



Définition des séquences sonde pour la PCR et pour les puces à ADN

Hubert Charles, Federica Calevro

► To cite this version:

Hubert Charles, Federica Calevro. Définition des séquences sonde pour la PCR et pour les puces à ADN. Bioinformatique : Principes d'utilisation des outils, Editions Quae, 272 p., 2010, Savoir Faire (Quae), 978-2-7592-0870-8 978-2-7592-0870-8. hal-02813659

HAL Id: hal-02813659

<https://hal.inrae.fr/hal-02813659>

Submitted on 1 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

~~FICHE 55~~

Fiche 54

Définition des séquences sonde pour la PCR et pour les puces à ADN

Hubert Charles et Federica Calevro

Une sonde est un fragment d'ADN simple brin de courte taille (10 à 100 bases) qui est utilisé en biologie moléculaire pour « pêcher » des gènes spécifiques selon le principe de l'hybridation : deux molécules d'ADN complémentaires s'associent pour former des doubles brins stables. L'optimisation de sondes (« probe design ») est une problématique qui a vu le jour avec le développement de la technique PCR (*Polymerase Chain Reaction*). Tout d'abord réalisée de façon empirique, l'optimisation bio-informatique de la séquence de la sonde s'est révélée nécessaire avec l'apparition des techniques de biologie moléculaire dites à « hauts débits », comme les puces à ADN ou la PCR multiplexe⁵⁰.

Si les séquences nucléotidiques (ARN ou fragments d'ADN) ont été préalablement séparées sur un gel d'électrophorèse, puis transférées sur une membrane, la sonde marquée par de la radioactivité est alors déposée en solution sur la membrane, et l'hybride est finalement détecté par radiophotographie ; ce sont les techniques du Southern blot (ADN / ADN) et du northern blot (ADN / ARN). Si la sonde est fixée sur un support de verre (ou une membrane de nylon), ce sont les gènes qui sont alors marqués avec de la fluorescence (ou de la radioactivité) et déposés en solution sur le support ; c'est la technique des puces à ADN (ou des microarrays). Enfin, les sondes peuvent être utilisées comme amorces d'une polymérase pour amplifier spécifiquement un fragment d'ADN au cours d'une réaction PCR. Mais finalement, quelle que soit la technique utilisée, le choix de la séquence optimale d'une sonde sera toujours guidé par les deux critères essentiels qui la caractérisent, l'affinité et la spécificité pour son gène cible.

La mesure de l'affinité d'une sonde pour sa cible passe par une analyse thermodynamique incluant les paramètres physico-chimiques de la séquence elle-même, mais aussi de l'expérimentation (température et salinité, par exemple). La mesure de la spécificité est un problème d'analyse comparative entre les séquences des gènes d'un même génome ou entre séquences de gènes d'espèces différentes. Nous verrons au cours de ce chapitre l'importance relative de ces deux paramètres et les solutions bio-informatiques apportées selon les différentes techniques expérimentales.

Affinité d'une sonde pour sa cible

L'affinité d'une sonde pour son gène cible définit le rendement d'hybridation, c'est-à-dire la fraction d'hybrides double brins cible-sonde présente dans le milieu⁵¹. Cette affinité est liée à la température de la solution, à la concentration saline, à la séquence de la molécule d'ADN (composition et voisinage des bases A, C, G T) et à sa concentration. Enfin, la présence éventuelle d'agents dénaturants dans la solution, comme l'urée ou le formamide, aura une influence forte sur l'affinité cible-sonde. L'optimisation thermodynamique des sondes est un critère primordial pour les puces à ADN, afin de maximiser l'homogénéité du signal de l'ensemble des gènes de la puce. En PCR, et tout particulièrement pour la PCR multiplexe, cette optimisation est souvent la clé de la réussite.

L'affinité d'une sonde peut se définir par la « température de fusion » T_m (*Temperature of melting*). Il s'agit de la température à laquelle la moitié des acides nucléiques est sous forme

⁵⁰ Réaction PCR au cours de laquelle plusieurs gènes sont amplifiés parallèlement grâce à des mélanges de couples d'amorces.

⁵¹ Williams, 1989, *Biotechniques*, 7:762-769.

double brin, comme le montre la Figure 1.

Figure 54.01
Manquante



Figure 1. Courbe de dissociation réversible de la molécule d'ADN. Le point d'inflexion de la sigmoïde correspond à la valeur T_m où 50% des molécules sont sous forme double brin. La dissociation est mesurée par la variation de l'absorbance à 260 nm de la molécule d'ADN entre les états simple et double brin (effet hyperchrome).

Détermination de la température de fusion

Il existe une méthode empirique très simple pour le calcul des T_m d'amorces courtes, bien connue des biologistes moléculaires adeptes de la PCR :

$$T_m(^{\circ}\text{C}) = 4GC + 2AT$$

Cette formule est adéquate pour obtenir un ordre de grandeur de la T_m de sondes comprises entre 18 et 25 bases, il suffit de compter 4 fois le nombre de bases G ou C et d'ajouter 2 fois le nombre de bases A ou T, mais elle ne doit pas être utilisée dans le cas d'une recherche de sondes ou d'amorces optimales.

Actuellement, le modèle le plus complet pour déterminer la T_m de courtes séquences d'ADN est le modèle thermodynamique dit du « plus proche voisin ». Au sein d'une molécule d'ADN double brin, ce modèle définit des interactions de voisinage entre deux nucléotides successifs, ce qui permet de prendre en compte à la fois la nature et la place des nucléotides. Ces paires de nucléotides sont associées à des valeurs distinctes d'enthalpie, d'entropie et d'énergie libre pour l'association des deux brins d'ADN. Le modèle se présente sous la forme suivante :

$$T_m(^{\circ}\text{C}) = \frac{\Delta H}{\Delta S + R \ln(C_T)} + 16,6 \log\left(\frac{[K^+]}{1 + 0,7[K^+]}\right) - 273,15.$$

R représente la constante des gaz parfaits ($R = 1,987 \text{ cal.K}^{-1}.\text{mol}^{-1}$) et C_T la concentration totale en acides nucléiques appariés, $[K^+]$ est la concentration en ions potassium de la solution. Les calculs de l'enthalpie (ΔH) et de l'entropie (ΔS) d'hybridation peuvent être décomposés en plusieurs termes additifs :

$$\Delta H_{total} = \Delta H_{initiation} + \Delta H_{symétrie} + \sum_x \Delta H_x$$

$$\Delta S_{total}[Na^+ 1M] = \Delta S_{initiation} + \Delta S_{symétrie} + \sum_x \Delta S_x$$

Les termes ΔH_x (et ΔS_x) représentent les interactions de voisinage de bases au sein de la séquence. Ces termes ont été déterminés expérimentalement pour chaque couple de paires de bases possibles, ainsi que les termes d'initiation et de symétrie. Ils sont disponibles dans la littérature⁵². L'enthalpie est indépendante de la concentration en sels. En revanche, pour l'entropie, la formule théorique considère une concentration sodique de 1 M, on introduit alors un terme correctif pour une séquence contenant N bases :

⁵² SantaLucia J, 1998, *Proc Natl Acad Sci USA*, 95:1460-1465.

$$\Delta S_{total} = \Delta S_{total}[Na^{+} 1M] + 0,368(N-1)\ln[Na^{+}]$$

Lorsque les séquences étudiées dépassent cinquante nucléotides, elles présentent de nombreux états intermédiaires d'appariements partiels, et la méthode du plus proche voisin n'est plus applicable. Il est donc préférable d'utiliser la formule suivante qui intègre le taux de GC (%GC), le taux de mésappariements entre les deux brins (%mésap.) et le nombre N de nucléotides de la séquence :

$$Tm(^{\circ}C) = 81,5 + 0,41(\%GC) + 16,6\log\left(\frac{[K^{+}]}{1 + 0,7[K^{+}]}\right) - (\%mésap.) - \frac{500}{N}$$

Énergie libre de formation des structures secondaires

Un autre critère thermodynamique peut interférer avec le processus d'hybridation, il s'agit de la formation de structures secondaires (Figure 2). Ces structures sont des autorepliements des sondes sur elles-mêmes (épingles à cheveux) ou des agencements tête-bêche (homoduplexes) ; elles sont d'autant plus stables que la valeur de leur énergie libre de formation est négative. La stabilité de ces structures est calculable avec le modèle du plus proche voisin, comme nous l'avons vu précédemment en intégrant quelques contraintes supplémentaires. Pour que la formation d'une épingle à cheveux soit possible, sa tige doit posséder au minimum deux liaisons entre bases complémentaires et sa boucle doit contenir au moins trois bases. En ce qui concerne les homodimères, leur formation est possible uniquement s'il existe au moins deux liaisons successives entre bases complémentaires. Le logiciel de référence pour le calcul de la stabilité des structures secondaires est le logiciel MFOLD⁵³. On peut citer également les logiciels OligoAnalyzer⁵⁴ et HyTher⁵⁵ qui présentent des interfaces conviviales et des paramètres mis à jour régulièrement.

⁵³ Zuker, 2003, *Nucleic Acids Res*, 31:3406-3415.

⁵⁴ <http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>

⁵⁵ Chavali *et al.*, 2005, *Bioinformatics*, 21:3918-25.

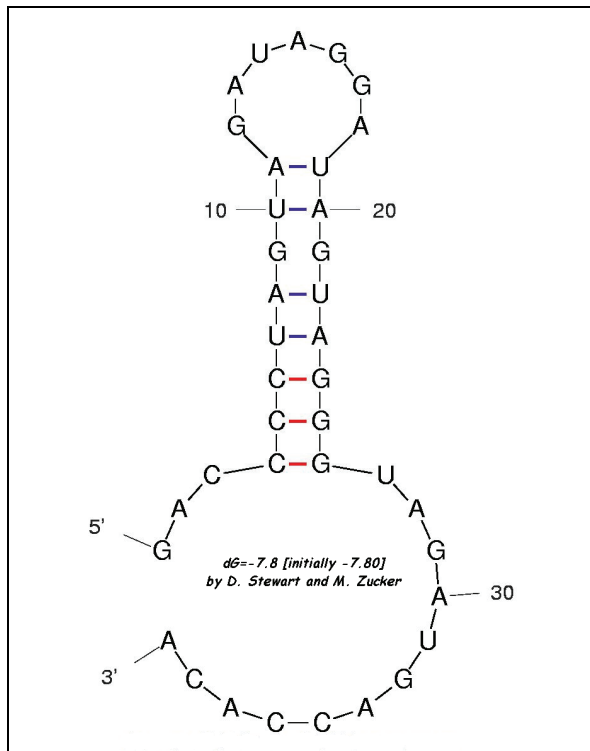


Figure 54.02

Figure 2. Modélisation du repliement en épingle à cheveux d'une sonde de 39 bases fournie par le logiciel MFOLD interfacé sur le site suivant : <http://www.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi>.

Autres critères thermodynamiques spécifiques

D'autres critères vont influencer la stabilité de l'hybridation. Ces critères sont beaucoup moins importants que la composition et le voisinage des bases, ou les structures secondaires. Ils vont, de plus, dépendre fortement des techniques pour lesquelles les sondes sont utilisées.

Les pentamères situés aux extrémités de la sonde (et plus précisément la première et la dernière base) jouent un rôle important sur l'affinité des sondes pour leurs cibles et influencent la stabilité de l'hybride qui en résulte. En effet, des ancrages GC aux extrémités des sondes stabilisent la formation des hybrides (selon le principe de la « fermeture Eclair ») et permettent d'obtenir de meilleurs rendements d'hybridation. Pour les puces à ADN, des pentamères riches en GC (stables) seront donc privilégiés aux deux extrémités des sondes. A l'inverse, en PCR, on favorisera les pentamères moins stables, riches en AT, en position 3' des sondes (même si des affirmations contraires sont souvent présentées sur internet), de façon à ce que la polymérase ne s'amorce pas sur une hybridation partielle et aspécifique de la sonde en 3'.

Les motifs de faible complexité (répétitions locales de bases) sont à éviter dans les séquences sondes (PCR et puces à ADN), d'une part, parce qu'ils sont très communs (problème de la spécificité) dans les régions transcrites et non traduites des gènes (5' et 3' UTR), et d'autre part, parce que la présence de bases successives est à l'origine d'un biais dans le calcul des énergies d'interaction avec le modèle du plus proche voisin.

Les valeurs d'énergie du modèle thermodynamique du plus proche voisin les plus récemment publiées ont été estimées très précisément pour des hybrides parfaits en solution. Ces mesures

sont donc très précises pour le calcul des T_m des amorces PCR. En revanche, dans les expérimentations sur puces, la sonde est immobilisée sur la lame. Il est donc probable que les entropies de la sonde et de l'hybride cible-sonde soient inférieures à celles qui sont calculées en solution. Peu de travaux ont été consacrés à ce problème. Néanmoins, dans cette technique, les valeurs absolues de T_m sont assez peu importantes car il s'agit plutôt de déterminer, pour l'ensemble du jeu de sondes, des valeurs de T_m homogènes.

Une autre limitation du modèle thermodynamique du plus proche voisin est liée à l'estimation des concentrations des molécules en présence. Ces concentrations sont estimables pour la PCR, mais une fois encore l'immobilisation des sondes dans les puces à ADN complique le problème. Cependant, sachant que les sondes sont déposées en large excès par rapport aux cibles, on peut considérer que la concentration en acides nucléiques appariés est proche de la concentration moyenne en cibles ; mais bien sûr la concentration en cible n'est pas connue puisque c'est le paramètre mesuré. Par conséquent, lorsqu'un jeu de sondes est défini avec une concentration constante, un biais est introduit : la T_m des gènes fortement exprimés est surestimée, alors que celles des gènes faiblement exprimés est sous-estimée puisque la concentration intervient au dénominateur de la formule. L'écart de signal de fluorescence est donc artificiellement amplifié entre les gènes fortement et faiblement exprimés.

Spécificité d'une sonde pour sa cible

La spécificité d'une sonde est sa capacité à se lier de façon unique sur son gène cible. L'optimisation de la spécificité dans une problématique de recherche de séquences sondes consiste donc en l'élimination des similitudes d'une sonde avec tout autre gène que le gène cible.

La spécificité et la taille des sondes sont deux facteurs liés. Plus une sonde est courte, plus elle pourra être choisie dans une région spécifique, et plus elle sera sensible à la présence de mésappariements. La contrepartie de cette forte spécificité des sondes courtes est leur faible affinité, puisque la T_m est inversement proportionnelle à la taille de la sonde.

La recherche de similarités de séquences est un problème très étudié en bio-informatique, notamment pour répondre aux besoins de la phylogénie et de la génomique comparative. Très naturellement, ce sont ces outils qui ont été utilisés pour l'optimisation de la spécificité des sondes. Le logiciel d'alignement le plus couramment utilisé est le logiciel BLAST (voir Fiche 13). Il faut noter tout de même que le paramétrage par défaut de l'outil BLASTn n'est pas compatible avec la recherche de sondes spécifiques de courtes tailles. Par exemple, pour remplir les critères de spécificité de Kane (*cf.* ci-dessous), il faut impérativement modifier les paramètres W (taille du mot d'ancrage) et q (pénalité du mésappariement) aux valeurs 7 et -2 respectivement (voir Fiche 16).

Spécificité des sondes pour les puces à ADN de génotypage

Les puces de génotypage (puces taxonomiques) sont des puces utilisées essentiellement comme outils diagnostiques. Un exemple d'application est l'utilisation de sondes « signatures » qui permettent d'identifier en une seule hybridation, l'ensemble des pathogènes présents dans un fluide biologique pour un diagnostic clinique. C'est sans doute la plus importante application industrielle des puces à ADN actuellement.

Lorsqu'on fait du génotypage, la spécificité des sondes est l'enjeu majeur. Chaque sonde doit être spécifique de son gène cible par rapport à ce même gène issu d'autres organismes également représentés sur la puce et susceptibles de se trouver dans l'échantillon analysé. Ainsi, les sondes choisies sont des sondes courtes (10 à 20 bases) dont l'hybridation est

sensible à un seul mésappariement central.

Le problème majeur du choix des séquences est celui de la détermination de « séquences signatures ». Ces signatures doivent être représentatives du groupe taxonomique visé, c'est-à-dire une souche, une espèce, un genre, ou encore une famille. Elles sont donc localisées dans des zones conservées pour tous les individus du groupe visé. Néanmoins, ces séquences doivent permettre de discriminer des groupes taxonomiques proches, et sont donc localisées dans des régions suffisamment variables pour présenter un polymorphisme au niveau taxonomique souhaité. Toute la difficulté provient du dilemme entre ces deux types de polymorphisme qu'on souhaite différencier. Enfin, les bases spécifiques des séquences ciblées doivent, si possible, être placées au milieu de la sonde, de façon à optimiser l'écart d'énergie libre d'appariement entre les sondes parfaitement complémentaires et les sondes aspécifiques

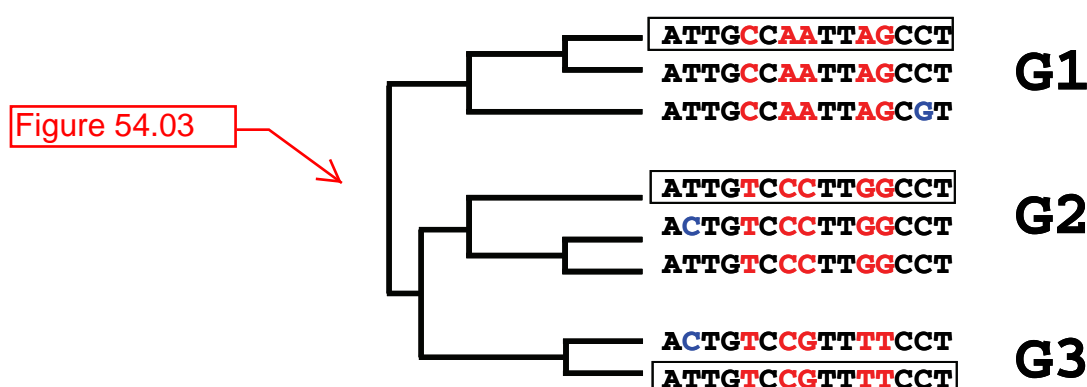


Figure 3. Alignement de séquences de 3 groupes taxonomiques (G1, G2, G3) phylogénétiquement proches. Les 3 séquences sondes sélectionnées sont encadrées. Les bases rouges, correspondant au polymorphisme inter-groupe, sont situées au centre pour garantir la spécificité des sondes. Les bases bleues correspondent à du polymorphisme intra-groupe, leur localisation externe permet une hybridation correcte de la sonde sur sa cible. Il est également possible de choisir plusieurs sondes par groupe pour intégrer ce polymorphisme.

Spécificité des sondes pour les puces à ADN transcriptomiques

En analyse du transcriptome, le problème de la spécificité des sondes est par nature plus simple que pour le génotypage. Néanmoins, l'analyse de l'expression de gènes issus de familles multigéniques ou de phénomènes d'épissages alternatifs, pourra venir sérieusement compliquer le problème. Du point de vue bio-informatique, il s'agit de définir des sondes qui soient strictement représentatives (uniques) de l'ARNm dont on veut étudier l'expression. Les sondes choisies sont généralement de taille comprise entre 35 et 70 bases (la technique Affymetrix n'est pas considérée ici) de façon à garantir une spécificité importante et une bonne affinité, permettant une mesure quantitative du rapport d'expression de chaque transcrit. Des critères simples ont été énoncés par Kane *et al*⁵⁶ et font références pour de nombreux logiciels :

⁵⁶ Kane *et al.* 2000, *Nucleic Acids Res.*, 28:4552.

- la séquence de la sonde ne doit pas présenter plus de 75% de similarité sur toute sa longueur avec une séquence non-cible de l'organisme ;
- la séquence de la sonde ne doit pas contenir plus de 15 bases consécutives strictement identiques à une séquence non-cible de l'organisme.

Le problème peut se compliquer si l'on souhaite définir des jeux de sondes pour des organismes dont on ne connaît pas l'intégralité du génome (hybridation aspécifique inconnue), ou si l'on souhaite analyser l'épissage alternatif chez les eucaryotes, ou encore si l'on travaille sur des échantillons présentant des mélanges d'espèces. C'est le cas lorsqu'on étudie le transcriptome d'organismes symbiotiques ou de parasites intracellulaires dont les ARNm sont mélangés à ceux de l'hôte.

Spécificité des amorces PCR

L'optimisation du choix des amorces PCR est plus un problème d'optimisation thermodynamique qu'un problème de spécificité : il s'agit d'égaliser les rendements d'hybridation des deux amorces 3' et 5'. Les amorces PCR sont généralement de très courtes tailles (18 à 25 bases), et il est donc souvent assez simple de les choisir dans des régions spécifiques. Néanmoins, le problème peut devenir plus complexe lorsqu'on souhaite dessiner des amorces dégénérées (ou intégrant des bases atypiques comme la désoxyinosine par exemple) pour rechercher des gènes dans des génomes inconnus, lorsqu'on connaît des homologues dans d'autres organismes. Peu de logiciels dédiés à l'optimisation des sondes PCR intègrent ce contrôle de spécificité et celui-ci nécessite souvent d'être réalisé dans un deuxième temps.

Des logiciels de choix de séquences sondes

Nous ne voulons pas ici tenter de dresser la liste de tous les logiciels libres dédiés au choix des séquences sondes pour les puces à ADN et pour la PCR, car ils sont extrêmement nombreux. Tous ne sont pas de qualité équivalente, et chacun sera adapté à des problématiques plus ou moins spécifiques.

Pour le choix de séquences sondes de génotypage, le logiciel libre ARB⁵⁷ est sans doute le plus utilisé. ARB permet de calculer localement des sondes spécifiques de groupes taxonomiques définis sur la base d'un arbre phylogénétique avec un positionnement optimisé des mésappariements. Même si le maniement de cet outil demande un peu d'investissement en temps, l'interface est très conviviale. L'outil est téléchargeable intégralement sur une machine locale et peut fonctionner sur n'importe quelle base de séquences alignées associée à un arbre phylogénétique. Une base de sondes précalculées est également disponible sur le serveur. On peut également citer le logiciel Primrose⁵⁸ qui permet le choix de sondes courtes ou d'amorces PCR. Ce logiciel est couplé au RDP (*Ribosomal Database Project*), et permet un choix d'oligonucléotides extraits de cette base d'ARN ribosomiques.

Pour le choix des sondes de puces à ADN destinées à la transcriptomique, on peut citer les logiciels OligoArray 2.0⁵⁹, Mprobe 2.0⁶⁰, OligoWiz⁶¹, OligoPicker⁶², ROSO⁶³ et GoArrays⁶⁴.

⁵⁷ Ludvig W *et al.*, 2004, *Nucleic Acids Research*, 32:1363-1371

⁵⁸ Ashelford *et al.*, 2002, *Nucleic Acids Research*, 30:3481-3489.

⁵⁹ Rouillard JM *et al.*, 2003, *Nucleic Acids Research*, 31:3057-3062.

⁶⁰ Li and Ying, 2006, *Appl Bioinformatics*, 5:181-186.

⁶¹ Wernersson and Nielsen, 2005, *Nucleic Acids Research*, 33:611-615.

⁶² Wang and Seed, 2003, *Bioinformatics*, 19:796-802.

⁶³ Raymond N *et al.*, 2004, *Bioinformatics*, 20:271-273.

⁶⁴ Rimour *et al.*, 2005, *Bioinformatics*, 21:1094-1103.

Ces logiciels présentent tous des spécificités propres qui les rendront plus performants pour des applications particulières. Ainsi, du point de vue de la stratégie globale de recherche de sondes spécifiques, les quatre premiers logiciels découpent chaque gène cible en sondes potentielles de la longueur désirée, à partir d'une extrémité fixée par l'utilisateur. Ces sondes potentielles sont alors testées successivement pour leur spécificité, puis pour leur affinité. Cependant, ces 4 logiciels n'utilisent pas les mêmes modèles thermodynamiques et ne recherchent pas les mêmes structures secondaires. Les sondes optimales sont finalement sélectionnées sur la base de valeurs seuils. Lorsque la première sonde valide est obtenue pour le gène cible analysé, les logiciels arrêtent la recherche et passent au gène suivant. Ces algorithmes permettent une optimisation du jeu de sondes intégralement automatisée. Mais en revanche, ils offrent très peu de souplesse vis-à-vis du choix de la localisation et de l'intervalle de Tm souhaité pour les sondes. Le logiciel Mprobe 2.0 permet une grande interopérabilité avec les bases de séquences et leurs divers formats. Il propose également un choix manuel des sondes (sans valeur seuil) selon la stabilité de leurs structures secondaires. Le logiciel OligoArray propose des informations sur les hybridations non spécifiques potentielles. OligoWiz est bien adapté au génome des eucaryotes et permet l'élimination de séquences introniques ; il permet également d'effectuer une pondération personnalisée des paramètres de recherche avec une interface graphique très performante.

Le logiciel ROSO propose un autre algorithme de recherche. Il va tout d'abord rechercher dans les gènes des zones de spécificité maximale ; puis dans ces zones, une démarche d'optimisation itérative est lancée, permettant ainsi d'obtenir plusieurs sondes pour les différents gènes, sans utiliser des seuils stricts. C'est la variation des paramètres sur l'ensemble du jeu de sondes qui est minimisée dans cette approche. Le découplage entre la partie recherche de spécificité (réalisée à l'aide du logiciel BLAST), très coûteuse en temps de calcul, et la partie thermodynamique, offre la possibilité d'une optimisation plus fine : l'utilisateur peut à tout moment demander un calcul intégral d'un nouveau jeu de sondes, en relaxant manuellement un paramètre de localisation ou une valeur d'énergie libre seuil de structure secondaire. Cette souplesse demande une part d'expertise manuelle beaucoup plus importante que pour les autres logiciels.

Enfin, le logiciel GoArray présente une démarche de recherche de sondes chimériques très originales. Les sondes produites sont constituées de deux parties spécifiques reliées entre elles par un « linker » non spécifique. Cette approche permet d'allier efficacement la spécificité de la sonde (propriété des sondes courtes) à la sensibilité de la détection (propriété des sondes longues). L'algorithme de recherche est assez classique puisqu'il procède par découpage d'une première portion du gène, en partant d'une extrémité définie. Cette première partie de sonde est testée pour sa spécificité et son affinité, en utilisant les logiciels BLAST et MFold et le modèle thermodynamique du plus proche voisin. Puis, le logiciel cherche la deuxième partie de la sonde en imposant une contrainte de longueur minimale et maximale pour le « linker ». Lorsqu'une sonde chimère est trouvée, le logiciel génère le « linker » pour assurer l'absence d'hybridation avec la cible ; la recherche est alors lancée sur le gène suivant. Du point de vue informatique, GoArray est le seul logiciel qui s'inscrive véritablement dans une démarche de développement pour la réutilisation des composants logiciels, et il intègre par exemple, le standard MAGE-OM pour la description des objets traités, ainsi que des modules plateformes indépendants fournis par le Consortium MGED⁶⁵.

Pour les expériences de PCR, on peut citer primer3⁶⁶ qui est un logiciel très utilisé dans la communauté bio-informatique, et qui a été récemment mis à jour pour tester les possibilités

⁶⁵ <http://www.mged.org/Workgroups/MAGE/mage-om.html>

⁶⁶ Untergasser et al., 2007, *Nucleic Acids Res.* 35:W71-W74.

d'amorçages aspécifiques dans les échantillons. Le logiciel propose un large choix de paramètres thermodynamiques et de paramètres du protocole expérimental.

Il existe également de très nombreux logiciels commerciaux adaptés à chacune de ces technologies. Ceux-ci ne seront pas présentés ici. Ils sont généralement de bonne qualité. Avant d'en acquérir un, chacun vérifiera : (1) qu'il propose une optimisation à la fois de la spécificité et de l'affinité des sondes, (2) qu'il intègre un nombre important de paramètres du protocole, (3) que la modélisation des structures secondaires soit intégrée et (4) qu'une interface conviviale (présentant par exemple de façon dynamique les énergies d'association) soit présente pour guider une sélection semi manuelle des amorces.