

GENOMA CLÍNICO

Carlos A. Moreira-Filho
Professor do Instituto de Ciências Biomédicas e Coordenador do Centro de Pesquisas em Biotecnologia da USP
cmoreira@icb.usp.br

Sergio Verjovski-Almeida
Professor titular de Bioquímica
Instituto de Química da USP
verjo@iq.usp.br

O impacto da genética genômica na prática médica

Fotos cedidas pelos autores

A genética genômica estuda a natureza física e o funcionamento dos genomas, isto é, do material genético contido no conjunto de cromossomos de cada espécie. O rápido avanço que vem ocorrendo na pesquisa genômica deve-se, em larga medida, ao surgimento de novas tecnologias para o seqüenciamento automatizado do DNA e para o estudo da expressão gênica, bem como à disponibilidade de poderosos recursos de informática para organização e análise dos dados obtidos em nível molecular (Lee & Lee, 2000). Essa trajetória tecnológica consolida um novo paradigma de pesquisa, por muitos denominado “Era Genômica”, com forte impacto nas áreas de aplicação das ciências biológicas, como a medicina e a agricultura. Neste artigo são abordados alguns usos correntes da genética genômica em medicina, com enfoque na genética do câncer.

A “Era Genômica” iniciou-se com a capacitação para seqüenciar genomas inteiros (inicialmente de microrganismos, depois genomas maiores e mais complexos) com razoável rapidez e colocar essa informação em bancos de dados *online*. Com o apoio da análise computacional, essas informações estão sendo ordenadas, permitindo mapear com precisão os genes e demais seqüências de DNA no genoma de diversos organismos. O progresso nesse campo de traba-

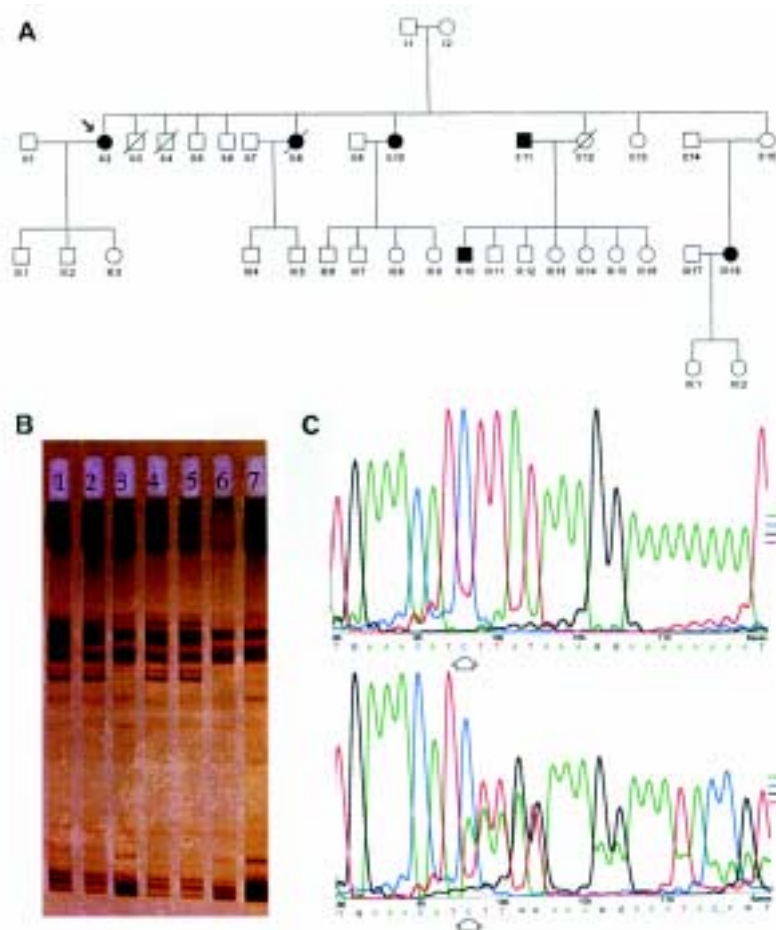


Figura 1- a) Heredograma mostrando recorrência familiar de câncer de mama. O paciente índice é indicado pela seta (caso referido por S. Simon, HIAE). **b)** SSCP-HA de um fragmento de 332bp do término 3' do exon 10 de BRCA2, mostrando alterações de padrão de migração eletroforética nas canaletas 1, 2, 4 e 5, correspondendo aos indivíduos II-2, III-3, II-10 e III-10. O controle (indivíduo normal não relacionado) aparece em 7. **c)** Seqüenciamento do DNA de controle normal (cromatograma superior) e do paciente-índice, II-2 (cromatograma inferior). A seta indica o ponto de início da mutação 2024del CTTAT em BRCA2

lho, conhecido como genômica estrutural, vai possibilitar a comparação da estrutura dos genomas de várias espécies, o que é essencial para se investigar a existência de leis gerais governando a organização dos genomas.

Num plano mais imediato, o conhecimento atual sobre a posição cromossômica e a seqüência de grande número de genes em diversos genomas já traz enormes benefícios. Basta considerar, como exemplo, o genoma humano: as seqüências de genes que têm importância médica por estarem associados a diferentes doenças têm sido reveladas em larga escala nos últimos anos. A partir dessa informação, diversos métodos podem ser empregados para detecção de mutações nesses genes, as quais são a causa primária de doenças, ou aumentam a susceptibilidade a elas.

Deteção e caracterização de mutações gênicas nos testes de predisposição ao câncer

A correlação entre mutações gênicas e susceptibilidade a doenças é particularmente importante no câncer. Cerca de 5% a 10% de todos os casos de câncer aparecem em indivíduos que herdaram alguma mutação que predispõe à doença e em suas famílias há outras pessoas com esse risco (Offit, 1998). A descoberta de genes cuja alteração por mutações aumenta, ou reduz, o risco de um indivíduo desenvolver um tumor maligno é uma das grandes conquistas da medicina. Ela permite identificar quem está sob risco aumentado e pode se beneficiar de práticas preventivas, de vigilância e de intervenção terapêutica. A detecção de mutações germinativas que predispõem ao câncer é indicada aos indivíduos cuja família tem um forte histórico de recorrência de casos de câncer e/ou com início precoce da doença. É também limitada às síndromes hereditárias onde está bem estabelecido o papel de um determinado gene, ou conjunto

de genes, no aparecimento do tumor. Entre essas síndromes estão os cânceres hereditários de mama e ovário, de cólon (com e sem polipose), o retinoblastoma familiar, a neoplasia endócrina múltipla (MEN) e a Síndrome de Li-Fraumeni (diversos tumores com aparecimento precoce em pacientes com mutações germinativas no gene supressor de tumor *p53*).

O desenvolvimento de testes genéticos para casos familiares de câncer de mama e ovário é um bom exemplo da aplicação direta da genética genômica na prática médica. Os tumores de mama e ovário são neoplasias extremamente frequentes na população afetando, em média, 1 em cada 8 mulheres. Estima-se que em cerca de 5% a 10% de todos os cânceres de mama e ovário ocorram mutações em genes autossômicos, com padrão de herança dominante (Newman e col., 1988). No início dos anos 90, isolaram-se dois genes associados aos tipos de câncer acima citados: BRCA1 e BRCA2. Em cerca de 70% das famílias com quatro ou mais afetados, mutações em um desses dois genes são detectadas. Mulheres portadoras dessas mutações têm risco de 84% de desenvolverem câncer de mama até os 70 anos de idade, contra 12% da população em geral (Ford e col., 1998). No caso de portadoras de mutações em BRCA1, o risco de câncer de ovário até os 70 anos é de 44%, contra 1% da população geral (Ford e col., 1994).

O gene BRCA1 está localizado no cromossomo 17 (17q21) e é responsável por aproximadamente metade de todos os casos familiares precoces de câncer de mama, bem como pela maioria dos casos familiares de câncer de mama e ovário. Esse gene possui 22 exons codificantes e 2 exons não-codificantes, estendendo-se por 100kb (100 mil bases) de seqüência genômica (Miki e col. 1994). Seu produto é uma proteína de 1.863 aminoácidos. Um grande número de mutações diferentes têm sido descritas nas várias

famílias estudadas. Essas mutações distribuem-se aleatoriamente pelos diversos exons: nenhum "hot spot" de mutação foi identificado (van Orsouw e col., 1999).

O gene BRCA2 está localizado no cromossomo 13(13q12-13) e compõe-se de 27 exons dispersos por 70 kb de DNA genômico e seu transcrito (10-12 kb) codifica uma proteína de 3.418 aminoácidos (Wooster e col. 1995). Mutações nesse gene são responsáveis por aproximadamente 35% dos casos familiares precoces de câncer de mama. Estão ainda associadas a um risco aumentado de câncer de mama e próstata em homens, e de câncer de ovário e pancreático. A distribuição de mutações nesse gene também é aleatória. A expressão dessas mutações é variável: uma mesma mutação pode estar associada a diferentes tipos de câncer numa mesma família.

Embora o espectro de mutações dos genes BRCA1 e BRCA2 ainda não tenha sido completamente caracterizado, dados preliminares indicam que o tipo e a freqüência das mutações apresentam uma distribuição étnica e geográfica variável (Offit, 1998). Recentemente, surgiram análises de correlação genótipo-fenótipo, vinculando a distribuição espacial das mutações ao longo dos genes BRCA1 e 2 ao tipo de câncer apresentado pelo paciente (Blakwood & Weber, 1998). Nesse campo, já é óbvia a contribuição da biologia molecular para o avanço dos estudos genético-clínicos (Ellisen e Haber, 1998). No momento, a relevância em se detectar mutações nos genes BRCA1 e 2 deriva de dois aspectos principais: 1) diagnóstico e critérios prognósticos; 2) aconselhamento genético beneficiando familiares do paciente.

A detecção de mutações nos genes BRCA1/BRCA2 é feita por técnicas de biologia molecular capazes de identificar a alteração de uma única base na seqüência de DNA. Devido ao grande tamanho das seqüências de BRCA1 e 2, pri-

Hibridização com duas cores em Chip de DNA

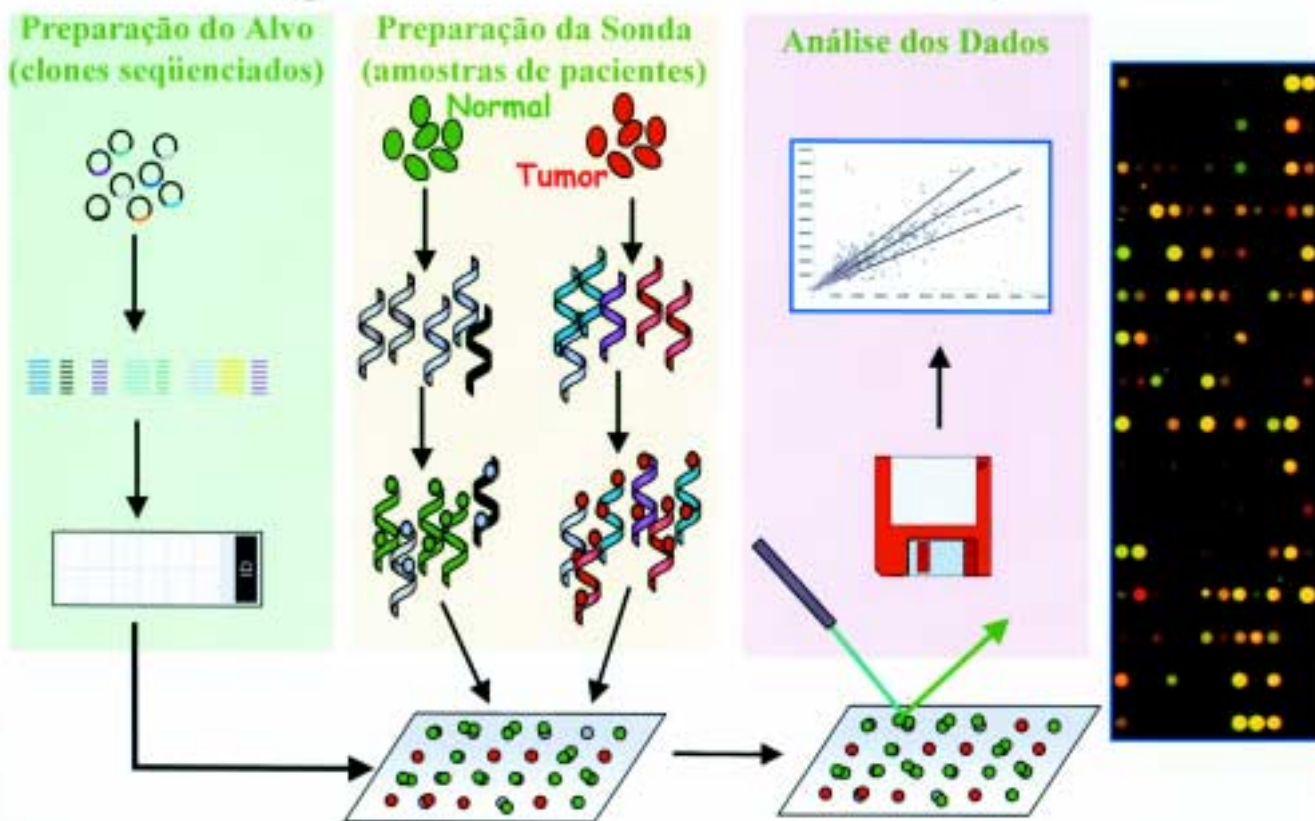


Figura 2 - Esquema de um experimento de análise de expressão gênica utilizando *DNA microarray*

meiramente busca-se identificar qual exon contém o sítio mutado antes de se proceder ao seqüenciamento do DNA. As seqüências de DNA correspondentes aos vários exons são amplificadas por PCR (reação em cadeia da polimerase) para a realização desses testes.

Para o *screening* inicial de mutações, uma técnica muito utilizada é a de SSCP (*single strand conformation polymorphism*), combinada com HA (*heteroduplex analysis*). A técnica de SSCP (Orita e col., 1989) baseia-se no fato de que a conformação tridimensional de fragmentos de DNA simples fita depende da seqüência de nucleotídeos, sendo que a diferença de um nucleotídeo altera o padrão de migração eletroforética. A técnica de HA (Nagamine e col., 1989) deriva da observação de que, em reações de PCR, onde estão presentes moléculas de DNA selvagens e mutantes, durante

os últimos ciclos, podem ser formados heteroduplexes entre essas duas espécies de moléculas, os quais terão um padrão de migração eletroforética diferente dos homoduplexes. Utilizando-se fragmentos de tamanho entre 100 e 350 pares de bases, a taxa de detecção de mutações é de aproximadamente 80%, tanto para o SSCP como para o HA, e a associação dos dois métodos pode elevar essa taxa para perto de 100% (sob condições bem controladas e processamento semi-automatizado).

Após a identificação do exon que abriga a mutação, esta é caracterizada pelo seqüenciamento do DNA. Conhecida a mutação, torna-se possível, também, por seqüenciamento, pesquisa-lo em outros membros da família do paciente, com vistas ao aconselhamento genético. A **Figura 1** mostra os procedimentos para detecção de muta-

ções no gene BRCA2 em indivíduos de uma família com histórico de câncer de mama (notar que há homens afetados) realizados no laboratório de um dos autores (C. A. M-F). Esse trabalho é feito em colaboração com a Unidade de Oncologia do Hospital Israelita Albert Einstein (equipe do Dr. Sergio Simon), onde os casos são selecionados – com base nos critérios recomendados pela American Society of Clinical Oncology (ASCO, 1996) – e os pacientes e seus familiares recebem aconselhamento genético.

A genômica funcional na pesquisa do câncer

A pesquisa de mutações associadas a variantes hereditárias de determinados tumores é apenas um aspecto da abordagem genética do câncer. A base genética da transformação maligna de células somáti-

cas tem um caráter combinatório, com mutações cumulativas no DNA. Embora tenha sido evidenciado há décadas que o câncer é causado por múltiplos fatores, só recentemente ficou claro que vários desses fatores se comportam como oncogenes (genes que contribuem para o aparecimento do câncer) não-dominantes cooperativos (Offit, 1998). Esses genes irão causar uma transformação maligna somente se alterados numa certa combinação e contra um certo *background*.

Os mecanismos genéticos e bioquímicos presentes num organismo multicelular têm sido tradicionalmente estudados com metodologias que fornecem apenas informações pontuais, sem abranger a complexidade das interações entre esses mecanismos. Barreiras tecnológicas fizeram com que se focalizasse apenas um pequeno conjunto de parâmetros biológicos, embora a maioria dos fenômenos investigados integre um conjunto mais complexo de fatores causais, mutuamente dependentes. Essa situação era particularmente limitante para a pesquisa dos mecanismos envolvidos na transformação de uma célula normal em célula cancerosa.

Desenvolvimentos recentes nas áreas de computação e tratamento de imagens (ver adiante) começaram a derrubar as barreiras que impediam a obtenção de informações sobre a expressão simultânea e relativa – ao nível do RNA mensageiro – dos genes expressos por um genoma. Nos organismos multicelulares esse conjunto de genes expressos varia conforme o tipo de célula e tecido. Essa variação também ocorre entre células normais e cancerosas de um mesmo tecido, e entre classes de tumores (Golub e col., 1999).

A análise diferencial da expressão gênica está hoje enormemente facilitada pela emergência da tecnologia de *DNA microarrays*, que permite mensurar simultaneamente o nível de expressão de milhares de

genes em um único ensaio de hibridação (Harrington e col. 2000). A **Figura 2** mostra um experimento típico com *DNA microarrays*. Cada *array* consiste em um conjunto reproduzível de milhares de seqüências de DNA (representativas dos genes expressos num microrganismo ou num tecido humano) depositado num suporte sólido, geralmente uma lamínula de vidro especial. DNA preparado a partir de RNA mensageiro (cDNA) de tecido normal ou tumoral é marcado por fluorescência com duas cores diferentes, hibridado ao DNA complementar no *array* e detectado através de leitura num *scanner* a laser. As intensidades de hibridação para cada seqüência de DNA no *arrays* são automaticamente analisadas e correlacionadas com o nível relativo de expressão gênica. Essas informações servem para o estudo de padrões de expressão, cuja variação pode indicar a ação de uma droga, um processo de transformação celular, etc.

A tecnologia de *DNA microarrays* é a chave para o estudo da genômica funcional, isto é, do conjunto completo de informações gênicas transcritas em RNA, ou transcriptoma, e do correspondente conjunto de proteínas codificadas, ou proteoma (Lee & Lee, 2000). A aquisição maciça e rápida de dados sobre o transcriptoma de um tecido é especialmente relevante, de duas maneiras, para a pesquisa sobre o câncer. Primeiro, o processo de *screening* dos genes que, isoladamente, sejam importantes para a biologia do câncer poderá ser significativamente mais rápido e mais eficiente. Segundo, como a informação sobre os diversos genes será analisada simultaneamente, pode-se esperar encontrar pistas importantes para a elucidação das vias que interagem na programação da transformação maligna das células.

É portanto provável que a análise em larga escala do padrão de expressão seja extremamente esclarecedora no que diz respeito às

modificações relativas a um tumor, em comparação com o tecido normal correspondente. Na verdade, os primeiros dados que surgiram na literatura, focalizando um pequeno grupo de genes do Projeto de Anatomia do Genoma Câncer, sugerem que o número de genes a ser expressados de modo altamente diferenciado (10 vezes ou mais) chegará a várias centenas.

O Projeto Genoma do Câncer Humano (PGCH)

Iniciado em abril de 1999 e conjuntamente financiado pela FAPESP e pelo Instituto Ludwig de Pesquisa do Câncer, o PGCH já produziu mais de 600 mil seqüências de clones de cDNA humano, cada qual com cerca de 300 a 500 pares de bases. Até dezembro de 2000 produzirá cerca de 1 milhão de seqüências. Os clones de cDNA que estão sendo seqüenciados são gerados a partir do mRNA do tecido tumoral de uma série de diferentes tipos de câncer.

O cDNA é gerado por uma nova técnica patenteada pelo Instituto Ludwig, que consiste, essencialmente, em amplificar por PCR a região central dos genes expressos, que é, em geral, a região codificante (Dias Neto e col. 2000). A fase piloto do projeto gerou, entre novembro de 1998 e março de 1999, mais de 10 mil seqüências do câncer de mama, das quais 25% eram seqüências novas, ou seja, representavam seqüências do DNA humano até então desconhecidas e não disponíveis nas bases de dados internacionais e públicas. As 1 milhão de seqüências que estão sendo geradas pelo PGCH são prontamente depositadas no GeneBank, tornando-se assim publicamente disponíveis. Considerando que haja atualmente cerca de 1,6 milhão de seqüências de EST humano disponíveis nas bases de dados internacionais, as seqüências do PGCH representarão uma fração importante das seqüências de EST disponíveis

no mundo.

Um dos cinco Centros de Sequenciamento que compõem o PGCH está localizado no Instituto de Química da USP e é coordenado por um dos autores (S. V-A). Todos os clones de cDNA fisicamente gerados no projeto de sequenciamento ficarão à disposição de seus membros, sendo que o Centro do Instituto de Química coordena também a armazenagem permanente e a distribuição de uma duplicata de cada um dos clones gerados nos cinco Centros do PGCH.

Estudo da expressão gênica em tumores de próstata e de pulmão

Embora ainda não tenham sido identificados os genes que determinam a suscetibilidade ao câncer de próstata e de pulmão, alguns candidatos surgiram nos últimos anos (Offit, 1998). A busca de marcadores genéticos com potencial prognóstico - extremamente importantes, pois podem auxiliar no tratamento de casos de câncer (na indi-

cação de cirurgia, radioterapia, quimioterapia etc.) - padece da mesma escassez de informações.

No câncer de pulmão, o interesse recai sobre os tumores epiteliais, do tipo não-pequenas-células (NPC), que representam 80% de todos os casos. Protocolos de estadiamento molecular estão sendo desenvolvidos a partir de estudos de detecção de mutação e expressão gênica abrangendo vários *loci*: p53, K-ras, MRP-1/CD9, KA11/CD82, bcl-2, erb-b2, RB1 (D'Amico e col. 1999; Miyake e col., 1999). A mesma abordagem está sendo testada para o câncer de próstata — estão em estudo os genes PSA (Gao e col. 1999), p53 e bcl-2 (Moul e col. 1999). Não há, porém, nenhum protocolo de estadiamento molecular estabelecido com valor prognóstico efetivo para NPC ou câncer de próstata.

Nesse cenário, surge como valiosa alternativa o estudo da expressão gênica diferencial em tumores de próstata e de pulmão NSC usando a técnica de *DNA microarrays*. Três grupos de pesquisa do Institu-

to de Química da USP juntaram-se para colaborar nesse estudo, explorando as informações novas produzidas sobre as seqüências do DNA humano: os grupos dos Drs. Mari Sogayar, Hamza El-Dorry e Sérgio Verjovski-Almeida. Esse estudo está sendo realizado em parceria com o corpo clínico de dois grandes hospitais, o Hospital Israelita Albert Einstein (Unidade de Oncologia, chefiada pelo Dr. Sérgio Simon) e o Instituto Nacional do Câncer (Departamento de Urologia, chefiado pelo Dr. Franz Campos), ambos dotados de larga experiência clínica em oncologia, juntamente com dois de seus colaboradores científicos associados ao Instituto de Ciências Biomédicas da USP, os Drs. Marcello Barcinski e Carlos A. Moreira-Filho. Todo esse grupo acima faz parte do subprojeto Genoma do Câncer, coordenado pelo Dr. Sérgio Verjovski-Almeida, dentro de um projeto temático multidisciplinar aprovado e financiado pela FAPESP, denominado Cooperação para a Análise dos Genes e sua Expressão - CAGE. O subprojeto CAGE/Genoma do Câncer irá concentrar-se no câncer de próstata e de pulmão, além de outros a ser oportunamente definidos. No momento, estão sendo selecionados os clones para construção dos primeiros microarrays, na unidade de Microarrays do projeto CAGE (**Figura 3**), sob a coordenação da Dra. Aline Silva, do Instituto de Química, com a tecnologia "Generation III system" desenvolvida pela divisão Molecular Dynamics da Amersham Pharmacia Biotech. O projeto CAGE é um dos primeiros centros acadêmicos no mundo, e o primeiro da América Latina, a ter acesso ao Microarray System Program da Molecular Dynamics, que até o início de 2000 estava restrito apenas às grandes empresas farmacêuticas e de biotecnologia. Paralelamente, está sendo criado um banco de amostras de tecido fresco congelado, extraído de tumores coletados nos dois hospitais, que fornecerá o

Robô para construção de Chips de DNA

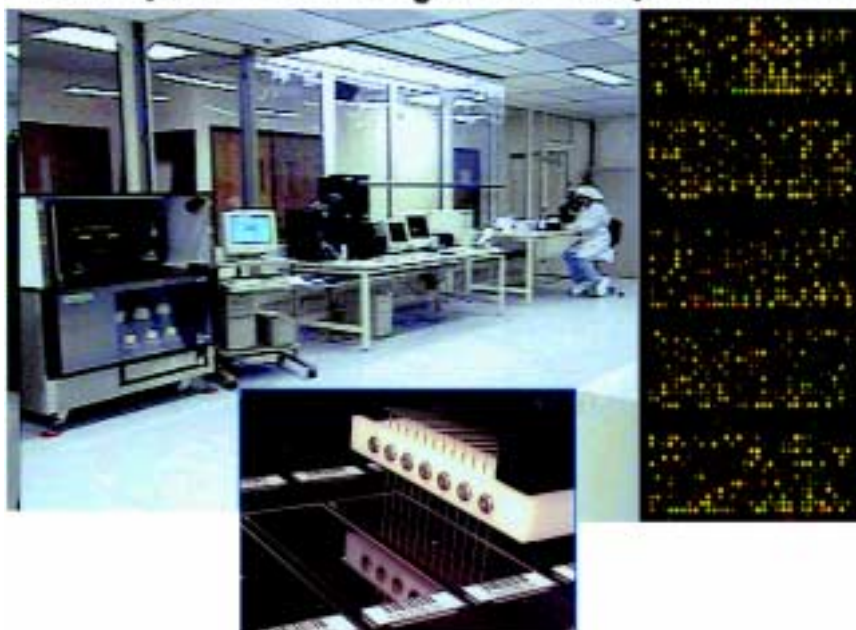


Figura 3 - Unidade de *microarrays*, onde se emprega a tecnologia Generation III da Amersham Pharmacia Biotech (foto publicada com permissão da Amersham Pharmacia Biotech)

material para extração de mRNA a ser usado na hibridação.

A investigação cuidadosa de padrões de expressão gênica em tecidos neoplásicos é bem capaz de revelar as seqüências do conjunto de múltiplos genes que determinam o resultado do processo tumoral. Além disso, será possível esclarecer o papel das mutações na linhagem germinativa de possíveis genes do câncer, associando-os a determinados padrões de expressão gênica diferencial. Esse processo começa a ser elucidado nas formas hereditárias de câncer de mama, onde já se demonstrou que mutações do gene BRCA1 interferem na supressão de vias de transcrição dependentes de estrogênio relacionadas com a proliferação epitelial mamária (Fan e col., 1999) e que o produto do gene BRCA2 inibe a atividade de transcrição do p53 (Marmorstein e col. 1998).

Agradecimentos.

Os autores agradecem à FAPESP pelo apoio financeiro.

Referências

ASCO (1996). Statement of the American Society of the Clinical Oncology. Genetic Testing for cancer Susceptibility. **Journal of Clinical Oncology** 14: 1730-1736.

Blackwood A & Weber BL. BRCA1 and BRCA2 (1998). From molecular genetics to clinical medicine. **Journal of Clinical Oncology** 16: 1969-1977.

D' Amico TA, Massey M, Herndon JR, Moore MB, Harpole DH Jr. (1999). A biologic risk model for stage I lung cancer: immunohistochemical analysis of 408 patients with the use of ten molecular markers. **J Thorac Cardiovasc Surg** 117:736-43

Dias Neto e col. Shotgun sequencing of the human transcripto-

me with ORF expressed sequence tags (2000). **Proc Natl Acad Sci USA** 97:3491-3496.

Ellisen LW & Haber D (1998). Hereditary Breast Cancer. **Annual Rev. Medicine** 49:425-436.

Fan S e col. (1999) BRCA1 Inhibition of Estrogen Receptor Signaling in Transfected Cells. **Science** 284:1354-1356.

Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE (1994). Risks of cancer in BRCA1 mutation carriers. Breast Cancer Linkage Consortium. **Lancet** 343: 692-695.

Ford D e col. (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. **American Journal of Human Genetics** 62: 676-689.

Gao CL e col. (1999). Detection of circulating prostatic antigen expressing prostatic cells in the bone marrow of radical prostatectomy patients by sensitive reverse transcriptase. **J Urology** 161: 1070-1076.

Gayther SA e col. (1997). Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. **Nature Genetics** 15:103-105.

Golub TR e col. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. **Science** 286: 531-537.

Harrington CA, Rosenow C, Rietz J (2000). Monitoring gene expression using DNA microarrays. **Current Opinion in Microbiology** 3:285-291.

Lee PS & Lee KH (2000). Genomic Analysis. **Current Opinion in Biotechnology** 11:171-175.

Marmorstein LY, Ouchi T, Aa-

ronson SA (1998). The BRCA2 gene product functionally interacts with p53 and RAD51. **Proc Natl. Acad. Sci USA** 95:13869-1374.

Miki Y e col.(1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. **Science** 266: 66-71,1994.

Miyake M. e col. (1999). A novel molecular staging protocol for non-small cell lung carcinoma. **Oncogene** 18:2397-2404.

Moul JW. (1999). Angiogenesis, p53, bcl-2 and ki-67 in the progression of prostatic cancer after radical prostatectomy. **Eur Urol** 35:399-407.

Nagamine CM, Chan K, Lau CYF (1989). A PCR artifact: Generation of heteroduplexes. **American Journal of Human Genetics** 45: 337- 339.

Newman B, Austin MA, Lee N, King MC (1988). Inheritance of human breast cancer : Evidence for autosomal dominant transmission in high-risk families. **Proc. Natl. Acad. Sci. USA** 85: 3044-3048.

Offit K (1998). **Clinical Cancer Genetics**. Wiley-Liss, New York.

Orita M, Iwahana H, Knazawa H, Hayashi K, Sekiya T (1989). Detection of polymorphisms of human DNA by gel electrophoresis as single – strand conformation polymorphisms. **Proc. Natl. Acad. Sci USA** 86: 2766 – 2770, 1989.

van Oursow NJ, Dhanda RK, Elhaji Y, Narod SA, Li FP, Eng C, Vijg J (1999). A highly accurate, low cost test for BRCA1 mutations. **J Med Genet** 36:747-753.

Wooster R e col. (1995) Identification of the Breast Cancer Susceptibility Gene BRCA2. **Nature** 378: 789-792.

