Methods and biomarkers for the diagnosis and prognosis of cancer and other diseases: Towards personalized medicine

Judy E. Anderson d,f,1, Lise Lotte Hansen a,1, Frank C. Mooren b, Markus Post b, Hubert Hug c, Anne Zuse e, Marek Los d,e,f,g,∗

a Institute of Human Genetics, University of Aarhus, Denmark
b Department of Sports Medicine, Institute of Sport Sciences, University Giessen, Germany
c DSM Nutritional Products Ltd, Research & Development, Kaiseraugst, Switzerland
d Department of Human Anatomy and Cell Science, University of Manitoba, Winnipeg, Canada
e Manitoba Institute of Cell Biology (MICB), CancerCare Manitoba, 675 McDermot Ave. Rm. ON6010, Winnipeg, Man. R3E 0V9, Canada
f Manitoba Institute of Child’s Health (MICH), University of Manitoba, Winnipeg, Canada
g Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Canada

Received 6 August 2006; received in revised form 22 August 2006; accepted 23 August 2006

Abstract

The rapid development of new diagnostic procedures, the mapping of the human genome, progress in mapping genetic polymorphisms, and recent advances in nucleic acid- and protein chip technologies are driving the development of personalized therapies. This breakthrough in medicine is expected to be achieved largely due to the implementation of “lab-on-the-chip” technology capable of performing hundreds, even thousands of biochemical, cellular and genetic tests on a single sample of blood or other body fluid. Focusing on a few disease-specific examples, this review discusses selected technologies and their combinations likely to be incorporated in the “lab-on-the-chip” and to provide rapid and versatile information about specific diseases entities. Focusing on breast cancer and after an overview of single-nucleotide polymorphism (SNP)-screening methodologies, we discuss the diagnostic and prognostic importance of SNPs. Next, using Duchenne muscular dystrophy (DMD) as an example, we provide a brief overview of powerful and innovative integration of traditional immuno-histochemistry techniques with advanced biophysical methods such as NMR-spectroscopy or Fourier-transformed infrared (FT-IR) spectroscopy. A brief overview of the challenges and opportunities provided by protein and aptamer microarrays follows. We conclude by highlighting novel and promising biochemical markers for the development of personalized treatment of cancer and other diseases: serum cytochrome c, cytokeratin-18 and -19 and their proteolytic fragments for the detection and quantitation of malignant tumor mass, tumor cell turn-over, inflammatory processes during hepatitis and Epstein–Barr virus (EBV)-induced hemophagocytic lymphohistiocytosis and apoptotic/necrotic cancer cell death.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: CYFRA 21-1; Cytokeratin fragments; Cytochrome c; Duchenne muscular dystrophy; Lab-on-the-chip; TPS; Personalized cancer therapy

1. Introduction

In recent years, tremendous progress has been made in treating cancer and other diseases due to rapid advances on multiple fronts. Currently, experimental drugs can be designed and targeted to different tissues in the body, or to specific types of cancer (Booy et al., 2005; Gills et al., 2004; Hauff et al., 2005; Kreuter et al., 2004; Mendoza et al., 2005; Kleespies et al., 2006). Strategies under development will not only allow for selective identification of known malfunctioning signaling pathways (Adjei and Rowinsky, 2003; Hashemi et al., 2005), but also of entities that differ between normal and malignant phenotypes without necessarily knowing the molecular changes that are causing a malignancy (Maddika

∗ Corresponding author at: Manitoba Institute of Cell Biology (MICB), CancerCare Manitoba, University of Manitoba, 675 McDermot Ave. Rm. ON6010, Winnipeg, Man. R3E 0V9, Canada. Tel.: +1 204 787 2294 (O)/1403 (Lab); fax: +1 204 787 2190.
E-mail address: losmj@cc.umanitoba.ca (M. Los).
1 Both authors contributed equally to the manuscript.

1368-7646/$ – see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.drup.2006.08.001
et al., 2005, 2006; Mitrus et al., 2005; Robert et al., 2002). In this review, we present selected approaches and their combinations that are likely to have a marked impact on future medical diagnosis and treatment of cancer and other diseases (see next paragraphs).

Changes or shifts in cellular metabolism or sub-cellular structure occurring in pathologies, adaptations or carcinogenic processes are often measured and applied as an index of functional change. Such studies require a significant sample size, often in the format of histological sections prepared through special stains or through monitoring the content of a particular protein or lipid deposited as a result of the pathology. Techniques, including immuno-staining and in situ hybridization, are now used to visualize the level of gene expression, abundance of abnormal proteins or gene transcripts, or to identify the tissue lineage or germ layer origin with high specificity in diagnostic molecular-pathology laboratories. The development of these tools has kept close pace with the rapid progress in basic science and applied clinical research (Table 1). In addition, laboratories have benefited

### Table 1
Novel biochemical markers for the ex vivo-monitoring of cell death

<table>
<thead>
<tr>
<th>Biochemical marker</th>
<th>Disease or pathology</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytochrome c</td>
<td>Cancer therapy</td>
<td>Barczyk et al. (2005), Renz et al. (2001a,b)</td>
</tr>
<tr>
<td></td>
<td>- detected in serum and plasma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- accompanies cancer therapy in vivo</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- may indicate in vivo tumor cell turn-over and tumor mass</td>
<td></td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>Epstein–Barr virus infection</td>
<td>Kawada et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>- detected in serum</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correlates with onset of in vivo-apoptosis triggered during Epstein–Barr virus-caused hemophagocytic lymphohistiocytosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correlates with Epstein–Barr virus-load</td>
<td></td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>Liver disease</td>
<td>Barczyk et al. (2005), Ben-Ari et al. (2003), Ghavami et al. (2005), Sakaida et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>- detected in serum and plasma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- strongly correlates with clinical symptoms of fulminant liver damage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correlates with severity of hepatic coma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- strongly correlates with biochemical marker levels for liver damage: serum hepatocyte growth factor, aspartate aminotransferase (AST), lactate dehydrogenase (LDH), and alkaline phosphatase (ALP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- negatively correlates with serum alpha-fetoprotein (AFP) and total plasma bilirubin levels</td>
<td></td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>Neuronal damage</td>
<td>Ahlemeyer et al. (2002), Satchell et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>- detected in cerebrospinal fluid</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- associated with inflicted traumatic brain injury, observed more frequently in female patients</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- detected in extracellular space upon staurosporine-induced neuronal cell death, may mediate bystander effect under these experimental conditions</td>
<td></td>
</tr>
<tr>
<td>Caspase-1</td>
<td>Neuronal damage</td>
<td>Satchell et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>- detected in cerebrospinal fluid</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- associated with traumatic brain injury</td>
<td></td>
</tr>
<tr>
<td>Caspase-2, -3, 7</td>
<td>detected in extracellular space</td>
<td>Cowan et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>- participate in degradation of excess extracellular matrix during remodeling in development and disease (in vitro data)</td>
<td></td>
</tr>
<tr>
<td>Cytokeratin-18</td>
<td>Malignancies of epithelial origin, prostate-, breast, other cancers</td>
<td>Kramer et al. (2004), Sheard et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>- detected in serum and plasma</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- cytosolic pool of soluble full-length cytokeratin-18 released during necrosis; apoptosis is associated with significant release of caspase-cleaved cytokeratin-18 fragments, mainly the 13 kDa fragment (TSP)</td>
<td></td>
</tr>
<tr>
<td>Cytokeratin-19</td>
<td>Malignancies of epithelial origin, prostate-, breast, other cancers</td>
<td>Sheard et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>- detected in serum and plasma, culture supernatants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- ~30 kDa fragment of cytokeratin-19, called also “water-soluble cytokeratin-19 fragment” or CYFRA 21-1, is an indicator in vivo apoptotic cell death</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- generation of CYFRA 21-1 during apoptosis, can be inhibited by caspase inhibitors</td>
<td></td>
</tr>
</tbody>
</table>
from the commercial availability of a wide range of reagents and kits, high-throughput strategies for screening, and highly focused training programs in medicine, laboratory sciences, clinical chemistry, pathology, and imaging. These advances have raised the expectations of investigators, technologists, trainees, and the consuming public for levels of specificity and sensitivity that bring high predictive value to a diagnostic test.

In addition to immuno-histological testing, recent advances in high-throughput screening techniques including whole transcriptome evaluation (mRNAs), proteome (proteins), metabolomic approaches, and analysis of single-nucleotide polymorphisms (SNPs), either restricted to selected loci or together in haplotype blocks, appear to offer a powerful predictive tool set for both monitoring related to prophylaxis and prevention, and for treatment selection (Alaou-Jamali et al., 2004; Pusztai et al., 2004; Kroczak et al., 2006).

A mutation is denoted polymorphic if the minor allelic frequency exceeds 1% in a population. SNPs are estimated to occur with one in 300 bp in the human genome and it is predicted that each person carries from 24,000 to 40,000 non-synonymous SNPs (Cargill et al., 1999; Judson and Stephens, 2001; Kruglyak and Nickerson, 2001; Reich et al., 2003). In the search for specific selective criteria, Zhu et al. compared 166 studies of more than 400 SNPs within 129 cancer-related and DNA-repair genes. One hundred and three SNPs were found to be non-synonymous and highly conserved among species, and a strong significant association was found between the odds ratio for cancer risk and the level of conservation of the individual SNP (Zhu et al., 2004).

Despite comprehensive studies, only few SNPs have been found to be statistically strong prognostic markers for breast cancer. In one study, 25,000 SNPs were selected from the 125,799 SNPs mentioned previously, and the genotypes were determined in 254 German breast cancer patients and in 268 age-matched women without malignant disease. The region 14q24.3–q31.1 associated weakly with breast cancer status and was further mapped using high-density SNPs. The genotype of just one SNP within intron 1 of the zinc-finger gene DPF3/CEMD4 correlated significantly with breast cancer status (OR = 1.6; p = 0.003), increased lymph-node metastasis (p = 0.006), age of onset (p = 0.01) and tumor size (p = 0.01). In a similar study of 25,000 SNPs located within or close to 16,000 genes, a 20 kb region at 19p13.2 was identified to be associated to the risk of acquiring breast and prostate cancer (OR = 3.4; p = 0.001) (Hoyal et al., 2005). The risk was strongest for patients with a family history of cancer. Further mapping of the region revealed a strong, significant correlation between one SNP in the Intercellular Adhesion Molecule 5 Precursor (ICAM5) and disease progression and prognosis (Kammerer et al., 2004). Specific SNPs in p53, estrogen receptor and progesterone receptor were found to be associated with a decreased risk of breast cancer, whereas SNPs in HRAS, GSTM1 and CYP19, as well as genes involved in DNA repair XRCC1, XRCC3, ERCC4/XPF, BRCA2 and RAI correlated with an increased risk of breast cancer (de Jong et al., 2005; Dimitrescu and Cotarla, 2005; Nexo et al., 2003).

SNP genotypes are important fingerprints of our genome, and once determined they are straightforward to analyze on a large scale. SNPs can be used as prognostic and predictive markers either alone or together in haplotype blocks. Furthermore, multiple SNPs can be used to identify new cancer-susceptibility genes via population or family-based linkage-disequilibrium studies. Below, we have chosen breast cancer as an example to demonstrate the potential of SNP screening for prognostic and therapeutic purposes.

2. Single-nucleotide aberrations as prognostic markers in breast cancer

How might informative SNPs be selected for prognostic studies? Most SNPs in the database are not validated, and many are not polymorphic. Arrays comprising targets of 204,200 SNPs extracted from the public databases have been constructed (Nelson et al., 2004). Only 125,799 SNPs (61.6%), which fulfilled the criterion of being located within 10 kb of a known or predicted gene with a minor allelic frequency of more than 2%, were polymorphic in the analyzed population. In addition, the genotype frequency may differ considerably between populations, as seen in studies of the HER-2 SNP I655V in which the minor allele frequency ranged from 0 to 24% among different populations (Ameyaw et al., 2002). In the search for specific selective criteria, Zhu et al. have raised the expectations of investigators, technologists, trainees, and the consuming public for levels of specificity and sensitivity that bring high predictive value to a diagnostic test.

In addition to immuno-histological testing, recent advances in high-throughput screening techniques including whole transcriptome evaluation (mRNAs), proteome (proteins), metabolomic approaches, and analysis of single-nucleotide polymorphisms (SNPs), either restricted to selected loci or together in haplotype blocks, appear to offer a powerful predictive tool set for both monitoring related to prophylaxis and prevention, and for treatment selection (Alaou-Jamali et al., 2004; Pusztai et al., 2004; Kroczak et al., 2006).

A mutation is denoted polymorphic if the minor allelic frequency exceeds 1% in a population. SNPs are estimated to occur with one in 300 bp in the human genome and it is predicted that each person carries from 24,000 to 40,000 non-synonymous SNPs (Cargill et al., 1999; Judson and Stephens, 2001; Kruglyak and Nickerson, 2001; Reich et al., 2003). These SNPs are especially interesting as they may affect the translation, exon splicing and the function of a protein. Equally important is the identification of SNPs in non-coding regions as they may influence the level of transcription or the correct exon splicing.

SNP genotypes are important fingerprints of our genome, and once determined they are straightforward to analyze on a large scale. SNPs can be used as prognostic and predictive markers either alone or together in haplotype blocks. Furthermore, multiple SNPs can be used to identify new cancer-susceptibility genes via population or family-based linkage-disequilibrium studies. Below, we have chosen breast cancer as an example to demonstrate the potential of SNP screening for prognostic and therapeutic purposes.

2. Single-nucleotide aberrations as prognostic markers in breast cancer

How might informative SNPs be selected for prognostic studies? Most SNPs in the database are not validated, and many are not polymorphic. Arrays comprising targets of 204,200 SNPs extracted from the public databases have been constructed (Nelson et al., 2004). Only 125,799 SNPs (61.6%), which fulfilled the criterion of being located within 10 kb of a known or predicted gene with a minor allelic frequency of more than 2%, were polymorphic in the analyzed population. In addition, the genotype frequency may differ considerably between populations, as seen in studies of the HER-2 SNP I655V in which the minor allele frequency ranged from 0 to 24% among different populations (Ameyaw et al., 2002). In the search for specific selective criteria, Zhu et al. compared 166 studies of more than 400 SNPs within 129 cancer-related and DNA-repair genes. One hundred and three SNPs were found to be non-synonymous and highly conserved among species, and a strong significant association was found between the odds ratio for cancer risk and the level of conservation of the individual SNP (Zhu et al., 2004).

Despite comprehensive studies, only few SNPs have been found to be statistically strong prognostic markers for breast cancer. In one study, 25,000 SNPs were selected from the 125,799 SNPs mentioned previously, and the genotypes were determined in 254 German breast cancer patients and in 268 age-matched women without malignant disease. The region 14q24.3–q31.1 associated weakly with breast cancer status and was further mapped using high-density SNPs. The genotype of just one SNP within intron 1 of the zinc-finger gene DPF3/CEMD4 correlated significantly with breast cancer status (OR = 1.6; p = 0.003), increased lymph-node metastasis (p = 0.006), age of onset (p = 0.01) and tumor size (p = 0.01). In a similar study of 25,000 SNPs located within or close to 16,000 genes, a 20 kb region at 19p13.2 was identified to be associated to the risk of acquiring breast and prostate cancer (OR = 3.4; p = 0.001) (Hoyal et al., 2005). The risk was strongest for patients with a family history of cancer. Further mapping of the region revealed a strong, significant correlation between one SNP in the Intercellular Adhesion Molecule 5 Precursor (ICAM5) and disease progression and prognosis (Kammerer et al., 2004). Specific SNPs in p53, estrogen receptor and progesterone receptor were found to be associated with a decreased risk of breast cancer, whereas SNPs in HRAS, GSTM1 and CYP19, as well as genes involved in DNA repair XRCC1, XRCC3, ERCC4/XPF, BRCA2 and RAI correlated with an increased risk of breast cancer (de Jong et al., 2005; Dimitrescu and Cotarla, 2005; Nexo et al., 2003).

The proto-oncogene HER-2 is overexpressed in 10–34% of primary breast carcinomas, and the level of HER-2 over-expression is a strong prognostic and predictive marker. One intragenic SNP I655V has been intensively studied in breast cancer patients, although without conclusive results. To explore the possibility of establishing a prognostic haplotype, six SNPs assigned to one haplotype block within HER-2 were analyzed in a breast cancer cohort. None of the SNPs alone correlated with any prognostic parameter. However, for the six SNPs together, one specific haplotype was found to correlate with increased expression of HER-2 (p = 0.009) and poor survival (p = 0.03) (Han et al., 2005). There was also significant correlation between a specific haplotype of Aurora-A and increased risk of breast cancer, despite there being no relation between disease parameters and individual SNPs in the gene (Lo et al., 2005). Identification and characterization of many more informative SNPs, anticipated to occur over the next few years, will provide an important tool to identify new susceptibility genes via linkage disequilibrium studies and to assess new prognostic and predictive markers via single SNPs or haplotypes.
3. Methods to screen for new SNPs and large-scale determination of SNP genotypes

As indicated in Section 1, SNPs constitute the most common variation in the human genome, contributing to a variety of phenotypes, predisposition to complex diseases and response patterns to therapy (Hirschhorn and Daly, 2005; Robert et al., 2005). More than 1 million of the 30 million estimated SNPs present in the human genome have been thoroughly characterized using 269 individuals from four different populations via the International Haplotype Mapping Project (http://www.hapmap.org/). The genome databases have published more than 2 million SNPs, the majority of which has to be validated in larger, more divergent, study cohorts.

The methodologies to analyze SNPs can be divided into two groups depending on whether the SNP is well-characterized or the method focuses on the search for still-unknown SNPs. The overall aim is to establish the genotype of each SNP in large populations. In the search for unknown mutations, a large spectrum of mutation-detection methods is available. Many methods are based on the detection of mismatched nucleotides or differences in melting conditions after heteroduplex formation. Denaturing and recoiling of the DNA strands, either from the same or from two different individuals, will form a heteroduplex. Two techniques, Denaturing Gradient Gel Electrophoresis (DGGE) and Denaturing High Performance Liquid Chromatography (DHPLC) use the principle that strand separation of a heteroduplex is faster under the influence of an increasing denaturing gradient than is a homoduplex (Donohoe, 2005; Kuklin et al., 1997; Myers et al., 1985a,b). Both methods require a thorough design of primers that considers the melting domains of the DNA fragment. For DHPLC, the conditions of temperature and denaturation used in analysis are crucial. Once the conditions are established for a specific DNA fragment or for a mutation, analysis is fast and has a high throughput. Still, the genotype has to be established.

The Chemical Cleavage of Mismatches (CCM) and enzyme-based methods cleave the DNA strand if mispaired bases are present in a heteroduplex. CCM is based on the principles of Maxam and Gilbert sequencing (Cotton, 1989; Cotton et al., 1988). A mispaired thymine is modified by OsO₄ or KMnO₄ and a mispaired cytosine by hydroxylamine. The sugar phosphate backbone is cleaved by piperidine at the respective modified bases. The two DNA strands are initially labeled with two different fluorescent dyes, and separation by electrophoresis after mutation-specific cleavage provides information on the location of the mutation and identifies the mispaired base, thereby providing the genotype (Hansen and Justesen, 2003). This method is highly sensitive; close to a 100% of all mutations are detected. Even the presence of only 5% DNA with another genotype (at a specific locus) for a specific mutation can be detected, and the method is suitable for screening large fragments (up to 1.500 bp) (Hansen et al., 1996).

Enzymatic cleavage of mispaired bases in a heteroduplex by, for example, the endonuclease CEL 1 is a highly efficient mutation-detection method. CEL 1 is isolated from celery and has a high specificity for insertions, deletions and base-substitution mismatches (Kulinski et al., 2000; Oleykowski et al., 1998). Fragments up to 2 kb can be screened with a high throughput. Both DNA strands of a heteroduplex are labeled with a fluorescent dye, and the nuclelease CEL 1 cleaves the DNA strand at a mispaired base. The cleavage is not complete and consequently several mutations in one fragment are detected by the final electrophoresis. This method reveals close to 100% of all mutations, but the genotype still has to be determined by another methodology.

Several different methods to establish the genotype of a SNP are based upon the primer-extension principle. The Sanger-sequencing principles are used in the semi-automatic primer-extension method, in which a primer is designed to anneal the 3′-end to the nucleotide preceding the SNP (Makridakis and Reichardt, 2001). The primer is extended by a polymerase chain reaction (PCR) with the nucleotide complementary to the polymorphism. In contrast to the common PCR, only dideoxynucleotides, labeled with a base-specific fluorescent dye, are present. Multiple SNPs can be analyzed in one primer-extension reaction. The primers are designed to differ in length and will be separated by electrophoresis. Commercial kits are available for this assay.

For high throughput analysis the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) is an efficient method that allows accurate SNP genotyping and sequence validation (Nelson et al., 2004; Smylie et al., 2004). The technique requires specific equipment, and is then cost-effective. DNA fragments up to 450 bp can be analyzed. The method is based upon the primer-extension principle, as above. The primer anneals 3′ to the SNP and is extended by incorporating unlabeled dNTPs. The 5′ phosphodiester bonds of each newly incorporated pyrimidine nucleotide are then replaced by acid-labile phosphoamidite (P–N) bonds. The template strand is attached to magnetic beads through biotin-streptavidin binding. The P–N bonds are cleaved by hydrolysis and the small fragments are subjected to MALDI-TOF separating the fragments according to size.

Oligonucleotide-based or allele-specific primer-extension arrays are designed for large-scale genotyping of known SNPs (Pastinen et al., 2000; Syvanen, 2005). The methodology requires specific equipment and commercially available arrays are expensive, which often restricts the analysis to small populations. The low sample number thereby limits correlation and comparative studies. The advantage of these methods is that the genotype of thousands of SNPs is identified from each individual, thereby providing a screen across pathways and gene families, depending on the design of the arrays.

A second advantage of the array-based methodology is that it provides the possibility of screening for allelic imbalance or loss of heterozygosity in the tumor genome (Mei et al.,...
2000). DNA from tumor and non-malignant tissue from the same individual is analyzed, and the intensity of the signals obtained from each allele is compared between two tissues (samples). Use of micro-dissected tumor DNA provides a clear picture of allelic loss for each heterozygous SNP locus. The high throughput technologies will provide a large quantity of results, comprising population-specific mutations, new SNPs and genotype frequencies. One SNP may be highly polymorphic in one population but solely homozygous in another. There is no doubt that SNP characterization will prove an important tool in the identification of predictive and prognostic markers in many diseases.

Finally, flow cytometry-based methods have been developed, mainly for clinical detection of selected-, pre-identified SNPs. While these methods may provide lower throughput compared to array-based methods, the availability of flow cytometry equipment in virtually every clinical diagnostic lab, makes bead-based assay where DNA-structures are attached to beads, very attractive and enables easy implementation. In one configuration of the bead-based assay, biotinylated primers that stop one base before the SNP site are attached to avidin-coated beads. DNA from the test subjects, either genomic DNA or PCR-amplified segments that span the SNP site, is hybridized to the primers. At this point, beads with numerous attached copies of the primer/template duplexes are divided into four aliquots to which dideoxy nucleotides are added along with DNA polymerase. In each of the aliquots, only one of the terminating nucleotides is fluorescently labeled. Thus, upon analysis of each of the four bead sets in flow cytometer for a homozygous SNP, only one bead type will be fluorescent. This assay can be multiplexed by the addition of unique DNA-sequence tags to the primers and sequences that are complementary to the tags attached to beads, that are different for each bead type of the multiplex set (Cai et al., 2000). After thermal cycling the single-base primer-extension reaction occurs at multiple sites on the target DNA, the tagged DNA primers are melted from the template DNA, and the tagged beads pick up their respective extended primers. Analysis in a flow cytometer of the bead-associated fluorescence determines the specific nucleotide located at the SNP site. Sequence tags or “ZipCodes” have also been used for SNP analysis by oligonucleotide ligation (OLA) (Iannone et al., 2000) and single-base extension (Chen et al., 2000) assays. A fluorescently labeled oligonucleotide reporter sequence is added to a “capture” probe by OLA. Capture probes are designed to hybridize both to genomic “targets” amplified by the polymerase chain reaction and to a separate complementary DNA sequences that has been coupled to a microsphere. These sequences on the capture probe are called “ZipCodes”. OLA-modified capture probes are hybridized to ZipCode complement-coupled microspheres. The use of microspheres with different ratios of red and orange fluorescence makes possible a multiplexed format, in which many SNPs may be analyzed in a single tube. Flow cytometric analysis of the microspheres simultaneously identifies both the microsphere type and the fluorescent green signal associated with the SNP genotype (Iannone et al., 2000). In the single-base extension assay, the ZipCode at the 5’-end of the capture fluorescent oligonucleotide probe allows the DNA-polymerase reaction product to be captured by its complementary sequence (cZipCode), which has been coupled to a specific fluorescent microsphere (Chen et al., 2000). Both assays can be analyzed using standard flow cytometry, or with the much less expensive bead-analyzer (Luminex systems).

4. Combination of traditional immuno-histochemistry with non-invasive techniques, like NMR and metabolic studies—the example of muscular dystrophies diagnosis

Many higher resolution methodologies take advantage of recent advances in the understanding of basic molecular, genetic and biochemical etiology of a particular pathology. As we learn the details of molecular signal-transduction pathways of the genetic processes that build proteins and sub-cellular structures, it is clear that measures of a biochemical attribute are demonstrating function. It becomes, then, an issue of understanding the meaning of a sophisticated measure or index, and the impact on function of observing some measure outside the normal range.

4.1. Combining methodological approaches

Combinations of classical cell biology methods, including autoradiography and in situ hybridization, with biophysical methodologies such as nuclear magnetic resonance (NMR) or infrared spectroscopy, have successfully been applied to monitor changes in physiology. By such combination, it is possible to observe patterns of global changes in tissue biochemistry as a tissue shifts from normal to pathological. For example, a lower than normal concentration of taurine and carnitine in dystrophic muscles from the muscular dystrophy (mdx) mouse model of Duchenne muscular dystrophy (DMD) was revealed using proton NMR spectroscopy, while taurine concentration was increased in regenerating muscle (DMD) was revealed using proton NMR spectroscopy, while taurine concentration was increased in regenerating muscle. Further studies showed that steroid treatment of mdx mice reduced the severity of disease progression and also increased the levels of taurine and carnitine (McIntosh et al., 1998a,c).

These findings are consistent with literature reports on diet-induced cardiomyopathy with or without neurological signs such as epilepsy (Barbeau, 1975) in cats (Michel, 2005) and humans (Fosslien, 2003; Georges et al., 2003; Karmaniolas et al., 2002; Vescovo et al., 2005). These cases too are deficient in taurine and/or carnitine due to genetic mutations that affect taurine and carnitine transport or metabolism. Recent reports on myopathies including DMD and mdx mouse muscular dystrophy have incorporated a taurine treatment arm (De Luca et al., 2002; Voisin and de la Porte, 2005), with considerable therapeutic benefit.
Results were also in accordance with the findings of other metabolic and pathological indices of disease progression. These included changes in the lipid components of membranes (Gillet et al., 1993), pH and energy metabolism using proton and phosphorous NMR spectroscopy in muscle (Dunn et al., 1991, 1992; Dunn and Radda, 1991; Radda, 1999) and/or brain (Rae et al., 1998; Tracey et al., 1996a,b), and treatment responses that were observed in skeletal muscle including the diaphragm, and in cardiac muscle (Anderson et al., 1998, 1996; McIntosh et al., 1998a,b,c, 1994; Skrabek and Anderson, 2001).

Importantly these studies employed a tandem approach to the hallmark phenotypic changes (i.e., those marking the dysfunction in a particular pathology) during disease progression and indices of cell or molecular biology. First, they allowed a statistical correlation of the relationship between a prominent histological feature of skeletal muscle regeneration, the fiber central nucleation index (CNI), with features of the proton-NMR spectrum. In adult muscle, with the exception of a few intra-fusal (muscle spindle) fibers, muscle fibers display peripherally-located myonuclei. Central myofiber nuclei are apparent only in those parts of muscle fibers that have undergone one or more regenerative events, and CNI is used as an index of accumulated disease progression (Karpati et al., 1988). An early NMR spectroscopy study correlated a number of metabolites including peaks from taurine in a muscle sample, with CNI in sections of the same muscle (McIntosh et al., 1998a). This was an essential step in developing an understanding that metabolite studies by NMR could provide a powerful tool for non-invasive diagnostic studies of muscle pathology and treatment effects.

Second, the expression of myf5, a muscle regulatory gene with a key role in early development and regeneration of skeletal muscle, was also correlated with taurine levels. This supported the notion that taurine levels measured by NMR spectroscopy were a reasonable index of muscle-specific processes including the regenerative response to treatment by dystrophic muscle (McIntosh et al., 1998a,c); and possibly also in other muscle pathologies. The use of magnetic resonance imaging in detecting the presence and extent of muscle lesions has been examined in dystrophy in mdx mice (McIntosh et al., 1998b), and after stem cell therapy (Walter et al., 2004). A similar study using Fourier-transformed infrared (FT-IR) spectroscopy of mdx muscles was able to follow the progression of fibrosis (collagen content and accumulation) during dystrophy, and also traced the partial alleviation of collagen accumulation after treatment with steroids (Shaw et al., 1996). These studies from the muscle literature, illustrate how a correlation of measures taken independently from different approaches can extend the impact of the conclusions that are reached, and similar combination studies of cancer diagnostics, screening, prognostic and treatment outcome have been shown highly advantageous (McIntosh et al., 1999; Kirkham et al., in press; Shah et al., 2006; Delorme and Weber, 2006; Weber et al., 2006; Huzjan et al., 2005).

4.2. Challenging the phenotype

Other approaches to the goal of finding a “phenotype” in a whole transgenic or mutant strain (one created by knowing for example, that a particular gene has an important function in normal regulatory mechanisms) include challenging of the system, either a tissue in culture or in vivo, to perform some function, and to observe the capability for a tissue to achieve an expected outcome. Often this is a sensitive approach to “finding a phenotype” where one might not be obvious. An example of this approach is exploring skeletal muscle function after a knockout mutation in one of the muscle regulatory genes. While the animals that were 5MyoD−/− developed normally and had normal muscle by histological examination, studies of muscle challenged to regenerate revealed that there was a phenotypic change that delayed the process of muscle regeneration (Megeney et al., 1996). The further challenge to muscle without MyoD expression, by testing muscle function and regeneration on the background of muscular dystrophy (in the mdx mouse) showed that the impact of muscular dystrophy was far greater in both skeletal and cardiac muscle that did not express MyoD (McIntosh et al., 1998b; Megeney et al., 1999). Finding the sensitive and specific test that will reveal a particular genetic, biochemical, or functional phenotype in a highly predictable manner is a significant challenge to investigators, and one to which imaging techniques and metabolomic approaches have contributed rapid and powerful advances.

5. Protein arrays and antibody microarrays as versatile screening tools

With the completion of the Human Genome Project, focus has now shifted to understanding what is the function of our ~20,000–30,000 genes. Thus, technologies that globally study the structure, function and expression of proteins are gaining importance not only in research settings but increasingly, also in clinical practice. Several techniques can be used to identify proteins present in cells, tissues, body fluids, or an organism. Among these, antibody-microarrays, protein-microarrays, peptide-microarrays, and even aptamer-based microarrays (see below) are gaining importance, as versatile, rapid, and high-throughput screening methods. The aptamer-, peptide-, and antibody-microarrays are easier for development than other types of protein-based microarrays, in part due to the significant knowledge about interaction principles typical for these molecules. It is worth noticing that aside from diagnostic procedures and antibody array development, antibodies, either naked or in a modified form are also increasingly being used in the clinic for therapeutic purposes (Booy et al., 2006; Brown and Gibson, 2005; Johnston et al., in press; Krzemieniacki et al., 2006). The above-mentioned microarrays detect specific interactions between proteins or proteins and other molecules. Below, we discuss various aspects related to the development and implementation of
this versatile technology, for the detection of therapeutic drug effects, for example.

The different components of protein and antibody microarray technology, i.e., microarray surfaces, immobilization techniques, probe and substrate design, detection systems and knowledge-management systems to handle the generated data are all under rapid development (Pavlickova et al., 2004; Wingren and Borrebaeck, 2004; Zangar et al., 2005). Compared to traditional assay systems advantages of microarrays are higher sensitivity and parallelization. There are attempts to increase the number of probes to up to 10,000 per protein microarray (Wingren and Borrebaeck, 2004). One of the biggest challenges is finding equal binding conditions for the various probes and substrates since the physical and chemical parameters of protein–protein (antibody–antigen) interactions can vary considerably for different proteins and cannot as easily be predicted as Watson–Crick base pairing in the case of nucleic acid micro-arrays. Very often protein binding requires co-factors or post-translational modifications. For these reasons parallelization is more difficult than for DNA microarray technology. But the advantage is that proteins can be directly linked to phenotypes and disease status (Bodovitz et al., 2005).

An alternative to microarrays may be the use of beads for immobilization of the probes or the capture and detection of proteins. Bead technology has already allowed the analysis of approximately 100 different targets simultaneously (Angenendt, 2005). In addition, small nucleic acid molecules called aptamers, that specifically bind proteins, have been developed for protein detection (Angenendt, 2005). Aptamers, due to their stability, binding specificity and well-developed, automated manufacturing process, hold great promise for the development of new classes of protein arrays, even arrays for the combined detection of proteins and nucleic acids.

Approaches to identify and to quantify proteins involved in cancer formation and development are in progress. The initial applications for antibody microarrays within cancer research have been described in 2001 (Knezevic et al., 2001; Sreekumar et al., 2001). Using antibody microarrays, five serum proteins that had significantly different levels in prostate cancer patients compared to normal sera could be identified (Miller et al., 2003). A useful and potentially powerful resource for evaluating gene expression in cancer samples to explore genes and pathways deregulated in a particular cancer type is Oncomine-database (http://www.oncomine.org). Antibody-microarrays are one step towards personalized cancer diagnosis and treatment (see also the next paragraph). For example, a test has recently been developed for the diagnosis of breast cancer based on three biomarkers: urokinase-type plasminogen activator (U-PA), plasminogen activator inhibitor-1 (PAI-1) and vascular endothelial growth factor (VEGF) (Weissenstein et al., 2006). Similar tests, that incorporate larger number of markers are being developed.

Changes in intracellular protein phosphorylation pattern constitute another useful marker for cancer diagnosis. Somatic mutations in protein kinases leading to changes in the protein phosphorylation pattern are prominently involved in tumor progression. Changes in the phosphorylation level, particularly in the status of specific phosphorylation sites, may not easily be accessed by microarray technology due to lack of antibodies that recognize specifically phosphorylated forms of a given protein; hence, combination of methods might be more appropriate. For example, a combination of protein microarrays and two-dimensional separation of proteins has been developed for determination of the tyrosine phosphorylation state in human breast cancer cells (Pal et al., 2006).

6. Proteins released by dying cells as indicators for in vivo-ongoing cell death by cancer therapy and during other pathologic processes

During necrosis, the cell-membranes rupture and the cellular contents are released to the environment in a largely indiscriminate fashion (Johar et al., 2004; Kroemer et al., 1998; Los et al., 2002; Rohn et al., 2005). The apoptotic form of cell death, on the contrary, is a much more orderly and tightly regulated sequence of biochemical events (Brouckaert et al., 2005; Lavrik et al., 2005; Philchenkov et al., 2004).

Apoptosis results in the extracellular release of a limited number of proteins including: cytochrome c (Ahlemeyer et al., 2002; Barczyk et al., 2005; Renz et al., 2001a; Satchell et al., 2005), some caspases (Cowan et al., 2005; Satchell et al., 2005), cleaved cytokeratin-18 and -19 (Sheard et al., 2002) (Table 1), and possibly a few other proteins. The apoptotic cascade can be triggered by a variety of signaling mechanisms, the death-receptor pathway (Los et al., 1995; Muzio et al., 1996) and the mitochondrial pathway (Ceconci et al., 1998; Los et al., 1999) being the most common. Mitochondrial cytochrome c release is the key event, which is critical for the initiation of the formation of the apoptosome complex that is essential for the activation of caspase-9 and initiation of the mitochondria-dependent apoptotic cascade. Below, we focus on cytochrome c and proteolytic fragments of cytokeratin-18 as the most promising ex vivo-indicators of the intensity and type of ongoing cell death.

6.1. Cytochrome c: an apoptosis-specific marker of ongoing in vivo cell death induced by cancer therapy and inflammatory processes

Most anticancer drugs kill target cells by the induction of apoptosis (Barry et al., 1990; Los et al., 1997), and the mitochondrial/apoptosome-dependent death pathway is usually the final, common, signaling cascade activated by cancer therapeutics (Kim, 2005; Los and Gibson, 2005; Makin, 2002; Sharma et al., 2005). Cytochrome c normally resides loosely attached to the outer side of the inner mitochon-
drial membrane and participates in electron transport between complex-III and complex-IV of the mitochondrial respiratory chain. Upon apoptosis induction, cytochrome c is released to the cytoplasm by a yet fully elucidated mechanism and prompts formation of the apoptosome, a complex critical for caspase-9 activation and apoptosis propagation.

Several research groups, including our own, have reported that cytochrome c is released from dying cells (Ahlemeyer et al., 2002; Kawada et al., 2006; Renz et al., 2001b). The extracellular cytochrome c release appears to be specific for apoptosis (Renz et al., 2001a), and its presence can easily be detected in body fluids including serum (Barczyk et al., 2005; Ben-Ari et al., 2003; Kawada et al., 2006; Renz et al., 2001a), and cerebro-spinal fluid (Satchell et al., 2005) (Table 1). In cancer patients, increased levels of cytochrome c in the serum correlate with the onset of in vivo ongoing cell death induced by cancer therapy (Barczyk et al., 2005; Renz et al., 2001a), and preliminary results indicate that it correlates with long-term survival (Barczyk et al., 2005). Some patients displayed a sharp increase in serum cytochrome c level within the first 24 h of anticancer therapy, returning to the basal level (∼5–20 ng/ml) within the next 7–10 days. In a subset of patients, the initial cytochrome c level (prior to the start of cancer therapy) was already very high and sometimes reached a few hundred nanograms per milliliter. In these patients, when therapy started there was no further increase of serum cytochrome c level, but a steep decrease, with levels approaching the control range towards the end of the therapy cycle. The high initial cytochrome c level may be indicative of the overall tumor mass and the incidence of the residual spontaneous apoptosis (tumor cell turn-over) (Barczyk et al., 2005; Renz et al., 2001a).

An increase of the serum cytochrome c level is not unique to cancer therapy. Direct correlation of serum cytochrome c with the onset of in vivo apoptosis triggered during the course of Epstein–Barr virus (EBV)-caused hemophagocytic lymphohistiocytosis (HLH) and infectious mononucleosis, was recently reported (Kawada et al., 2006). Interestingly, in patients with EBV-associated HLH, serum EBV loads and serum cytochrome c were correlated with each other. These results indicate that the apoptosis observed in the course of HLH is dependent on, and proportional to the EBV load (Kawada et al., 2006).

An elevated cytochrome c level has also been reported in patients suffering from fulminant hepatitis. A strong correlation between the clinical symptoms of fulminant liver damage and serum cytochrome c level has been described (Ben-Ari et al., 2003; Ghavami et al., 2005; Sakaida et al., 2005). The serum cytochrome c level also seemed to parallel the severity of hepatic coma (Sakaida et al., 2005). In patients suffering from fulminant hepatitis, the serum cytochrome c level significantly correlated to serum hepatocyte growth factor, aspartate aminotransferase (AST), lactate dehydrogenase (LDH), and alkaline phosphatase (ALP), while it was negatively correlated to serum alpha-fetoprotein (AFP) and total plasma bilirubin (Ben-Ari et al., 2003; Sakaida et al., 2005). These results suggest serum cytochrome c as a new, potent diagnostic and prognostic marker for acute liver failure that is in most cases accompanied by a massive hepatocyte apoptosis (Ghavami et al., 2005; Sakaida et al., 2005).

Two research groups reported apoptosis-triggered release of cytochrome c from neuronal cells: in vitro upon staurosporine-induced cell death (Ahlemeyer et al., 2002) and in vivo after traumatic brain injury (TBI) resulting from child abuse (Satchell et al., 2005). During the in vivo study, cytochrome c concentration was determined by ELISA in 167 ventricular cerebrospinal fluid (CSF) samples from 67 infants and children with TBI. Controls included lumbar CSF from 19 infants and children without trauma or meningitis. Increased CSF cytochrome c was independently associated with abuse-related TBI (p = 0.0001) and female gender (p = 0.04) but not age, Glasgow coma-scale score, or survival (Satchell et al., 2005). Other apoptosis-related proteins including Fas and caspase-1 were also higher in CSF after TBI, but did not independently discriminate between accidental and inflicted TBI. This discovery paves the way for the development of strategies targeting apoptosis after TBI, particularly in female victims of child abuse.

6.2. Cytokeratin-18 and -19 as indicators for cell-death type in vivo

Cell-death type dependent, differential processing of cytokeratin-18 can serve as a novel approach to differentiate between apoptosis and necrosis in vivo. Cytokeratin-18 and -19 are intermediate filaments, which are normally insoluble elements of the cytoskeleton expressed in epithelia and their malignant counterparts. Upon induction of apoptosis, their proteolytic fragments are released from dying cells and can serve as cell-death markers in vivo. In particular, two markers have been frequently used in the clinic for over 10 years: a caspase-generated, 13 kDa proteolytic fragment of human cytokeratin-18, also known as “serum tissue polypeptide-specific antigen” (TPS), and a ∼30 kDa fragment of cytokeratin-19, also called “water-soluble cytokeratin-19 fragment” or CYFRA 21-1 (Romero et al., 1996). Both markers have been traditionally considered as biochemical indicators of tumor burden and cell death within the epithelium, since they do not discriminate between normal and malignant cells.

Interestingly, the comparison of the serum abundance and size of cytokeratin-18 fragments can be used as a marker for the identification of the type of programmed death (Kramer et al., 2004). During necrosis, the cytosolic pool of soluble full-length cytokeratin-18 is released, whereas apoptosis is associated with significant release of caspase-cleaved cytokeratin-18 fragments, mainly the 13 kDa TSP form (Kramer et al., 2004). Thus, the assessments of different forms of cytokeratin-18 in patient sera could be used to examine cell death modes. For example, monitoring of different cytokeratin-18 forms in peripheral blood during chemotherapy of prostate cancer patients showed individ-
ual differences in the patterns of release (Kramer et al., 2004). TPS and CYFRA 21-1 are abundantly released into the extracellular space during the intermediate stage of epitelial cell apoptosis when induced by cancer therapy and other stimuli.

7. Closing remarks

The apoptosis-specific release of cytokeratin proteolytic fragments, cytochrome c, or specific caspasess to the extracellular space after chemotherapy and other stimuli are applied in vitro and in vivo (Ahlemeyer et al., 2002; Cowan et al., 2005; Kramer et al., 2004; Renz et al., 2001a; Satchell et al., 2005). The molecular mechanisms responsible for the trans-membrane translocation are not known in detail. A number of proteins, which lack a signal peptide, are exported by alternative and still poorly-defined pathways. Death ligands for example, such as CD95L and TRAIL, are stored in microvesicles that are released “on demand” upon activation or apoptosis (Albanese et al., 1998; Martinez-Lorenzo et al., 1999). The biological significance of the presence of these molecules in extracellular space, or in plasma, has not been fully elucidated. While future research will focus on a possible immunomodulatory role of cytochrome c (prevention of inflammatory response induction by apoptotic cells), at least in neural tissue cytochrome c release from dying cells contributes to the “bystander effect” (Ahlemeyer et al., 2002). That is, the released cytochrome c re-enters other cells and induces secondary cell death. Furthermore, it has been described recently that cytochrome c may activate K⁺ channels and thus directly contribute to apoptosis-related cell shrinkage (Platoshyn et al., 2002). Nevertheless, the significance of cytochrome c as a mediator of the “bystander effect” still awaits to be resolved.

Serum cytochrome c is a specific indicator of cell death episodes in vivo (Barczyk et al., 2005; Renz et al., 2001a). Similar to lactate dehydrogenase (LDH), a common clinical marker for cell damage, cytochrome c is also released to the extracellular medium, but earlier than LDH and in larger quantities (Renz et al., 2001a). Given the 10-fold difference in their molecular mass, cytochrome c (∼14 kDa) is much more likely to be cleared through the kidneys than LDH (tetramer’s mass: ∼140 kDa). The serum level of cytochrome c therefore changes very dynamically and is directly linked to the real-time ongoing cell death events. Since the two molecules are located in different cellular compartments – cytochrome c in the mitochondrial intermembrane compartment and LDH in the cytoplasm – their release is governed by different cellular processes. The translocation of cytochrome c to the cytoplasm is a prerequisite for the initiation of the apoptosome-dependent apoptotic cascade; therefore, serum cytochrome c level is an indicator of apoptotic rather than necrotic cell death. LDH, on the other hand, is separated from the extracellular space by the single lipid bilayer membrane of the cell. Additionally, the release mechanisms of the two molecules may differ significantly: cytochrome c release is unlikely to occur by simple cell lysis, since LDH release occurs at later time points (Barczyk et al., 2005; Renz et al., 2001a).

Prognostic markers help to predict the outcome of disease, and thus can aid the selection of high-risk patients for more aggressive and/or experimental therapy. In addition to the here presented new concepts for therapy monitoring, several markers exist, that are mostly useful for the therapy evaluation in a single disease or confined to a cluster of malignancies (Diamandis and Yousef, 2002; Kok et al., 2003; Niitsu et al., 2002; Riley et al., 2003; Trieb and Kotz, 2001). Therefore, to cover a broad spectrum of diseases a typical clinical laboratory would need to be furnished with reagents for hundreds of different types of tests. While this is feasible, it would be too costly and labor-intensive, thus often beyond the range of many clinical pathology units. The emerging “lab-on-the-chip” technology will allow for hundreds, if not thousands of assays to be screened with a single sample supplied by the patient for a number of parameters including nucleic acids, proteins, other bio-molecules and their metabolites. As efforts to develop lab-on-the-chip technology are under way, the combination of biochemical apoptotic markers, together with indicators specific for various tissues and/or development stages and the here discussed genetic diagnostics, SNP (or haplotype block) mapping, histological, and NMR tests will one day allow not only precise detection and localization of cell death in vivo, but will facilitate semi-automatization of the diagnostic process and treatment-related decision making (Baldwin et al., 2002; Jain, 2003; Kroczak et al., 2006; Weigl and Hedine, 2002).

Acknowledgements

J.E. Anderson acknowledges operating grant support from the Muscular Dystrophy Association (Project #3968), Parent Project Muscular Dystrophy, the Canadian Institutes for Health Research and the Canadian Space Agency, and support for students and trainees from Manitoba Health Research Council, Natural Sciences and Engineering Research Council, the Manitoba Institute for Child Health, the Paul H.T. Thorlakson Foundation, the Manitoba Medical Services Foundation, the Health Sciences Centre Foundation (Winnipeg); M. Los thankfully acknowledges the support by the Canada Research Chair program, Prostate Cancer Research Foundation of Canada, Manitoba Medical Services Foundation, CancerCare Manitoba Foundation, Canadian Institute of Health Research, and Manitoba Health Research Council.

References


