina* C. Koch cv 'Kanota'. *Theor Appl Genet*, 83: 650-654;
- **Linares, C., Gonzalez, J., Ferrer, E., Fominaya, A.** (1996). The use of double fluorescence *in situ* hybridization to physically map the positions of 5s rDNA genes in relation to the chromosomal location of 18s-5.8s-26s rDNA and a C genome specific DNA sequence in the genus *Avena*. *Genome* 130(2): 535-541;
- **Linares, C., Ferrer, E., Fominaya, A.** (1998). Discrimination of the closely related A and D genomes of the hexaploid oat *Avena sativa* L. *Proc Natl Acad Sci USA*, 95: 12450-12455;
- **Linares, C., Irigoyen, M., Fominaya, A.** (2000). Identification of C-genome chromosomes involved in intergenomic translocations in *Avena sativa* L., using cloned repetitive DNA sequences. *Theor Appl Genet*, 100: 353-360;
- **Loskutov, I.** (2008). On evolutionary pathways of *Avena* species. *Genet Resour Crop Evol* 55: 211-220;
- **Nikoloudakis, N., Skaracis, G., Katsiotis, A.** (2008). Evolutionary insights inferred by molecular analysis of the ITS1 – 5.8S – ITS2 and IGS *Avena* sp. Sequences. *Molecular Phylogenetics and Evolution*, 46: 102-115;
- **Peng, Y., Baum, B., Ren, C., Jiang, Q., Chen, G., Zheng, Y., Wei, Y.** (2010). The evolutionary pattern of rDNA ITS in *Avena* and phylogenetic relationship of the *Avena* species (Poaceae: Avenae). *Hereditas*, 147: 183-204;
- **Richard, G., Kerrest, A., Dujon, B.** (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiology and Molecular Biology Reviews*, 72, n.4: 686-727;
- **Rozen, S., Skaletsky, H.** (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386;
- **Sereno, M., Bodanese-Zanettini, M., Carvalho, F., Neto, J.** (1998). Variabilidade genética em *Avena sativa*, *Avena sterilis* e em seus híbridos, por meio de avaliações agronômicas. *Pesquisa Agropecuária Brasileira*, Brasília, 33, n.10: 1601-1607;
- **Solano, R., Hueros, G., Fominaya, A., Ferrer, E.** (1992). Organization of repeated sequences in species of the genus *Avena*. *Theor Appl Genet*, 83: 602-607;
- **Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S.** (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, 28: 2731-2739;
- **Tavares, M., Zanettini, M., Carvalho, F.** (1993). Origem e Evolução do Género *Avena*: suas implicações no melhoramento genético. *Pesquisa Agropecuária Brasileira*, Brasília, 28, n.4: 499-507;
- **Varela, A.M.S.** (2011). Avaliação da diversidade genômica de diferentes espécies do género *Avena*, Projecto do 1º Ciclo em biologia ISA, UTL.
- **White, T., Bruns, T., Lee, S., Taylor, J.** (1991). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide to Methods and Applications*: Academic Press, Inc., pp. 315-322.

Anexos

Anexo 1 – Alinhamento das sequências ITS1 obtidas de pTa71 e de diferentes espécies de

Avena. (As posições variáveis em *Avena* encontram-se sublinhadas a azul).

```

      .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                10          20          30          40          50
pTa71  TGACCCTGAC CAAAACAGAC CGCGCACGCG TCATCCAATC C--GTCGGCG
strO   TGACCCTGAC CAAAACAGAC CGAGCACGCG TTATCTATTC CTGCTGAGTG
satK   TGACCCTGAC CAAAACAGAC CGAGCACGCG TTATCTATTC CTGCTGAGTG
satM   TGACCCTGAC CAAAACAGAC CGAGCACGCG TTATCTATTC CTGCTGAGTG
satE   TGACCCTGAC CAAAACAGAC CGAGCACGCG TTATCTATTC CTACTGAGTG
steO   TGACCCTGAC CAAAACAGAC CGAGCACGCG TTATCTATTC CTACTGAGTG
steE   TGACCCTGAC CAAAACAGAC CGAGCACGCG TTATCTATTC CTGCTGAGTG

```

```

      .....|.....| .....|.....| .....|.....| .....|.....|
                60          70          80          90          100
pTa71  ACGGCACCGT CCGTCGCTCG GCCAATGCCT CGACCACCTC CCCTCCTCGG
strO   GCGGCAACGT C-GTTGCTCA GCCAAATCCT CGATAACCTC CTCTCCTTGG
satK   GCGGCACCGT C-GTCGCTCA GCCAAATCCT CGATAACCTC CTCTCCTTGG
satM   GCGGCACCGT C-GTCGCTCA GCCAAATCCT CGATGACCTC CTCTCCTTGG
satE   GCGGCACCGT C-GTCGCTCA GCCAAATCCT CGATAACCTC CTCTCCTCGG
steO   GCGGCACCGT C-GTCGCTCA GCCAAATCCT CGATAACCTC CTCTCCTCGG
steE   GCGGCAACGT C-GTTGCTCA GCCAAATCCT CGATAACCTC CTCTCCTTGG

```

```

      .....|.....| .....|.....| .....|.....| .....|.....|
                110         120         130         140         150
pTa71  AGCGGGTGGG GGCTCGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG
strO   AG----TGGG GGCTTGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG
satK   AG----TTGG GGCTTGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG
satM   AG----TTGG GGCTTGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG
satE   AG----TGGG GGCTTGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG
steO   AG----TGGG GGCTTGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG
steE   AG----TGGG GGCTTGGGGT AAAAGAACCC ACGGCGCCGA AGGCGTCAAG

```

```

      .....|.....| .....|.....| .....|.....| .....|.....|
                160         170         180         190         200
pTa71  GAACACTGTG CCTAACCCGG GGGCATGTCT AGCTTGCTAG CCGTCCCCTG
strO   GAACACTGTG CCTAGCTTGG GGACACGACT GGCTTGCTGG CCGTCCCCT
satK   GAACACTGTG CCTAGCTTGG GGACACGACT GGCTTGCTGG CCGTCCCCT
satM   GAACACTGTG CCTAGCTTGG GGACACGACT GGCTTGCTGG CCGTCCCCT
satE   GAACACTGTG CCTAGCTTGG GGACACGACT GGCTTGCTGG CCGTCCCCT
steO   GAACACTGTG CCTAGCTTGG GGACACGACT GGCTTGCTGG CCGTCCCCT
steE   GAACACTGTG CCTAGCTTGG GGACACGACT GGCTTGCTGG CCGTCCCCT

```

```

      .....|.....| .....|.....| ...
                210         220
pTa71  TGTTGCAAAG CTATTTAATC CAC
strO   TGCTGCAAAG CTATTTAATC CAC
satK   TGCTGCAAAG CTATTTAATC CAC
satM   TGCTGCAAAG CTATTTAATC CAC
satE   TGCTGCAAAG CTATTTAATC CAC
steO   TGCTGCAAAG CTATTTAATC CAC
steE   TGCTGCAAAG CTATTTAATC CAC

```

Anexo 2 – Números de acesso das sequências das espécies vegetais utilizadas na análise filogenética.

Espécie	Número de acesso do GenBank
<i>Arabidopsis thaliana</i>	X52322
<i>Arrhenatherum elatius</i> subsp. sardoum	AJ632248.1
<i>Avena byzantina</i>	DQ092755.1
<i>Avena fatua</i>	DQ092756.1
<i>Avena occidentalis</i>	EU833781.1
<i>Avena sativa</i>	AY520821.1
<i>Avena sterilis</i>	EU833810.1
<i>Brachypodium sylvaticum</i>	AJ608155.1
<i>Hordeum vulgare</i>	FJ593180.1
<i>Oryza sativa</i>	AF169230.1
<i>Secale cereale</i>	JF489233.1
<i>Triticum aestivum</i>	FJ609737.1
<i>Zea mays</i> subsp. mexicana	AF019817.1

Anexo 3 – Homologia das sequências ITS1 em percentagem.

%	pTa71	strO	satK	satM	satE	steO	steE
pTa71	100						
strO	85	100					
satK	85	98	100				
satM	85	98	99	100			
satE	86	98	98	98	100		
steO	86	98	98	98	100	100	
steE	85	100	98	98	98	98	100

Anexo 4 – Alinhamento das sequências ITS2 obtidas de pTa71 e de diferentes espécies de *Avena*. (As posições variáveis em *Avena* encontram-se sublinhadas a azul).

	10	20	30	40	50	
pTa71	CGCCAAAACA	CGCTCCCAAC	CACCCTCAAC	GGGATTCGGG	ATGCGGCATC	
strO	CGCTAAA-CA	CGCTCCCAAC	C--CCTTACG	GGGGAACAGG	ATGCGGCATT	
satK	CGCTAAA-CA	CGCTCCCAAC	C--CCTTACG	GGGGAACAGG	ATGCGGCATT	
satM	CGCTAAA-CA	CGCTCCCAAC	C--CCTTACG	GGGGAACAGG	ATGCGGCATT	
satE	CGCTAAA-CA	CGCTCCCAAC	C--CCTTACG	GGGGAACAGG	ATGCGGCATT	
steO	CGCTAAA-CA	CGCTCCCAAC	C--CCTTACG	GGGGAACAGG	ATGCGGCATT	
steE	CGCTAAA-CA	CGCTCCCAAC	C--CCTTACG	GGGGAACAGG	ATGCGGCATT	

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          60          70          80          90          100
pTa71    TGGTCCCTCG TCTCTCAAGG GACGGTGGAC CGAAGATTGG GCTGCCGGCG
strO     TGGCTCCCCG TCACCCAAGG G-CGGTGGGC CGAAGATATG GCTGCCGGCG
satK     TGGCTCCCCG TCACCCAAGG G-CGGTGGGC CGAAGATATG GCTGCCGGCG
satM     TGGCTCCCCG TCACCCAAGG G-CGGTGGGC CGAAGATATG GCTGCCGGCG
satE     TGGCTCCCCG TCACCCAAGG G-CGGTGGGC CGAAGATATG GCTGCCGGCG
steO     TGGCTCCCCG TCACCCAAGG G-CGGTGGGC CGAAGATATG GCTGCCGGCG
steE     TGGCTCCCCG TCACCCAAGG G-CGGTGGGC CGAAGATATG GCTGCCGGCG

```

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          110         120         130         140         150
pTa71    TACCGCGCCG GACACAGCGC ATGGTGGGCG TCCTCGCTTT A-TCAATGCA
strO     CATCGTGTCTG GACACAGCGC GTGGTGAGCG TCCTCGCTAT ACTTACCGCA
satK     CATCGTGTCTG GACACAGCGC GTGGTGAGCG TCCTCGCTAT ACTTACCGCA
satM     CATCGTGTCTG GACACAGCGC GTGGTGAGCG TCCTCGCTAT ACTTACCGCA
satE     CATCGTGTCTG GACACAGCGC GTGGTGAGCG TCCTCGCTAT ACTTACCGCA
steO     CATCGTGTCTG GACACAGCGC GTGGTGAGCG TCCTCGCTAT ACTTACCGCA
steE     CATCGTGTCTG GACACAGCGC GTGGTGAGCG TCCTCGCTAT ACTTACCGCA

```

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          160         170         180         190         200
pTa71    GTGCATCCCG CGCGCAGCTG GCATTATGGC CTTTGAACGA CCCAACAAAC
strO     GTGTCTCCGA CACGTAGCCG GCGTGATGGC CTA-GAATGA CCCTTGTAAC
satK     GTGTCTCCGA CACATAGCCG GCGTGATGGC CTA-GAATGA CCCTTGTAAC
satM     GTGTCTCCGA CACGTAGCCG GCGTGATGGC CTA-GAATGA CCCTTGTAAC
satE     GTGTCTCCGA CACATAGCCG GCGTGATGGC CTA-GAATGA CCCTTGTAAC
steO     GTGTCTCCGA CACATAGCCG GCGTGATGGC CTA-GAATGA CCCTTGTAAC
steE     GTGTCTCCGA CACATAGCCG GCGTGATGGC CTA-GAATGA CCCTTGTAAC

```

```

.....|.....| .....|.....| .....
          210         220
pTa71    AAAGCGCACG TCGCTTCGAC CGC
strO     GGTGCGCATG ACGCTCCGAC CGC
satK     GGTGCGCATG ACGCTCCGAC CGC
satM     GGTGCGCATG ACGCTCCGAC CGC
satE     GGTGCGCATG ACGCTCCGAC CGC
steO     GGTGCGCATG ACGCTCCGAC CGC
steE     GGTGCGCATG ACGCTCCGAC CGC

```

Anexo 5 – Homologia das sequências ITS2 em percentagem.

%	pTa71	strO	satK	satM	satE	steO	steE
pTa71	100						
strO	78	100					
satK	78	99	100				
satM	78	100	99	100			
satE	78	99	100	99	100		
steO	78	99	100	99	100	100	
steE	78	99	100	99	100	100	100

Anexo 6 – Alinhamento das sequências pAs120a obtidas de pAs120 e de diferentes espécies de *Avena*. (As posições variáveis em *Avena* encontram-se sublinhadas a azul).

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          10          20          30          40          50
pAs120a  CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACAATTGTA CCTGACTAGT
strO     CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACCATTGTA CCTGACTAGT
satK     CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACAATTGTA TCTGACTAGT
satM     CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACCATTGTA TCTGACTAGT
satE     CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACCATTGTA TCTGACTAGT
steO     CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACAATTGTA TCTGACTAGT
steE     CAACAACTGC TTGCTTGGTT TTGTCGCTTGC GACAATTGTA TCTGACTAGT

```

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          60          70          80          90          100
pAs120a  AAGTATTGCA ACTTCCACGG AAATGGCTGAA TAAAAGTACC AACAACTG-T
strO     AATTATTGGCA ACTACAACGG AAATGGCTGAA TAAAAGTACC AACAACTGC
satK     AAGTATTGCA ACTACAACGG AAATGGCTAAA TAAAAGTACC AACAACTGC
satM     AAGTATTGCA ACTACAACGG AAATGGCTGAA TAAAAGTACC AACAACTGC
satE     AATTATTGGCA ACTACAACGG AAATGGCTAAA TAAAAGTACC AACAACTGC
steO     AAGTATTGCA ACTACAATGG AAATGGCTAAA TAAAAGTACC AACAACTGT
steE     AAGTATTGCA ACTACAACGG AAATGGCTAAA TAAAAGTACC AACAACTGC

```

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          110         120         130         140         150
pAs120a  GGGTTTT--G GGTTTATCTC ATATCTCTG TACTCGATTA GTAATTGTTG
strO     GGGTTTT--T GGGTTATCTC AATGCTAATGG TATCTGACTA GTAAGTGTTG
satK     GGGTTTT--T GGGATATCTC AATGCTATCTA TACCTGATTA GTAAGTGTTG
satM     GGGTTTTTG GGTTTAACTC AATGTTATCCG TATCTGATTA GTAATTGTTG
satE     GGGTTTT--T GGGATATCTC AATGCTATTTA TACCTGATTA GTAAATGATG
steO     GGGTTTT--G GG-TTATCTC AATGCTTATTG TATCTGATTA GTAAGTGTTG
steE     GGGTTTT--G GGGTTATCTT AATGCTAATTG TATCTGACTA GTAAGTGCTG

```

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          160         170         180         190         200
pAs120a  TAACTACAAC GGAATGGTTA ACTAAAACCGG CCCACTGCTT ATTTTTATCAG
strO     CAACTACAAC GGAATGGTTA ACTAAAAGTG CCCACTGATT ATTTTTATCAG
satK     TAACTAGAAC GGAATGGTTA ACTAAAAATG CCCACTGTTT ATTTTTATCAG
satM     CAACTGCAAC GGAATGGTTA ACTAAAACCGG CCCACTGCTT ATTTTTATCAG
satE     CAACTACAAC GGAATGGTTA ACTAAAAATG CCCACTGCTT ATTTTATCAG
steO     CAACTATAAC GGAATGATTA ACTAAAAGTG CCCACTGCTT ATTTTTATCAG
steE     CCACTATAAT GGAATGGTTA ACTAAAATGA CCCACTGCTT ATTTTTATCAG

```

```

.....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
          210         220         230         240         250
pAs120a  GACGATTATC TATACCTGAC TAGTAATTGC TGCAACGATA ACCGAATAAA
strO     GATGATTATC TATACCTGAC TAGTAATTAC TGCA-CGATA ACCGAATAAG
satK     GATGATTATC TATACCTGAC TAGTAACTAC TGCAACGATA ACTGAATAAA
satM     GACGATTATC TATACCTGAC TAGTAATTGT TGCAACGATA ACTGAATAAA
satE     GATGATTATC TATACCTGAC TAGTAATTAC TGCAACGATA ACTGAATAAA
steO     GACTATTATC TCTACCTGAC TAGTAATTGC TGTAACGATA ACTGAATAAA
steE     GATGCTTATC TATACCTGAC TAGTAATTGC TGCAACGATA ACCGAATAAA

```

.....|.....||.....||.....||.....|

		260		270		280	
pAs120a	TAAAA-TGCC	AACT	GTGTGA	GTTCTCGGT	CGACTTCCT		
strO	TAAAAGTGCC	AACT	GTGTGA	GTTCTCGGT	CGACTTCCT		
satK	TAAAAGTGCC	CACGGT	GTGTGA	GTTCTCGGT	CGACTTCCT		
satM	TAAAAGTGCC	AACGGT	GTGTGA	GTTCTCGGT	CGACTTCCT		
satE	TAAAAGTGCC	CACT	GTGTGA	GTTCTCGGT	CGACTTCCT		
steO	TAAAAGTGCC	AACGGT	GTGTGA	GTTCTCGGT	CGACTTCCT		
steE	TAAAAGTGCC	AACT	GTGTGA	GTTCTCGGT	CGACTTCCT		

Anexo 7 – Homologia das sequências pAs120a em percentagem.

%	pAs120a	strO	satK	satM	satE	steO	steE
pAs120a	100						
strO	91	100					
satK	92	92	100				
satM	95	92	93	100			
satE	92	93	96	93	100		
steO	92	91	93	93	92	100	
steE	92	93	92	91	93	94	100

Anexo 8 – Amplificação da sequência pAm1 de várias espécies de *Avena* com os primers Am1 forward e reverse (pAm1 gentilmente cedido por A. Fominaya, Departamento de Biologia Celular e Genética, Universidad de Alcalá (UAH) – Madrid, Espanha)

