Mechanisms of Ageing and Development 151 (2015) 114–121

View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Aston Publications Explorer

Mechanisms of Ageing and Development

journal homepage: www.elsevier.com/locate/mechagedev

Novel ageing-biomarker discovery using data-intensive technologies

Griffiths H.R. [a,*], Augustyniak E.M. [a], Bennett S.J. [a], Debacq-Chainiaux F. [b], Dunston C.R. [a], Kristensen P. [c], Melchjorsen C.J. [c], Navarrete Santos A. [d], Simm A. [d], Toussaint O. [b]

[a] Aston Research Centre for Healthy Ageing and Life & Health Sciences, Aston University, Birmingham, B4 7ET, UK
[b] NARILIS-URBC, University of Namur, Namur, Belgium
[c] Aarhus University Department of Engineering Gustav Wieds Vej 10, 8000 Aarhus, Denmark
[d] Martin Luther University of Halle-Wittenberg, Clinic for Cardiac und Thoracic Surgery, Germany

ARTICLE INFO

Article history:
Received 12 December 2014
Received in revised form 1 May 2015
Accepted 28 May 2015
Available online 6 June 2015

Keywords:
Biomarker discovery
Proteomics
miR array
Hyperoxia
Endothelial progenitor cells

ABSTRACT

Ageing is accompanied by many visible characteristics. Other biological and physiological markers are also well-described e.g. loss of circulating sex hormones and increased inflammatory cytokines. Biomarkers for healthy ageing studies are presently predicated on existing knowledge of ageing traits. The increasing availability of data-intensive methods enables deep-analysis of biological samples for novel biomarkers. We have adopted two discrete approaches in MARK-AGE Work Package 7 for biomarker discovery; (1) microarray analyses and/or proteomics in cell systems e.g. endothelial progenitor cells or T cell ageing including a stress model; and (2) investigation of cellular material and plasma directly from tightly-defined proband subsets of different ages using proteomic, transcriptomic and miR array. The first approach provided longitudinal insight into endothelial progenitor and T cell ageing.

This review describes the strategy and use of hypothesis-free, data-intensive approaches to explore cellular proteins, miR, mRNA and plasma proteins as healthy ageing biomarkers, using ageing models and directly within samples from adults of different ages. It considers the challenges associated with integrating multiple models and pilot studies as rational biomarkers for a large cohort study. From this approach, a number of high-throughput methods were developed to evaluate novel, putative biomarkers of ageing in the MARK-AGE cohort.

Crown Copyright © 2015 Published by Elsevier Ireland Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction 115
1.1. Biomarkers of ageing discovery 115
2. Rationale for adopting high data-intensity, hypothesis-generating methods 115
3. Data-intensive methods in ageing biomarker discovery 115
4. MARK-AGE approaches 116
4.1. Proteomic identification of biomarkers in plasma and on the cell surface using two older and younger healthy male comparator subpopulations 116
4.2. Recombinant antibody technology as a tool to explore endothelial ageing biomarkers 117
4.3. Culture of T cells under physiological and atmospheric oxygen conditions 117
4.4. microRNA arrays in biomarker discovery 118
5. Challenges in examining large sample numbers of low volume 119
6. Converting data-intensive methods to robust high throughput analyses – experimental and statistical challenges 119
7. Post-MARK-AGE high throughput technologies – future ageing biomarker discovery tools 119
8. Conclusions 120
Acknowledgments 120
References 120

* Corresponding author.
E-mail address: h.r.griffiths@aston.ac.uk (H.R. Griffiths).

http://dx.doi.org/10.1016/j.mad.2015.05.010
0047-6374/Crown Copyright © 2015 Published by Elsevier Ireland Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Biomarkers of ageing discovery

We have been aware of the normal age-associated changes in human physiology, for example loss of arterial elasticity, since the days of the ancient Greeks; these changes may in some individuals predispose to cardiovascular events for which clinical intervention is essential (Lee and Oh, 2010). This sort of epidemiological evidence, of age-associated characteristics and diseases, has provided the foundation for creating hypotheses about the ageing process (Jacob et al., 2013). For example, protein oxidation has been implicated in increased crosslinking of the extracellular matrix with age and which contributes to skin wrinkling and poor vascular tone, neither of which normally demand an intervention for survival of the ageing organism (Baraibar et al., 2013; Ratnayake et al., 2013). Increased oxidative DNA damage in short-lived animals provides further support for a sub-category of the "damage" or "error theory of ageing" (Sastre et al., 2000; Barja, 2013). Other observations in the physiology of ageing include an increase in the appearance of senescent cells; this observation is one of several used to inform the "programmed theory of ageing" (Kirkwood and Melov, 2011; van Deursen, 2014). From such theories, a number of ageing-biomarkers have been proposed and investigated, many within the MARK-AGE project. Importantly, ageing is a normal process of development and is not inherently a disease requiring treatment. Therefore, by understanding and being able to define normal ageing better, we may also be better placed to predict those people whose health is deviating from normal age-associated change and who are therefore are at increased risk for age-associated disease. Hypothesis-driven approaches have been commonplace for scientists over centuries, however, the advent of high data intensity approaches has fuelled the opportunity to identify novel biomarkers and possibly also develop new hypotheses about ageing.

2. Rationale for adopting high data-intensity, hypothesis-generating methods

Biomedicine, as with all areas of biology, has undergone a rapid change in the last 15 years through the development of technologically advanced methods providing large datasets. We can now question gene expression patterns simultaneously, the profile of regulatory microRNA expression, mRNA and the study of protein forms, their frequency and location or interactions at a particular moment in time, offering unprecedented insight into many aspects of human biology. Developing large scale datasets, e.g. of biological ageing in the case of MARK-AGE, provides opportunities for discovering hitherto unrecognised biomarkers and can also offer increased understanding of the ageing process, generating new hypotheses *in silico*. These biomarkers can then be tested through the associated development of high throughput screening methods such as qPCR, ELISA or flow cytometry for proteins. Here, we describe the approaches, challenges and validation of methods that have been applied in the discovery of novel putative biomarkers of ageing. The data arising from the discovery process has been introduced into the larger MARK-AGE dataset for statistical evaluation and will be published elsewhere. All tissues are used here in biomarker discovery (T cells, circulating endothelial cells, plasma, peripheral blood mononuclear cells) are available from a normally collected human blood sample. It is not anticipated that any new biomarkers of healthy ageing which are discovered would be consistent between different tissues and indeed the differences may provide further insight into healthy ageing.

3. Data-intensive methods in ageing biomarker discovery

The sensitivity of mRNA and miR analysis afforded by their relatively simple amplification together with the immense capability for data mining, has contributed to their popularity in biomarker discovery. Microarrays are typically used to identify new targets in a low or medium throughput scale. After identification of specific mRNA or miRNAs as biomarkers, real-time PCR is used for validation and high throughput application. Microarray platforms enable the comparison of thousands of genes expressed in a cell at any given time which are simultaneously monitored; they have been applied to study the effects of development and ageing on gene expression. However, the variation in gene expression changes with age, probably due to declining transcriptional efficiency during ageing remains a problem (de Magalhaes et al., 2009).

Microarray has been used to study healthy ageing-related changes in gene expression in a several studies since the concept of MARK-AGE was developed (ElSharawy et al., 2012; Nakamura et al., 2012; Lazuardi et al., 2009; Uehara et al., 2006). These studies have largely been based on cross-sectional cohorts. Varying approaches have been adopted previously for the use of microarray in biomarker discovery including a pooling approach, where samples from patients are pooled and compared to controls. This offers advantages in terms of normalizing outliers but correspondingly restricts the data that can be generated from the system (Rudolf et al., 2013). The comparative nature of the analyses also presents a problem when samples are blinded as was the case during MARK-AGE; to overcome this we adopted two different approaches. First, single-blinded analysis was undertaken after the sample codes were broken by the biobank, so enabling leukocytes to be pooled for miR analysis from multiple donors within a given decile of years of age. Second, a subset of donors was recruited to serve as a source of larger cell numbers of known ages for comparative profiling at rest and under hyperoxic stress which is known to induce senescence. This approach allows us to follow longitudinal changes in CD4+ T cells that have been exposed to controlled "ageing".

A range of proteomic approaches are available to investigate the expressed proteome. These include one-dimensional and two-dimensional polyacrylamide gel electrophoresis with options for specific labelling using 2D-DIGE to enable simultaneous sample separation; gel-free high approaches are also adopted such as isotope-coded affinity tag ICAT; SILAC; isobaric tagging for relative and absolute quantitation (iTRAQ); and shotgun proteomics which is proving increasingly popular in recent times owing to the high resolving power for complex samples through innovations in mass spectrometry hardware and methods (Nikolov et al., 2012). MARK-AGE began in 2008 and use of 2DE with LC–MS was commonplace at that time and so was adopted here. The article by Capri et al. (Capri et al., 2015) in this edition, has reviewed the few proteomic ageing biomarker studies to date. These include a differential plasma protein pattern in subjects that associates with amyloid deposition in the brain; a differential expression of ApoE and antioxidant proteins was observed in the plasma of 10 japanese supercentenarians; and our own work has shown age-related changes to the transferrin glycoform (Dunston et al., 2012; Thambisetty et al., 2010; Miura et al., 2011).

Two different proteomic strategies were explored as sources of putative ageing biomarkers in the MARK-AGE project; (1) comparing the proteomes of a small healthy population of younger and older adults, which was carefully controlled for health status; (2) exploring endothelial markers identified during PROTEOMAGE by fishing with antibodies. The workflow is illustrated in Fig. 1 and the approaches used are described in detail in Section 3 below.

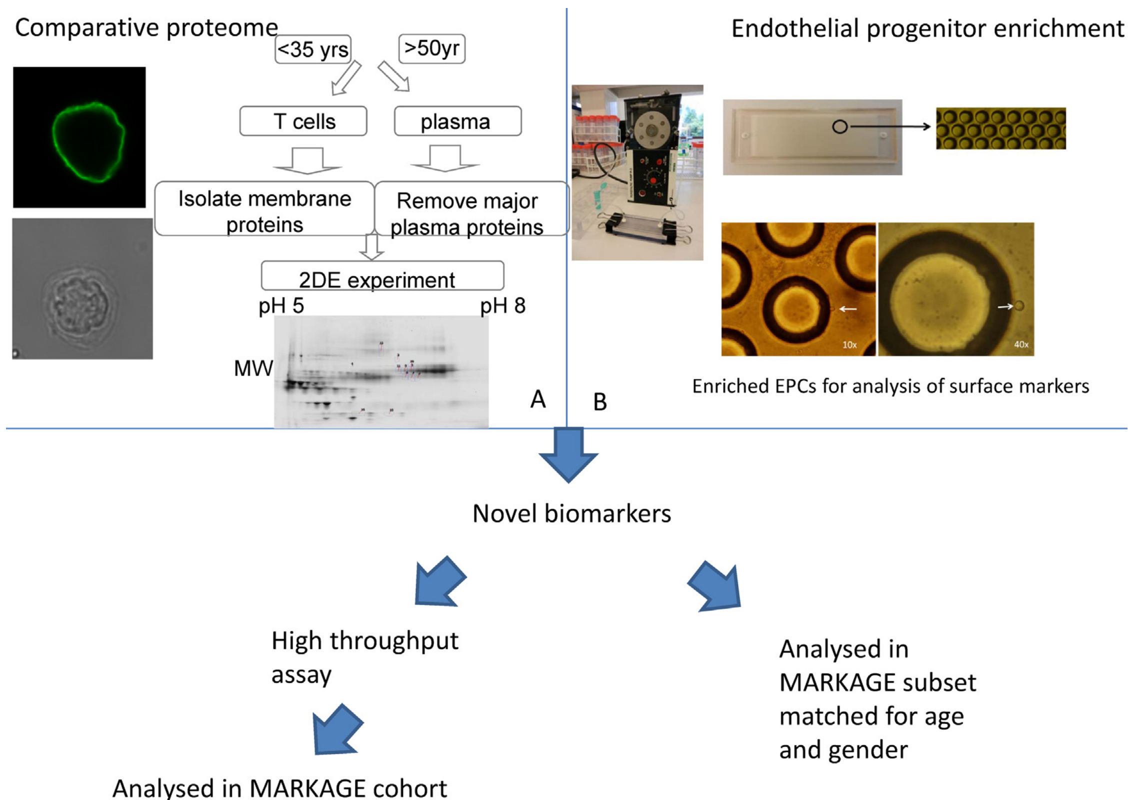

Fig. 1. Strategies to identify protein biomarkers of ageing by high data methods.

4. MARK-AGE approaches

4.1. Proteomic identification of biomarkers in plasma and on the cell surface using two older and younger healthy male comparator subpopulations

Age-related changes can arise in intracellular, cell surface and extracellular proteins. Changes in specific protein concentration, localisation and post-translational modification are commonly reported across a range of physiological systems. For example, incorrect protein secretion patterns by retinal pigment epithelial cells has been associated with the development of age-related macular degeneration possibly due to uncontrolled trafficking (Kay et al., 2013). In the ageing immune system, re-expression of CD45RA on the surface of senescent memory T cell accompanies their reactivation (Di Mitri et al., 2011); and accumulation of an intracellular domain of amyloid precursor protein has been observed in ageing neurones due to alterations in proteolytic processing (Zhang et al., 2012). Alongside, the age-associated changes in protein localisation, the pattern of post-translational protein modifications (PTMs) is also influenced during ageing. Increased ubiquitination of cellular proteins has been described as a result of impaired proteasomal degradation and in tissue, the collagen matrix shows evidence of increased advanced glycation endproduct formation (Scharf et al., 2013). The greatest oxidative changes to proteins are expected in the matrix where protein turnover is low. Indeed non-invasive methods of analysing skin pigment due to advanced glycation end-products are applied to measure ageing (Scharf et al., 2013). The ageing proteome is diverse and offers a rich source of potential biomarkers that may be uncovered through unbiased proteomic analysis using a 2DE and LC–MS approach. 2DE not only identifies proteins that have either increased or decreased in concentration within a specific biological fluid or tissue at a given time but can also identify spots moving in 2D location. This may be due to PTMs that cause an alteration in the pI or mass of a protein isoform. PTMs that confer a charge to a protein are most likely to cause an alteration in spot location on a 2D gel and commonly include phosphorylation (Zhu et al., 2005) and the presence of *N*-acetyl neuraminic acid (NANA also known as sialic acid) residues (Barrabés et al., 2010).

Our initial approach focussed on the plasma proteome from healthy young men (mean age = 21.4 ± 1.5 years) and healthy mid-life men (mean age = 57.0 ± 1.6 years). Their characteristics are shown in Table 1.

We identified twelve spots including transferrin, complement C3b and transthyretin that were quantitatively different between the age groups. Transferrin spots showed an acidic shift in older males (Dunston et al., 2012) which was due to loss of sialic acid.

Table 1
Demographics of subjects who donated plasma samples for comparative 2DE proteomics.

	Young adult(*n* = 10)	Mid-life adult(*n* = 10)	
Mean age	21.4 ± 1.5 years	57.0 ± 1.6 years	*P* = <0.0001
BMI	22.5 ± 2.4	25.4 ± 2.7	NS
LDL (mM)	3.82 ± 0.65	4.75 ± 0.6	NS
HDL (mg/dL)	54.4 ± 35	62.0 ± 3.5	NS
Blood pressure(diastolic) mm/Hg	119 ± 9.7	122 ± 7.5	NS
Blood pressure (systolic) mm/Hg	75 ± 8.2	85 ± 10.3	NS

A high throughput lectin based enzyme linked assay was developed to detect the galactosylation status of transferrin as a putative biomarker of ageing (Dunston et al., 2012) and the significant difference in spot intensity by the 2DE LC–MS approach based on pooled samples, was confirmed in in individual plasma samples by ELLA. Using this ELLA methodology in the MARK-AGE cohort, asialylation was analysed in excess in 2500 plasma samples. Coefficients of variation between assays were <10% and around 50 samples were analysed per day. These data will be reported in relation to other MARK-AGE data.

Compared to the plasma, a more diverse, and therefore potentially richer source of available biomarkers arises within and on the surface of circulating leukocytes. Progressive decline in immune defence appear to contribute to impaired resistance to infection and reduced tumour surveillance over neoplastic lesions in older adults (Castle, 2000; Lang et al., 2013). However, crude proteomic analysis of total leukocyte populations for age-related biomarkers is likely to be confounded by age-related changes in specific subset frequency. Therefore, to reduce this risk and to retain focus on cells which decline functionally with age, CD4+ T cells were selected for further analysis. Previous reports have suggested a shift in the cytokine profile secreted from T cells during ageing so contributing to changes in the circulating plasma protein pool, but membrane proteins have not been investigated in detail. Taking forward this putative, novel source of ageing biomarkers, membrane markers in circulating cells are attractive targets to identify since they can be easily analysed by high throughput approaches such as flow cytometry and require low sample volumes.

To identify reliable immune cell aging markers in an hypothesis-free approach, in Work Package 7 we adopted a membrane enrichment approach of the $CD4^+$ T cell membrane proteome (purified by cell surface labelling using Sulfo-NHS-SS-Biotin reagent) from young ($n = 9$, 20–25y) and mid-life male ($n = 10$; 50–70y) groups. This approach has been detailed and successfully applied elsewhere (Carilho Torrao et al., 2013). The requirement to isolate T cells from subset of healthy subjects was because of the low number of cells that could be isolated from MARK-AGE probands donating to the biobank; it was not possible to achieve proteomic analysis based on 2D-DIGE and mass spectrometry following isolation of CD4+ cells from MARK-AGE donors. However, collection of blood samples from an independent subset of volunteers, followed by 2DE separation yielded several membrane protein spots with age-dependent density differences ($p < 0.05$ and >1.4 fold change) which were identified by LC–MS/MS.

A flow cytometric protocol was established for determining expression of target proteins identified by proteomics on the surface of CD4+ T cells. Colour compensation was applied was robust with less than 0.5% drift between days. Biological controls, were included in each batch to check reproducibility of the analysis. The between and within batch CV of our analysis based on the analysis of >10 biological control samples was 18% and 8% for the two markers selected for study and this is suitable for a clinical chemistry analytical method which would typically show inter- and intra-batch CV of less than 20% and 10% respectively. Although we found some difficulties in assessing about 10–15% of the samples, due to debris, we were able to obtain clear and distinctive peaks for CD4+ population for 85–90% of the samples. These data will be reported in relation to other MARK-AGE data.

4.2. *Recombinant antibody technology as a tool to explore endothelial ageing biomarkers*

A functional vascular system is vital for all organs of our body; however a full understanding is hampered by the heterogeneous nature of the system. Within our human body cellular complexity is a general theme in seemingly identical cells. Individual cells may respond differently to external stimuli, thereby giving rise to functional distinct cells, examples of this being the various subpopulation of endothelial progenitor cells (EPCs) which circulate in the blood. This population has been shown to home to particular sites of vascular injury or insult, depending on the age and functionality of the EPCs and the external stimuli, the EPCs will secrete different factors which may aid in e.g. vascular regeneration (Estes et al., 2010; Denizot and Nathan, 2010). Another example is the pericyte, which has been shown to regulate blood flow in the microvasular system of e.g. the brain (Jespersen and Ostergaard, 2012). Although less studied, if a few of the pericytes surrounding the vessels start to be senescent, this could have marked effect on local mechanotransduction and hence oxygenation of tissue. Based on the above it is clear that single cell analysis and high throughput technologies are needed to gain further insight into the complexity (Sørensen and Kristensen, 2011). Today many of the technologies used, e.g. in the proteomics field, suffer from a sensitivity problem demanding that millions of cells to be analysed together. This will result in an average picture is obtained of the functional states. Such an approach will miss out on the influence of the single or few ageing cells, which may exert global changes on the body.

Allowing single cell proteomic analysis, a novel methodology was developed, based on recombinant antibody technology, which allow identification of biomarkers on single cells in a heterogeneous population (Sørensen and Kristensen, 2011). Antibody phage display relies on genetic engineering of antibodies in bacteriophages. Subsequent ligation of the variable heavy and light PCR products into a phage display vectors is followed by repeated rounds of antigen-guided selection and phage propagation. Monoclonal antibodies of virtually any specificity can be selected for and produced semi-synthetically. By performing selection using large diversity antibody libraries (Mandrup et al., 2013) (more than 10^8 different antibodies) were presented as genetic packages on the surface of filamentous bacteriophage. Using this approach, it was possible to isolate antibodies which bind sub-populations of cells, such as senescent cells. The technology has been optimised, allowing selection of antibodies binding specifically to one identified cell in a heterogenous population (Sørensen and Kristensen, 2011; Mandrup et al., 2013). Following identification of the cognate antigens recognised by the isolated antibodies, it has been possible to test if the identified antigens qualify as biomarkers of ageing, as we have done in the MARK-AGE project. Within MARK-AGE we have in particular focused on an antigen, which was identified in a FP6 EU supported project Proteomage and which were identified by selection of recombinant antibodies on human endothelial cells allowed to age in culture (Boisen et al., 2010; Boisen and Kristensen, 2010).

The analysis of plasma for proteins which have been secreted by endothelial cells as biomarkers of ageing was undertaken using 4 antibodies isolated using Phage Display technology for application in ELISA. In addition a panel of antibodies recognising T cells was being analysed for their potential to discriminate ageing T cells using flow cytometric analysis.

4.3. *Culture of T cells under physiological and atmospheric oxygen conditions*

Atmospheric oxygen pressure generates oxidant stresses on human and murine cells cultivated in vitro and non-exposed in vivo to the atmosphere, altering the expression of genes (Zglinicki et al., 2003; Parrinello et al., 2003) and causes premature senescence (Toussaint et al., 2004). Cell culture under atmospheric oxygen concentration must then be assessed in respect of the cellular conditions studied, especially in studies on ageing in which ROS strongly influence the process. T cells were shown to be particularly sensitive to oxygen concentration (Castle, 2000). If oxygen pressure in bloodstream is evaluated at 5% (Larbi et al., 2007), it is estimated

Fig. 2. Low oxygen cell culture platform.

that the mean oxygen level experienced by T cells varies from 1% to 10% according to their localization (Larbi et al., 2010). Culturing T cells at physiological oxygen concentration (5%) generates cellular responses closer to those detected in vivo (Atkuri et al., 2007). In order to detect new biomarkers of T cell ageing, we studied gene expression changes following activation of T cells cultured under atmospheric or physiological oxygen pressure.

To simulate physiological oxygen tension, T cells were cultured in a low oxygen culture platform developed within the European project PROTEOMAGE. This culture platform consists of three interrelated parts: a culture hood open to the outside (E1), a closed culture hood allowing to manipulate cells in a controlled oxygen concentration and equipped with a robot (E2) and an automated cell incubator maintained at a controlled oxygen concentration (E3) (Fig. 2). Therefore, any cell introduced in this platform will stay at a defined oxygen concentration during the whole process, with no oxygen variation between handling steps (medium change, splitting) and incubation. PBMC cells were isolated from 10 young volunteers (35–39 years) and 10 aged volunteers (70–74 years) recruited in WORK PACKAGE 1. After isolation, cells were cultivated at 5% oxygen pressure (physiological conditions) or at 21% oxygen pressure (atmospheric conditions). T cells were then activated with CD3/CD28 dynabeads® or not, and cultivated at 5% oxygen pressure or at 21% oxygen pressure for 72 h. Total RNA was then isolated, retro-transcribed and PCR using specific TaqMan Low Density Array (TLDA-Applied Biosystems) were achieved. Cross-comparisons of gene expression change between the younger vs older groups and between cells grown at 5% vs 21% oxygen revealed significant differences that distinguished changes associated with age, with oxygen pressure or both. Specific on-demand TLDA were designed to allow analysis of the expression of 48 genes in total including 3 housekeeping genes (18s, B2 M and RPLP0) and 45 genes specifically related to the activation of T lymphocytes. The three housekeeping genes were chosen as robust after using the Human Endogenous Control Array (Applied Biosystems) that compared the expression of 16 housekeeping genes. The genes selected appeared stable and by referencing to the multiple internal control genes with reliable quantification of gene expression changes (Vandesompele et al., 2002). The further analysis of these selected genes showed that their expression was modulated by activation, the donor age and/or the oxygen concentration.

4.4. *microRNA arrays in biomarker discovery*

MicroRNAs are small non-coding RNA molecules, 21 to 25-nucleotides in length, mainly involved in post-transcriptional gene regulation (translational repression, mRNA cleavage, deadenylation). Since the discovery of MicroRNAs (miRNAs) in C. elegans in 1993 (Lee et al., 1993), many efforts have been taken to further elucidate the cellular function of new disclosed miRNAs. The actual number of microRNA loci annotated in miRBase has risen from 15,172 loci in 142 species (release 16, October 2010) to 24,521 loci in 206 species (release 20, June 2013); (Kozomara and Griffiths-Jones, 2014). The number of *Homo sapiens* sequences listed in the miRNAs database is 1872 (June 2014; www.mirbase.org). In 2008, Mitchell et al. published, for the first time, the use of miRNA-141 as biomarker distinguishing patients with prostate cancer from healthy controls in serum or plasma. The results established the measurement of tumor derived miRNAs in serum or plasma as an important approach for the blood-based detection of human cancer (Mitchell et al., 2008). In addition, circulating miRNAs can be taken up by recipient cells, modifying the cellular behaviour by miRNA induced silencing of target mRNAs. Actually, a huge number of miRNAs have been reported to be specific biomarkers in a variety of diseases such as cancer, neurodegeneration, diabetes and myocardial disease. In a recently published review, Chevillet et al. summarized miRNA biomarkers for neurodegeneration, cardiovascular and metabolic diseases (Chevillet et al., 2014). The specimens used for the isolation of miRNAs included peripheral blood mononuclear cells (PBMCs), cerebrospinal fluid, serum, plasma, whole blood and urine, which reflect the wide incidence of miRNAs in body fluids. Regarding miRNA profiling and their use as biomarkers of ageing in the old *"healthy"* population, few studies are published. The Grillari group showed that miR-17, miR-19b, miR-20a, and miR-106a are down-regulated in human aging (Hackl et al., 2010). In a study by ElSharawy et al. using a microarray with

Table 2
PBMC samples analysed for miR expression.

Group	Explanation	# of samples	Age	Gender
GO	GEHA offspring of long living parents	31	68.4 ± 6.2	16 F/15 M
SGO	Siblings of the GEHA offspring	19	61.7 ± 8.5	9 F/19 M
RASIG	Normal population	140	53.8 ± 13.8	73 F/67 M

863 miRNAs, the expression profiles obtained from blood samples from 15 centenarian and nonagenarian (mean of age 96.4 years) were compared with those of 55 younger individuals (mean age 49.5 years). Eighty miRNAs showed aging-associated expression changes (ElSharawy et al., 2012). In another study, in which eleven healthy individuals aged 20, 80 and 100 years underwent miRNA plasma profiling, miR-21 was identified as a possible biomarker of inflammation and thus ageing (inflammaging) (Olivieri et al., 2012).

In the context of the MARK-AGE subset, the primary approach for the establishment of miRNAS as biomarkers of ageing consisted in the isolation of total RNA from CD45+ PBMCs. Total RNA from 190 subjects was analysed in the age range from 30 to 74 years using the miRNA2.0 array system (Affymetrix). The subjects' characteristics are shown in Table 2. The false discovery cut-off was set to <5%.

We demonstrated age- as well as group-specific miRNA healthy ageing biomarkers, which are under evaluation as new ageing biomarkers.

5. Challenges in examining large sample numbers of low volume

Searching for biomarkers with small volumes can work well when high fidelity amplification of biological material is feasible e.g. for DNA, mRNA and miRNA. However, high throughput array technology is comparative and requires the establishment of defined groups which is not consistent with the double blind nature of the study. To address this, at the biobank, samples were decoded, grouped by age and recoded prior to dispatch to the analysis sites at Halle University so that the analytical sites were provided with a subset of PBMC samples from probands of specific age ranges.

For protein analyses where amplification of biological material was not possible i.e. for protein biomarker discovery, the major challenge was in recovering specific protein from any of the given samples. This was further complicated in cellular analyses by the requirement to study specific T cell subsets and therefore negate the likely contribution of age-related changes in major subset frequency to any novel biomarker – i.e. to avoid a series of complex analyses which only determined that the ratio of CD4:CD8 was changed with age. To overcome this, biomarker identification in the membrane proteomic discovery platform (Aston) was predicated on collection of samples from local cohorts in the first instance which enabled recovery of 5–10 times the number of CD4+ T cells compared to samples available in the biobank. To obtain sufficient protein for putative biomarker discovery in CD4+ membrane proteins required a minimum of T cell extraction from 40 ml blood and 10 volunteers per group, prior to validation through the MARK-AGE biobank.

PBMC quality and number were critical success factors for cell culture and miR isolation from PBMC. Analysis was not undertaken when cell viability was low in case the preparation, storage and handling of cells had contributed to alterations in miR expression.

To validate potential biomarkers resulting from proteomic analysis high throughput technology, such as ELISA and flow cytometry, is needed in order to quantitate the antigens. Even though the level of a given antigen may not differ in different age groups, the localisation within cellular compartments or the extracellular space may vary, thus requiring image analysis. In all of the above specific antibodies are needed which bind to specific epitopes in the native conformation. Even today this remains a challenge, as commercial available antibodies may not always exhibit the specificity claimed by the manufacture. Therefore one of the most crucial steps in developing assays for validation of biomarkers remain good and proper validation of the antibodies applied.

6. Converting data-intensive methods to robust high throughput analyses – experimental and statistical challenges

Before the big-data handling approaches that are available for protein-based antigen discovery through mass spectrometry that are available today, during the marker-development phase of MARK-AGE samples were pooled from multiple donors to reduce the data complexity in plasma 2DE proteomics. The donors were carefully selected from well-characterised groups with defined exclusion criteria as previously reported (Dunston et al., 2012). The two populations differed only in age. An advantage of this approach was the assumption that the differences due to age would be greater that the inter-individual variances during within each age population. However, there are drawbacks. First, the populations themselves do not represent the "typical" ageing seen in the general population. This may result in an atypically, healthy-aged population as the comparator group against younger healthy subjects. However, this approach does eliminate confounding effects of drugs and disease on biomarker differences seen during healthy ageing. Second, while the demographics within each of the populations were similar, there may still be underlying major genetic and environmental factors that affect protein expression or post-translational modification in a minority of the population. For these reasons, further validation of markers identified by 2DE was undertaken by ELISA.

The same approach, using samples pooled from well-defined, specific-aged healthy populations, was adopted in the discovery of cell surface marker changes. As ageing associates with differences in immune cell subsets, only membranes from CD4+ T cells were explored to reduce potential bias from over-representation of CD4+ compared to CD8+ cell membranes from older adults. After identification of novel cell surface marker antigens via a pooled-membrane approach, the individual whole blood, optilyse-stored samples were re-analysed to confirm that any CD4+ cell surface marker differences observed were not generated artefactually during sample processing.

An essential validation step for newly discovered biomarkers is to apply them within a much larger cohort of samples, however, the greater the number of potential markers considered, the greater the risk for false-positive discovery. An independent Work Package was established in MARK-AGE, Work Package 8: data analysis and bioinformatics that was based on using neural networks, correlation measures and boot-strapping tools to extract a robust set of biomarkers of human ageing and to derive a model for healthy ageing.

7. Post-MARK-AGE high throughput technologies – future ageing biomarker discovery tools

Robust non-gel based proteomics (commonly referred to as 'shotgun proteomics') offer certain benefits over 2-DE and have been used more widely in ageing biomarker discovery in vitro, in animal models and most recently in human tissue (Dunston and Griffiths, 2010; Won et al., 2012; Shevchenko et al., 2012). The entire complex protein mixture in samples is enzymatically digested and followed by separation with liquid chromatography and direct infusion into a mass spectrometer (MS). Quantitative differences can be uncovered using a mass spectrometer capable of MS/MS fragmentation. For the majority of known specifically modified proteins, second dimension MS fragments the parent ion, yielding one or more unique fragments for quantitation can be applied if a known post-translation modification is being sought. By zooming in on one or more specified parent ion(s) and selectively monitoring unique fragment ion transition(s), single or multiple

reaction monitoring (SRM or MRM) methods are established. SRM and MRM offer quantitative label-free MS approaches when MS/MS is available. Future opportunities lie using the most advanced triple MS in hyper-reaction monitoring and SWATH™ for determining digital fingerprints based on post-acquisition SRM data mining, however, this methodology is very much in its infancy. Label-free techniques have advanced and with improved mass accuracy do not require sample clean up to remove unreacted reporter tags, so do not suffer loss of material in the sample preparation steps.

Isobaric tagging remains a popular approach to compare different samples with peptides that have identical masses and chemical properties. Heavy and light "isotopologues" co-elute together and are then cleaved during MS/MS. These isobaric labels have the benefit that they fundamentally do not interfere with peptide ionisation, but the conjugation processes do require significant sample clean-up which is often inefficient, this can be a significant issue when dealing with valuable material such as small volumes of patient samples. A benefit of iCAT and TMT tags is the option for a purification step by biotinylation of iCAT or antibodies to TMT which enables purification using a streptavidin matrix. This extra step removes any unlabelled peptides which can reduce the presence of contaminants in the sample which can interfere with peptide ionisation.

Analysis of nucleotides has been aided over the last 20 years by the ease with which mRNA could be reverse-transcribed and DNA amplified to facilitate detection. Next generation sequencing (NGS) technologies have now accelerated to such an extent since the first descriptions of their use 10 years ago, that they are being considered in the development of clinical diagnostics and therapeutics (Xuan et al., 2013). One principle driver for this is the wealth of data created and reduced costs; hundreds of gigabases of nucleotide sequence per analysis is delivered with a reduction of sequencing cost by over five orders of magnitude since the approach was first developed. The promise of this technology lies in its ability to directly sequence single DNA or RNA molecules in biological samples without amplification, so it therefore minimizes sample handling, reduces sample manipulation, avoids amplification-induced bias, increases read length flexibility resulting in more accurate quantitation of nucleic acid molecules including alternative splice variants.

Nevertheless, targeted sequencing can yield much higher coverage of genomic regions of interest while reducing the sequencing cost and time following target-enrichment techniques such as PCR with array-based or solution-based hybridization. Indeed, the systematic analysis of miR, single nucleotide polymorphism and deep sequencing approaches has led to further insight into expression changes e.g. in the CNS (van der Brug et al., 2010).

A final emerging area of promise in the discovery of ageing biomarkers is the field of metabolomics. One of the early groups to work in this field has suggested the resistance of ageing muscle to anabolic stimuli which may underpin the loss of muscle strength with age (Collino et al., 2014). Evidence is also emerging for relationship between immune ageing and cell metabolism (Torrao et al., 2014). As both the technology and data handling capabilities improve over the next few years, the importance of metabolism to healthy ageing will be better characterised.

8. Conclusions

The overall objectives of this work package were to adopt new, robust, high throughput technologies (gene expression analysis, miRNA analysis and proteomics), to identify biomarkers from tissues, cells and/or serum which define the biological age of humans. Using high data-intensity, hypothesis-generating methods we have provided additional biomarkers that may prove use in fine-tuning ageing biomarkers either alone or in combination with other markers or for identifying specific age-associated diseases.

Acknowledgments

The authors acknowledge financial support from the European Union through the MARK-AGE project "MARK-AGE (EU FP7 Large-scale integrating Project HEALTH-F4-2008-200880)". O. Toussaint and F. Debacq-Chainiaux are respectively Senior Research Associate and Research Associate of the FNRS, Belgium.

References

Atkuri, K.R., et al., 2007. Importance of culturing primary lymphocytes at physiological oxygen levels. Proc. Natl. Acad. Sci. U S A 104 (11), 4547–4552.
Baraibar, M.A., Ladouce, R., Friguet, B., 2013. Proteomic quantification and identification of carbonylated proteins upon oxidative stress and during cellular aging. J. Proteomics 92, 63–70.
Barja, G., 2013. Updating the mitochondrial free radical theory of aging: an integrated view, key aspects, and confounding concepts. Antioxid. Redox Signal. 19 (12), 1420–1445.
Barrabés, S., et al., 2010. Effect of sialic acid content on glycoprotein pI analyzed by two-dimensional electrophoresis. Electrophoresis 31 (17), 2903–2912.
Boisen, L., Kristensen, P., 2010. Confronting cellular heterogeneity in studies of protein metabolism and homeostasis in aging research. Adv. Exp. Med. Biol. 694, 234–244.
Boisen, L., et al., 2010. Evaluation of endothelial cell culture as a model system of vascular ageing. Exp. Gerontol. 45 (10), 779–787.
Capri, M., et al., 2015. MARK-AGE population: from the human model to new insights. Mech. Ageing Dev.
Carilho Torrao, R.B., et al., 2013. Healthy ageing and depletion of intracellular glutathione influences T cell membrane thioredoxin-1 levels and cytokine secretion. Chem. Cent. J. 7 (1), 150.
Castle, S.C., 2000. Clinical relevance of age-related immune dysfunction. Clin. Infect. Dis. 31 (2), 578–585.
Chevillet, J.R., et al., 2014. Issues and prospects of microRNA-based biomarkers in blood and other body fluids. Molecules 19 (5), 6080–6105.
Collino, S., et al., 2014. Reprint of: musculoskeletal system in the old age and the demand for healthy ageing biomarkers. Mech Ageing Dev 136–137, 94–100.
de Magalhaes, J.P., Curado, J., Church, G.M., 2009. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics 25 (7), 875–881.
Denizot, Y., Nathan, N., 2010. Age-dependent endothelial progenitor cell mobilization and cardiac surgery. Cytokine 51 (3), 249, author reply 250.
Di Mitri, D., et al., 2011. Reversible senescence in human CD4+CD45RA+CD27 – memory T cells. J. Immunol. 187 (5), 2093–2100.
Dunston, C.R., Griffiths, H.R., 2010. The effect of ageing on macrophage Toll-like receptor-mediated responses in the fight against pathogens. Clin. Exp. Immunol. 161 (3), 407–416.
Dunston, C.R., Choudhury, K., Griffiths, H.R., 2012. Terminal galactose residues on transferrin are increased in midlife adults compared to young adults. Proteomics 12 (21), 3147–3153.
ElSharawy, A., et al., 2012. Genome-wide miRNA signatures of human longevity. Aging Cell 11 (4), 607–616.
Estes, M.L., et al., 2010. Identification of endothelial cells and progenitor cell subsets in human peripheral blood. Curr. Protoc. Cytom. 33, 1–11, Chapter 9: Unit 9.
Hackl, M., et al., 2010. miR-17, miR-19b, miR-20a, and miR-10 are down-regulated in human aging. Aging Cell 9 (2), 291–296.
Jacob, K.D., et al., 2013. Markers of oxidant stress that are clinically relevant in aging and age-related disease. Mech. Ageing Dev. 134 (3–4), 139–157.
Jespersen, S.N., Ostergaard, L., 2012. The roles of cerebral blood flow, capillary transit time heterogeneity, and oxygen tension in brain oxygenation and metabolism. J. Cereb. Blood Flow Metab. 32 (2), 264–277.
Kay, P., Yang, Y.C., Paraoan, L., 2013. Directional protein secretion by the retinal pigment epithelium: roles in retinal health and the development of age-related macular degeneration. J Cell Mol Med 17 (7), 833–843.
Kirkwood, T.B., Melov, S., 2011. On the programmed/non-programmed nature of ageing within the life history. Curr. Biol. 21 (18), R701–R707.
Kozomara, A., Griffiths-Jones, S., 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42 (Database issue), D68–D73.
Lang, P., Govind, S., Aspinall, R., 2013. Reversing T cell immunosenescence: why, who, and how. AGE 35 (June (3)), 609–620.
Larbi, A., Kempf, J., Pawelec, G., 2007. Oxidative stress modulation and T cell activation. Exp. Gerontol. 42 (9), 852–858.
Larbi, A., et al., 2010. Reduced oxygen tension results in reduced human T cell proliferation and increased intracellular oxidative damage and susceptibility to apoptosis upon activation. Free Radic. Biol. Med. 48 (1), 26–34.
Lazuardi, L., et al., 2009. Microarray analysis reveals similarity between CD8+CD28- T cells from young and elderly persons, but not of CD8+CD28+ T cells. Biogerontology 10 (2), 191–202.

Lee, H.Y., Oh, B.H., 2010. Aging and arterial stiffness. Circ. J. 74 (11), 2257–2262.
Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75 (5), 843–854.
Mandrup, O.A., et al., 2013. A novel heavy domain antibody library with functionally optimized complementarity determining regions. PLoS One 8 (October (10)), e76834, Accepted.
Mitchell, P.S., et al., 2008. Circulating microRNAs as stable blood-based markers for cancer detection. Proc. Natl. Acad. Sci. U S A 105 (30), 10513–10518.
Miura, Y., et al., 2011. Proteomic analysis of plasma proteins in Japanese semisuper centenarians. Exp. Gerontol. 46 (1), 81–85.
Nakamura, S., et al., 2012. Identification of blood biomarkers of aging by transcript profiling of whole blood. Biochem. Biophys. Res. Commun. 418 (2), 313–318.
Nikolov, M., Schmidt, C., Urlaub, H., 2012. Quantitative mass spectrometry-based proteomics: an overview. Methods Mol. Biol. 893, 85–100.
Olivieri, F., et al., 2012. Age-related differences in the expression of circulating microRNAs: miR-21 as a new circulating marker of inflammaging. Mech. Ageing Dev. 133 (11–12), 675–685.
Parrinello, S., et al., 2003. Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts. Nat. Cell Biol. 5 (8), 741–747.
Ratnayake, S., et al., 2013. Stabilising cysteinyl thiol oxidation and nitrosation for proteomic analysis. J. Proteomics 92, 160–170.
Rudolf, H., Pricop-Jeckstadt, M., Reinsch, N., 2013. Flexible pooling in gene expression profiles: design and statistical modeling of experiments for unbiased contrasts. Stat. Appl. Genet. Mol. Biol. 12 (1), 71–86.
Sørensen, M.D., Kristensen, P., 2011. Selection of antibodies against a single rare cell present in a heterogeneous population using phage display. Nat. Protoc., accepted.
Sastre, J., Pallardo, F.V., Vina, J., 2000. Mitochondrial oxidative stress plays a key role in aging and apoptosis. IUBMB Life 49 (5), 427–435.
Scharf, B., et al., 2013. Age-related carbonylation of fibrocartilage structural proteins drives tissue degenerative modification. Chem. Biol. 20 (7), 922–934.
Shevchenko, G., et al., 2012. Longitudinal characterization of the brain proteomes for the tg2576 amyloid mouse model using shotgun based mass spectrometry. J. Proteome Res. 11 (12), 6159–6174.
Thambisetty, M., et al., 2010. Proteome-based plasma markers of brain amyloid-beta deposition in non-demented older individuals. J. Alzheimers Dis. 22 (4), 1099–1109.
Torrao, R.C., et al., 2014. Does metabolic reprogramming underpin age-associated changes in T cell phenotype and function. Free Radic. Biol. Med. 71C, 26–35.
Toussaint, O., et al., 2004. Role of subcytotoxic stress in tissue ageing. In: In Model Systems in Aging. Springer, Berlin Heidelberg, pp. 269–294.
Uehara, T., et al., 2006. S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. Nature 441 (7092), 513–517.
Vandesompele, J., et al., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3 (7), 0034.1–0034.11.
van Deursen, J.M., 2014. The role of senescent cells in ageing. Nature 509 (7501), 439–446.
van der Brug, M., Nalls, M.A., Cookson, M.R., 2010. Deep sequencing of coding and non-coding RNA in the CNS. Brain Res. 1338, 146–154.
Won, C.H., et al., 2012. Comparative secretome analysis of human follicular dermal papilla cells and fibroblasts using shotgun proteomics. BMB Rep. 45 (4), 253–258.
Xuan, J., et al., 2013. Next-generation sequencing in the clinic: promises and challenges. Cancer Lett. 340 (2), 284–295.
Zglinicki, T.v., Petrie, J., Kirkwood, T.B.L., 2003. Telomere-driven replicative senescence is a stress response. Nat. Biotech. 21 (3), 229–230.
Zhang, H., et al., 2012. Proteolytic processing of Alzheimer's beta-amyloid precursor protein. J. Neurochem. 120 (Suppl. 1), 9–21.
Zhu, K., et al., 2005. Protein pI shifts due to post translational modifications in the separation and characterization of proteins. Anal. Chem. 77 (9), 2745–2755.