Setting up the LC-MS/MS approach for plasma proteomics in the DM4You Project

Eduardo da Costa Camilo

Thesis to obtain the Master of Science Degree in

Microbiology

Supervisors: Prof. Carla Maria Alexandre Pinheiro and Prof. Miguel Nobre Parreira Cacho Teixeira

Examination Committee

Chairperson: Prof. Jorge Humberto Gomes Leitão

Supervisor: Prof. Carla Maria Alexandre Pinheiro

Members of the Committee: Prof. Ana Luísa Almaça da Cruz Fernando

October 2024

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

The work presented in this thesis was performed at the Unidade de Ciências Biomoleculares Aplicadas, Portugal, University Nova of Lisbon (Lisbon, Portugal), during the period September-July 2024, under the supervision of Prof. Carla Maria Alexandre Pinheiro and Prof. Rita Gonçalves Sobral de Almeida. The thesis was co-supervised at Instituto Superior Técnico by Prof. Prof. Miguel Nobre Parreira Cacho Teixeira.

Acknowledgments

This project could not have been possible without the collaboration and contribution of extremely important people who, whenever I needed them, helped me unconditionally. Their support is sincerely gratefully acknowledged, and I could not thank each one of them enough:

Professor **Carla Pinheiro**, for the opportunity, for all the guidance, advice and dedication throughout this long journey. Thanks to her help, it was possible to develop and complete this work, always managing to find a solution to every adversity. A huge thank you.

Professor **Rita Sobral**, for her help, collaboration and availability during this work. And for supervising this work as well.

Professor **João Almeida**, for her help and collaboration with the analysis of the results.

To the entire **MolMicro of Bacterial Pathogens Lab**, a sincere thank you for your availability and friendliness. To **Fátima Cardoso**, without our partnership, the completion of this work would not have been possible.

To the entire **DM4You** team, for the help provided throughout this work and for the welcome in this team. Also, to "**Recovery and Resilience Plan"** for funding this work. (PRR-C05-i03-I-000152-LA1.3; PRR-C05-i03-I-000152-LA1.4) To all our **volunteers** in this work.

Professor **Maria da Graça Pacheco**, for her kindness throughout this work and for the opportunity she provided me through the **Research Scholarship (BI) in the fields of Agricultural Sciences, Nutrition Sciences, Food Studies, or related areas** (Mediterranean Diet: promoting and valuing the Mediterranean Diet).

To the **Luxemburg Institute of Science and Technology**, especially PhD. **Céline Leclercq**, for your help and availability during the proteomic analysis.

To my **closest friends**, for always believing in me and never letting me give up, despite all the difficulties encountered along the way. I can never thank everyone enough.

To my girlfriend, **Catarina Baptista**, for helping me in every step of this work. For your dedication and especially patience throughout this work. I love you my "tonta".

To my **family**, I would like to highlight my parents (**Rute Costa** and **Flávio Camilo**), sister (**Sofia Camilo**) and grandmother (**Domicília Costa**), because without their support and patience, this work would have been impossible to carry out. And finally, I want to dedicate this work to my grandfather, **Victor Costa**. Unfortunately, you will not be able to read this thesis. But I know that on the day of the discussion you will somehow be in the front row.

To **NOVA School of Science and Technology, Polytechnic Institute of Portalegre** and **Institute Técnico of Lisbon**

Abstract

An effective way to assess the nutraceutical impact of diet on health is through the analysis of biomarkers, such as plasma proteins, which are widely used as indicators of various diseases. However, many human biomarkers remain uncharacterised, and the dynamic range of protein abundance in plasma is poorly understood. This study aimed to test procedures to be employed within the DM4You project to minimise post-collection alterations. The inclusion of phosphatase and protease inhibitors, along with standardisation of the time between extraction and centrifugation, demonstrated a one-third preservation in protein concentration, as shown by the Bradford and Kalb and Bernlohr quantification methods. Regarding the reducing potential of the samples, two protocols (Folin-Ciocalteu and TEAC assay) were used, optimisation was achieved despite different results obtained in each assay. For proteomics analysis three preparation protocols were tested to demonstrate the impact of fractionation and enrichment. The cost-time relationship should be considered in future studies, particularly with 80 participants and multiple sample collections. This work provides data for the selection of an experimental protocol for future studies.

Keywords: Mediterranean diet, Human plasma proteome, Protein biomarkers, Plasma Protein Analysis, Post-collection alteration, Experimental protocol, Sample preparation.

Resumo

Uma forma eficaz de avaliar o impacto nutracêutico da dieta na saúde é através da análise de biomarcadores, como as proteínas plasmáticas, amplamente utilizadas como indicadores de várias doenças. No entanto, muitos biomarcadores humanos permanecem não caracterizados, e a gama dinâmica de abundância de proteínas no plasma é pouco compreendida. Este estudo teve como objetivo testar procedimentos a serem utilizados no âmbito do projeto DM4You para minimizar alterações pós-colheita. A inclusão de inibidores de fosfatases e proteases, juntamente com a padronização do tempo entre a extração e a centrifugação, demonstrou uma preservação de um terço na concentração proteica, conforme evidenciado pelos métodos de quantificação de Bradford e Kalb e Bernlohr. Em relação ao potencial redutor das amostras, foram utilizados dois protocolos (Folin-Ciocalteu e ensaio TEAC), tendo sido alcançada otimização apesar dos diferentes resultados obtidos em cada ensaio. Para a análise proteómica, foram testados três protocolos de preparação para demonstrar o impacto da fracionamento e enriquecimento. A relação custo-tempo deve ser considerada em estudos futuros, especialmente com 80 participantes e múltiplas coletas de amostras. Este trabalho fornece dados para a seleção de um protocolo experimental para investigações futuras.

Palavras-chave: Dieta Mediterrânica, Proteoma do plasma humano, Biomarcadores proteicos, Análise de proteínas plasmáticas, Alteração pós-colheita, Protocolo experimental, Preparação de amostras.

Oral and poster presentations

Dieta mediterrânea: Uma dieta à prova dos desafios futuros (2024)

Poster presented by Maria da Graça Pacheco at the 7º Simpósio de Produção e Transformação de alimentos em ambiente sustentável (2024) with authors Eduardo Costa-Camilo, Beatriz Rovisco Pais, Isabel Duarte, and Maria da Graça Pacheco. Doi: https://doi.org/10.6084/m9.figshare.27304404.v1

Consumption of soup in a circular economy (2024)

Presentation by Beatriz Rovisco Pais at the 7º Simpósio de Produção e Transformação de alimentos em ambiente sustentável (2024) with authors Beatriz Rovisco Pais, Eduardo Costa-Camilo, Isabel Duarte, and Maria da Graça Pacheco. Doi: https://doi.org/10.6084/m9.figshare.27304389.v1

Future-proof a mediterranean soup (2024)

Presentation by Eduardo Costa-Camilo at the International Conference on Water, Energy, Food and Sustainability (2024), with authors Eduardo Costa-Camilo, Beatriz Rovisco Pais, Carla Pinheiro, Isabel Duarte, and Maria da Graça Pacheco. This was considered the best presentation at the conference. To be published in ICoWEFS 2024 Sustainability Proceedings (Lecture Notes on Multidisciplinary Industrial Engineering), Springer Nature. Doi: https://doi.org/10.6084/m9.figshare.27304407.v1

Soup was an emotional trigger (2024)

Poster presented by Beatriz Rovisco Pais at the International Conference on Water, Energy, Food and Sustainability (2024), with authors Eduardo Costa-Camilo, Beatriz Rovisco Pais, Carla Pinheiro, Isabel Duarte, and Maria da Graça Pacheco. To be published in the Proceedings of the ICoWEFS 2024 Sustainability Proceedings (Lecture Notes on Multidisciplinary Industrial Engineering), Springer Nature. Doi: https://doi.org/10.6084/m9.figshare.27304407.v1

Session II Food and Health in A New Dimension Assessment of the Influence of the Mediterranean Diet in Human Life Quality above 40s (2023)

Presentation by Maria da Graça Pacheco at the 9th International Conference on Food Chemistry and Technology (2023), with authors Isabel Duarte, Carla Pinheiro, Eduardo Costa-Camilo, Pedro Moreira, Duarte Torres, Olívia Pinho, Elsa Lamy, and Maria da Graça Pacheco. DOI: 10.13140/RG.2.2.21121.02403.

Index

Declaration I
Preface I
Acknowledgments II
Abstract III
Resumo II
Oral and poster presentations III
Table Index VI
Equation Index VI
List of abbreviations VII
1. Objectives 1
2. Introduction 2
2.1 Food: more than nutrients 2
2.2 Dietary choices and sustainability 3
2.3 Diets as a major factor in healthier living 4
2.4 Mediterranean Diet 6
2.4.1 Mediterranean Diet pyramid 7
2.4.2 Mediterranean Diet Health Benefits 8
2.4.3 Mediterranean Diet Sustainability Benefits 10
2.5 Blood analysis to assess the health status 11
2.5.1 The functions of blood cells and plasma: Implications for Health and Disease 11
2.6 Biomarker: definition and classification 13
2.6.1 Plasma biomarkers 14
2.6.2 Challenges in Plasma Proteomics Research 17
2.7 Plasma characterization 19
2.7.1 Protein quantification 19
2.7.2 Plasma antioxidant capacity 21
2.7.3 Plasma proteomics 21
2.8 Ethical and Practical Considerations in Handling Human Biological Samples for Research 23
2.9 Potential of the Mediterranean Diet in increasing quality of life: + health + sustainability 23
3 Materials and Methods 24
3.1 Optimization of the protein quantification protocol 24
3.1.1 Samples 24
3.1.2 Protein Quantification 24
3.2 Optimization of the plasma extraction protocol 25
3.2.2 Samples 25
3.2.2 Protein quantification of samples 27
3.3 General analysis of the polypeptide electrophoresis pattern 27

3.4.1 Sample extraction 27
3.4.2 Folin-Ciocalteu 27
3.2.1 TEAC assay 28
3.3 Optimization of protein extraction 28
3.4 Optimization of running LC-MS/MS conditions 28
3.5 Statistical analysis 29
4 Results and Discussion 30
4.1 Selection of the protein quantification protocol 30
4.2 Optimization of the plasma sampling protocol 31
4.3 Protein plasma polypeptide profile 33
4.4 Antioxidant capacity 35
4.5 Characterization of the samples 37
4.5.1 Protein concentration 37
4.5.2 Extraction method signature 40
4.5.3 Protein Sources 42
4.5.4 Biomarkers analysis 44
4.5.5 Proposed protocol for DM4You 45
5. Conclusion 46
6. References 47
7. Annex 56

Figure Index

Figure 1: **Food waste in the year 2021 in some European countries** 3
Figure 2: **Countries with Mediterranean diet as recognized by UNESCO in 2011** 6
Figure 3: **Mediterranean diet pyramid** 8
Figure 4: **Health benefits of the Mediterranean diet** 9
Figure 5: **Food loss in the years 2016, 2020 and 2021** 10
Figure 6: **Blood constituents** 12
Figure 7: **The use of biomarkers in medicine** 13
Figure 8: **Biomarker distribution across the abundance range** 18
Figure 9: **Absorption spectrum for nucleic acids and proteins** 20
Figure 10: **Protocol used for collecting and processing blood in the pre-trial** 26
Figure 11: **Protein quantification of plasma samples using different quantification method** 30
Figure 12: **Protein quantification of plasma samples using various methods** 32
Figure 13: **Plasma protein samples analysis by SDS-PAGE stained with colloidal Coomassie G250** 33
Figure 14: **Plasma protein samples analysis by SDS-PAGE stained with silver nitrate** 34
Figure 15: **Antioxidant capacity results by Folin-Ciocalteu** 36
Figure 16: **Antioxidant capacity results by TEAC assay** 36
Figure 17: **Spectrum obtained from nLC-MS/MS analysis of a plasma sample** 38
Figure 18: **Spectra obtained for the different dilutions used** 39
Figure 19: **Analysis of the identification results** 41
Figure 20: **Percentages of Identified Protein Sources** 42

Table Index

Table 1: **Common diets**.....5
Table 2: **Biomarker types**. 14
Table 3: **Overview of proteins related with the Mediterranean diet** 15
Table 4: **Protease and phosphatase inhibitors used in this work**.....26
Table 5: **Protein concentration after the lysis process carried out at LIST** 37
Table 6: **Identified Biomarkers** 44

Equation Index

Equation 1: **Warburg and Christian equation.** ……...20

Equation 2: **Kalb and Bernlohr equation**…….20

List of abbreviations

CVD	Cardiovascular Diseases
MD	Mediterranean Diet
NCDs	Noncommunicable diseases
PREDIMED	Prevention with Mediterranean Diet
CRP	C-Reactive Protein
BCA	Bicinchoninic acid assay
EDTA	Ethylenediamine tetraacetic acid
UV-Vis	Ultraviolet-Visible Spectroscopy
WC	Warburg and Christian
KB	Kalb and Bernlohr
TEAC	Trolox Equivalent Antioxidant Capacity
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
LC-MS	Liquid Chromatography–Mass Spectrometry
nLC-MS/MS	Nanoscale liquid chromatography coupled to tandem mass spectrometry
LIST	Luxembourg Institute of Science and Technology
BSA	Bovine Serum Albumin
ENRICH	Enrich-iST Sample Preparation Kit
iST	iST Sample Preparation Kit
Add-on	iST Fractionation Add-on Sample Preparation Kit
PCA	Principal Component Analysis
SCFA	**S**hort-chain fatty acids

1. Objectives

Within the scope of the project DM4You the main objective of this work was to optimise the protocol for the characterisation of plasma samples through proteomics. The DM4You project, funded by the Recovery and Resilience Programme of the Innovation for Agriculture 20|30 «Future Earth» aims to assess how the plasma proteome is influenced by the diet and to select a set of proteins responsive to the diet. This approach will be used to value endogenous products, certified quality products and the Mediterranean Diet, providing evidence of the benefits of this diet.

To this end, a pre-trial to optimise the various steps was taken, including the sampling of blood necessary to assess human plasma proteomics from volunteers. With the goal to preserve the samples at a proteomic level, sampling protocols were assessed, and the samples were also analysed for protein concentration and antioxidant potential.

Due to the protein range intensity, the protocol for Nano LC-MS/MS must also be optimised in order to determinate if fractioning approaches increase the identification of peptides. A literature review of protein biomarkers related to diet was also conducted to compare the obtained results with existing studies.

Since many samples will be processed in the future (80 individuals and 4 moments of collection), the protocol must be optimised also based on time and cost.

2. Introduction

2.1 Food: more than nutrients

The provision of energy and nutrients is fundamental for the survival of all living organisms. In humans, these essential elements are acquired through dietary intake. The early ancestors of *Homo sapiens* embarked on extensive journeys to gather berries, hunt, or fish [1]. Eventually, instead of foraging, they transitioned to agriculture and animal husbandry [1, 2]. This shift marked the end of a nomadic lifestyle, leading to the establishment of small villages which eventually evolved into cities [2].

Over the centuries, the act of eating transcended mere survival and became a cultural phenomenon. It evolved into a sensory experience and gave rise to the early forms of gastronomy, the study of the interrelationships between food, culture, and society [3, 4, 5].

In the 21st century, food remains a cornerstone of human survival, providing necessary energy and organic matter [5]. Beyond sustenance, meals bring families together, fostering social bonds around the dining table. Food is deeply embedded in global culture, with each region developing its unique gastronomy based on locally available ingredients [5, 6]. This regional differentiation has led to the formation of distinct dietary patterns and food chains [7, 8].

Furthermore, food encompasses more than just nutrition and cultural significance. It plays a pivotal role in health, as specific diets can aid in weight loss and reduce the incidence of various diseases [8, 9, 10]. This is largely due to the presence of nutraceutical components in food, which possess properties such as antimicrobial and antioxidant activities [10].

2.2 Dietary choices and sustainability

Dietary choices play a crucial role in combating food waste [11, 12, 13]. Local consumption can significantly reduce the energy expended during transportation and storage processes [13]. Currently, food waste before reaching consumers represents one of the most significant global challenges. Each year, millions of tonnes of food are wasted, as illustrated in Figure 1. This trend has been escalating due to a combination of inefficient production practices and population growth [11, 12]. Addressing this issue requires the development of sustainable strategies to mitigate its impact effectively [11, 12].

Approximately 20% of the food produced is wasted [11, 12]. This issue spans the entire food lifecycle, from harvest and production to distribution and consumption. Therefore, it is crucial to develop and implement methods to repurpose what is currently deemed waste [12].

From a scientific perspective, food waste remains one of the most critical areas of interest. Daily, new strategies are developed to mitigate food waste [11, 12, 13]. Consequently, sustainable development has become a central focus within food sciences and related disciplines.

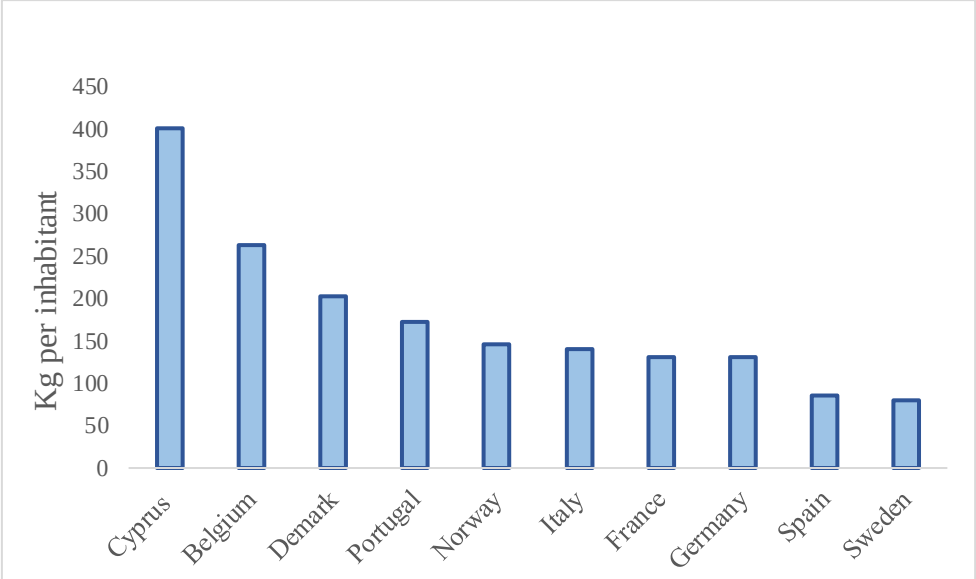


Figure 1: **Food waste in the year 2021 in some European countries.** Several kilograms of food are wasted annually per capita before reaching consumers, resulting in substantial quantities of wasted food at the national level. Portugal ranks as the fourth highest country in terms of food waste [12, 13]. Adapted from [14].

Moreover, food production is one of the industries that contributes most significantly to pollution through the release of greenhouse gases and excessive consumption of water and energy [11]. The extensive use of agrochemicals in large-scale production further leads to soil and aquatic habitat contamination [11, 12]. Additionally, food distribution often involves exportation, which increases transportation costs and pollution [11, 12].

Therefore, it is imperative to innovate and restructure these industries to meet food production needs while also combating pollution and food waste [12]. Reducing food waste concurrently reduces environmental pollution [13].

One effective approach is to disseminate information to the public, raising awareness about the problem and preventive measures. Advocating for more careful and efficient agricultural practices, along with promoting a balanced and sustainable diet, is essential [11, 12, 13].

2.3 Diets as a major factor in healthier living

A diet is based on principles that enable individuals to improve their health in a less invasive manner compared to surgical procedures or pharmacological interventions [12, 13]. Consequently, a well-structured diet can enhance one's quality of life [12, 14].

Typically, weight loss is a primary motivation for adopting specific dietary habits [15]. However, diets offer additional benefits as the consumption of certain foods can provide nutraceutical advantages. For instance, conditions such as cardiovascular diseases (CVD) can be managed through a healthy and regulated diet [12, 13, 14]. Planned dietary intake can reduce the incidence of various diseases, thereby increasing healthy longevity [12, 14, 16].

Many physicians advocate for dietary control to mitigate various health conditions, such as hypertension [17]. This condition can be alleviated by consuming foods like bananas, which are rich in potassium - a mineral that helps manage blood pressure by counteracting sodium's effects and reducing tension in blood vessel walls [18]. Similarly, watermelon, which is rich in citrulline, an amino acid that aids in the production of nitric oxide, improves blood flow by enhancing arterial flexibility, thereby lowering high blood pressure [18, 19].

There are numerous examples of foods that can reduce and control medical conditions. The establishment of dietary guidelines has led to the formulation of various diets worldwide (Table 1).

Table 1: **Common diets** [20]. Diets offer non-invasive health improvements compared to surgery or medications, enhancing quality of life and managing conditions like cardiovascular diseases through nutraceutical properties.

Diet	Acronym	Description
Dietary Approaches to Stop Hypertension	DASH	Prevents and reduces hypertension. Strict in the amounts ingested and the type of food consumed. It is a balanced diet and can be followed in the long term by the whole family.
Flexitarian Diet	FD	Diet that does not eliminate animal protein but gives priority to vegetables. Those who follow this diet can not only lose weight, but also see their general health improve, reducing the incidence of heart disease, diabetes and even some types of cancer.
Mediterranean Diet	MD	National cuisine in 6 countries. A geographical diversity that brings with it many foods for the same diet.
Mediterranean-DASH Intervention for Neurodegenerative Delay	MIND	This diet is a combination of the Mediterranean diet and DASH and was developed to delay neurological decline and the risk of developing dementia.
Therapeutic Lifestyle Changes	TLC	It aims to reduce cholesterol as part of a heart-healthy diet to prevent cardiovascular disease.

2.4 Mediterranean Diet

This work contributes to the DM4You project, with the main goal to promote the Mediterranean Diet (MD). MD is widely recognized for its significant benefits to health and quality of life, as endorsed by both the scientific community and the general public [14, 21]. The recognition of the MD's advantages dates to the mid-20th century [21, 22, 23], when it was observed that the population of the island of Crete exhibited higher longevity levels compared the population in North America [22, 23]. In fact, the term "Mediterranean Diet," was created by Ancel Keys following his observations of Mediterranean populations that had a low incidence of CVD and a higher life expectancy which was attributed to their dietary habits [21, 23, 24, 25].

In fact, a correlation was established between fat consumption and CVD mortality, with higher fat intake associated with increased risk [21, 23, 24, 25]. Populations in the Mediterranean Basin (Figure 2) were an exception: despite their high fat consumption, low CVD mortality was registered [23, 24]. Keys attributed this to the type of fat consumed, which in these populations was primarily unsaturated fat, derived mainly from olive oil [22, 23, 24].

Together with the prominent use of non-heated olive oil as the primary source of monounsaturated fat [21, 22, 26], the MD is characterized by a high intake of cereals, legumes, fresh fruits, nuts, fish, and a substantial presence of vegetables. Conversely, the consumption of potatoes, meat, dairy products, eggs, and sugar was moderate [23, 25, 26]. This dietary pattern reflects the traditional eating habits of several countries in the Mediterranean area (Figure 2) [27].

Figure 2: **Countries with Mediterranean diet as recognized by UNESCO in 2011** (blue stars). Seven emblematic communities comprise the representation for this cultural heritage: Cyprus, Croatia, Spain, Greece, Italy, Morocco, and Portugal. These communities both protect the history and cultural assets of this heritage, as well as help to promote its values and practices throughout the world. Adapted from [28].

Beyond its dietary and nutritional components, the concept of the MD also encompasses broader aspects such as social interaction, coexistence, biodiversity, and seasonality [21, 22, 27]. These elements collectively contribute to the MD being recognized as more than just a diet; it represents a comprehensive lifestyle [21, 23].

2.4.1 Mediterranean Diet pyramid

There is no universal consensus on the definition of the dietary pattern of MD, due to the diversity of diets practiced across different countries in the Mediterranean region, and even within various regions of the same country [21, 22, 23, 27]. However, certain common characteristics have been identified, including: daily consumption of minimally refined cereals, a high intake of both raw and cooked vegetables, extensive use of virgin olive oil as the primary fat source, high consumption of seasonal fruits, regular intake of fatty fish, moderate consumption of dairy products, and limited consumption of red meat and its derivatives [21, 22, 29].

A dietary guide translates the MD into a pyramid, illustrated in Figure 3, which outlines the recommended frequency and quantity of food intake, as well as lifestyle practices adapted to the contemporary era [23, 30, 31]. This MD pyramid represents an international consensus on the scientific evidence supporting this dietary pattern, specifically targeting individuals aged 18 to 65. The base of the pyramid features foods that should be consumed more frequently and in larger quantities, while the upper levels include foods that should be consumed in moderation and with limited frequency [31].

This guide also emphasizes the importance of physical activity and integrates cultural and social aspects of the Mediterranean lifestyle [32]. Its flexibility accommodates individual preferences, facilitating long-term adherence [30, 31, 32].

Figure 3: **Mediterranean diet pyramid.** In this scheme it is possible to observe the typical foods of this diet as well as the regularity with which they should be consumed. It is also possible to observe the indication for sports practice that must be accompanied with this diet but also characteristics beyond it that accompany it such as coexistence and seasonality. Adapted from [31].

The dietary and nutritional attributes of the MD are foundational to its health benefits [33, 34]. The MD's food composition ensures the provision of all essential micronutrients [32, 33]. Additionally, prioritizing the consumption of local, fresh vegetables maximizes the intake of antioxidant compounds, which can degrade during transportation and over time [33, 34]. The MD's seasonal nature, characterized by harvesting foods according to their natural growing season, also promotes nutritional variety and lower ecological footprint.

2.4.2 Mediterranean Diet Health Benefits

Adherence to the MD is generally associated with a reduced incidence of non-communicable diseases (NCDs). NCDs, which are not transmitted through infection or person-to-person contact, are typically linked to unhealthy behaviours and include conditions such as heart disease, cancer, chronic respiratory diseases, and diabetes. These conditions are the leading causes of death globally and represent a significant and growing threat to global health [21, 25, 35]. NCDs are responsible for 41 million deaths annually, accounting for over 70% of all deaths worldwide [35].

Numerous studies have demonstrated that adherence to the MD is strongly associated with increased longevity by reducing mortality risk [21, 25, 26]. This dietary pattern is particularly effective in reducing the incidence of diseases such as CVD and type 2 diabetes mellitus [36]. Figure 4 illustrates some of the diseases that can be managed through adherence to the MD.

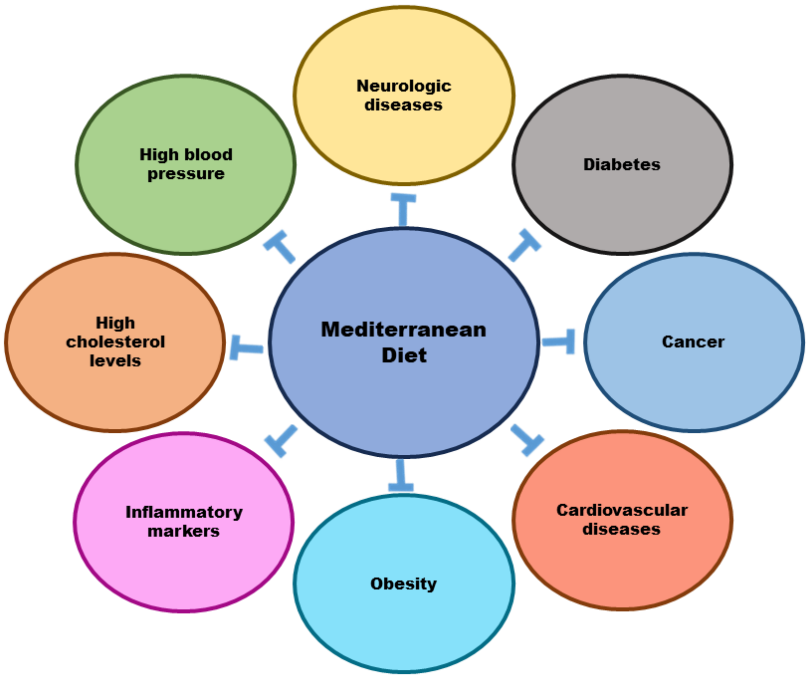


Figure 4: ***Health* benefits of the Mediterranean diet**. The consumption of this diet can lead to a reduction in the occurrence of certain pathologies. Adapted from [21].

It is estimated that 90% of type 2 diabetes mellitus cases and 70% of strokes could be prevented by adopting the MD in conjunction with regular physical activity [37]. The extensive benefits of the MD are well-documented in numerous studies, including the PREDIMED trial. This study was a large-scale, long-term clinical trial conducted in Spain, which aimed to investigate the effects of the MD on cardiovascular disease prevention. It enrolled over 7,000 participants at high risk of cardiovascular events, primarily aged 55-80. The main conclusion demonstrated that the MD effectively prevents CVD in both diabetic and non-diabetic populations [35, 38, 39].

Moreover, significant statistical and clinical improvements in glycaemic control have been observed in individuals adhering to the MD [40]. This dietary pattern has proven more effective in improving insulin resistance and glucose tolerance compared to control and low-fat diets, particularly in obese, non-diabetic individuals [40, 41]. Additionally, several studies have reported the use of olive oil, a key component of the MD, as a preventive measure against inflammation and oxidative stress, with a notable increase in plasma antioxidant capacity [36, 38, 39, 40].

Given these characteristics, the MD is not only a model for dietary and nutritional recommendations for the general population but also for individuals who have already developed certain diseases [37].

2.4.3 Mediterranean Diet Sustainability Benefits

In addition to its clinical benefits, the MD also offers significant ecological advantages. Numerous studies have shown that the MD has a lower environmental impact than other dietary patterns. Being predominantly plant-based with low consumption of animal products, it results in a smaller water footprint and lower greenhouse gas emissions [11, 21, 25]. The seasonality inherent to the MD promotes more sustainable consumption and reduces associated costs, contributing to lower energy consumption. The emphasis on local products further reduces transportation needs, leading to decreased fuel usage and lower prices [11, 12, 21, 25].

Another critical aspect is the MD's association with reduced food waste. Figure 5 illustrates the gradual increase in food waste over recent years. The Northern European region presents a higher food waste percentage than the other European regions. This fact can be explained by considering different factors, such as the economic level of each location and the local culture itself, for example the Eastern Europe represent poorer countries [25]. Additionally, it may be associated with the type of diet in each region, with the MD being more prevalent in southern European regions.

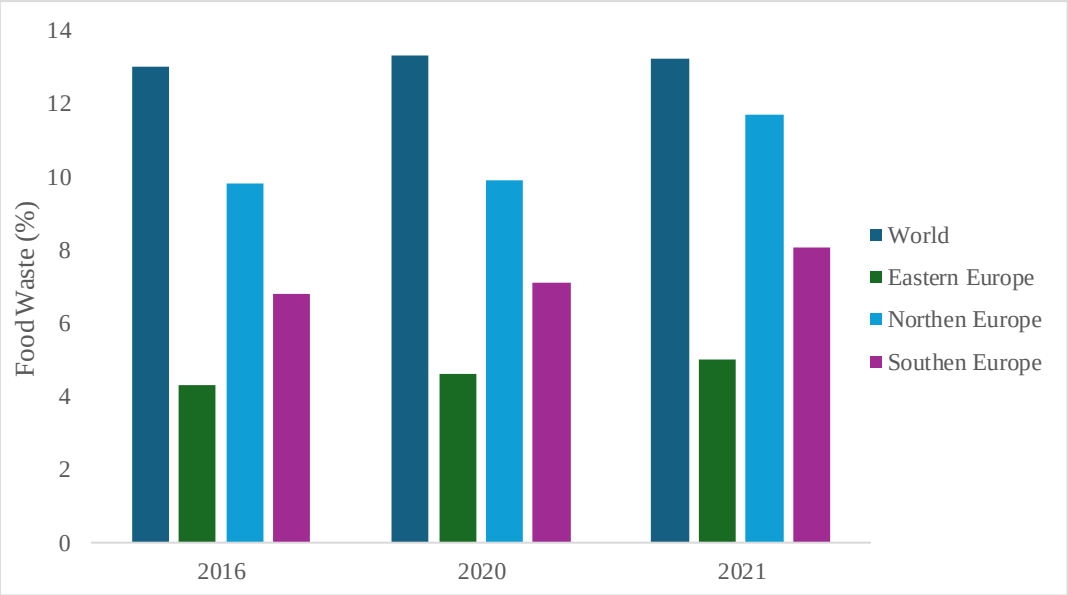


Figure 5: **Food loss in the years 2016, 2020 and 2021.** Several kilos are wasted per year per inhabitant before reaching consumers, which ultimately leads to tons of wasted food in each country. Adapted from [25].

2.5 Blood analysis to assess the health status

One of the primary methods for assessing human physiological functions is through clinical analysis of blood, which allows for the rapid diagnosis of numerous diseases due to its ability to provide comprehensive information about the body's internal state [42, 43]. Blood serves as a critical physiological fluid for evaluating various health parameters, as it contains vital biomarkers that reflect the overall health status of an individual. For instance, the presence above 0.3 mg/mL of C-reactive protein (CRP) in the blood can indicate inflammation and is often used to assess the risk of CVD, among other conditions [42, 43].

2.5.1 The functions of blood cells and plasma: Implications for Health and Disease

Blood cells serve a multitude of critical functions, ranging from transportation to protection [44]. Blood comprises both a fluid component (plasma) and cellular components, including three primary types of cells: erythrocytes, leukocytes, and thrombocytes. Erythrocytes, or red blood cells, are rich in haemoglobin (a protein containing iron) that is essential for the transport of oxygen and carbon dioxide throughout the body [43, 44]. Leukocytes, or white blood cells, are integral to the immune system, responsible for eliminating aged cells, cellular debris, and combating infectious agents and foreign substances [45]. Thrombocytes, commonly known as platelets, play a crucial role in blood clotting, preventing excessive bleeding [45].

These cells are suspended in blood plasma, the liquid portion of blood, which is characteristically yellow. Plasma constitutes approximately 55% of the total blood volume, as depicted in Figure 6 and is primarily composed of water (up to 95% by volume) [38, 46]. Plasma is essential in maintaining the intravascular osmotic balance, ensuring that electrolyte concentration remains stable, and in protecting the body from infections and other blood-related disorders [46, 47].

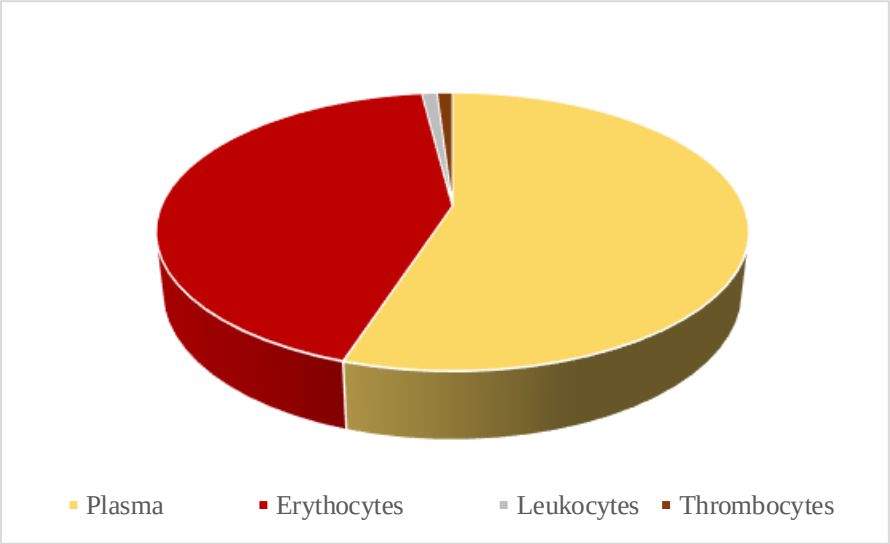


Figure 6: **Blood constituents.** Plasma is the main constituent representing approximately 55%. In second place, erythrocytes or red blood cells which represent 43% of the blood volume. And finally, with approximately 1% representation, thrombocytes and leukocytes. Data from [46, 47].

Plasma contains numerous proteins, minerals, and other components that can serve as biomarkers, indicators of potential abnormalities, such as cellular aging or disease processes, thereby making it an essential tool in medical diagnostics and monitoring [46, 47].

2.6 Biomarker: definition and classification

Biomarkers are defined as measurable characteristics that serve as indicators of normal biological processes, pathogenic activities, or responses to various exposures or interventions [47, 48]. These biomarkers are pivotal in the early detection, diagnosis, and prognosis of diseases, as well as in monitoring therapeutic responses. Biomarkers can encompass specific cells, molecules, genes, proteins, or even patterns of gene expression and metabolic processes [48].

To be of clinical utility (Figure 7), a biomarker must not only exhibit high reproducibility and a substantial signal-to-noise ratio but also demonstrate sensitivity and specificity for the condition being assessed. This ensures that the biomarker reliably reflects the biological state or disease process it is intended to measure. Additionally, it is crucial for a biomarker to change dynamically and consistently with the progression or regression of a clinical condition, allowing for accurate monitoring of disease status or treatment efficacy [47, 48]. They can be categorized into various types, including physiological, physical, molecular, or histological markers, as detailed in Table 2.

For effective integration into routine clinical practice, biomarkers must be easily measurable from accessible biological samples, utilizing cost-effective and standardized techniques [48]. Historically, plasma and tissue samples obtained from biopsies have been the cornerstone of biomarker discovery.

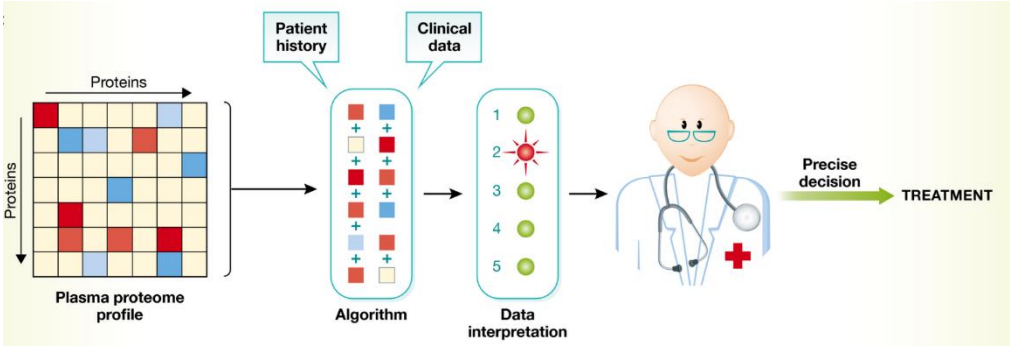


Figure 7: **The use of biomarkers in medicine.** Biomarkers will aid the physician in making more precise recommendations for treatment, while still taking patient history and other clinical data into account. Adapted from [48].

Table 2: **Biomarker types.** According to their applications, biomarkers can provide complementary information about the disease or the intervention under consideration. Biomarkers may be identified at any event occurring since the pathogenesis, the onset of first clinical manifestations, diagnosis, and treatment outcome. Data from [47].

Use	Description
Monitoring	This category includes biomarkers that are analysed at different time points to monitor the status of a disease or medical condition, and as a marker of the response to an intervention, including exposure to a medical product or an environmental agent.
Diagnostic	Biomarkers used to detect or confirm the presence of a disease or medical condition. The advent of the era of precision medicine emphasizes the fact that diagnostic biomarkers are useful not only to identify patients with a disease, but also to redefine its classification.
Safety	Any measure that can be assessed before and after the exposure to a medical intervention, or an environmental agent, allowing to identify the probability of developing signs of toxicity as an adverse event, to detect the presence of toxicity, and for monitoring its extension.
Predictive	A marker is considered a predictive biomarker when its presence or modification allows predicting which patient or group of patients are more likely to experience an effect as consequence of being exposed to a medical product or environmental agent.
Pharmacodynamic or Response	Proposed to be a potential useful tool in clinical practice providing useful information for patient management. A pharmacodynamic biomarker is modified in response to a medical condition or clinical intervention, including drug treatments.
Susceptibility or Risk	Risk measure to develop a disease or medical condition.

2.6.1 Plasma biomarkers

The following table (Table 3) was obtained from a review study realized by Fátima Cardoso, a member of the project DM4You, in which articles published between 1990 and 2024 were searched on the Web of Science using the Science Citation Index Expanded (Sci-EXPANDED) filter [49]. In this study, the keywords "Plasma Biomarkers" and "Mediterranean Diet" were used, resulting in a total of 231 published articles. Table 3 presents the biomarkers deemed relevant through this research. A thorough understanding of their functions and interactions is vital for advancing scientific knowledge in disease detection, therapeutic innovation, and the development of personalised medicine strategies [49-66].

Table 3: **Overview of proteins related with the Mediterranean diet.** Information extracted from Web of Science using the Science Citation Index Expanded (Sci-EXPANDED) filter [49-66].

Protein	Gene	Summary
Adipokines- Adiponectin	*ADIPOQ*	An adipokine secreted by adipose tissue that plays a role in regulating glucose levels and fatty acid breakdown. It has anti-inflammatory and insulin-sensitizing properties.
Adipokines- leptin	*LEP*	A hormone produced by adipose tissue that regulates energy balance by inhibiting hunger.
Angiotensin converting enzyme	*ACE*	An enzyme that converts angiotensin I to the potent vasoconstrictor angiotensin II, a critical component of the renin-angiotensin system. Inhibitors are commonly used in treating hypertension and heart failure.
Apolipoprotein E	*APOE*	Protein involved in lipid metabolism, particularly in the transport and clearance of cholesterol and other lipids from the bloodstream. Associated with different risks for cardiovascular disease and Alzheimer's disease.
Catalase	*CAT*	An enzyme that catalyses the decomposition of hydrogen peroxide into water and oxygen and is a crucial component of the antioxidant defence system.
Cellular communication network factor 5	*CCN5*	A matricellular protein that plays a role in cell proliferation, differentiation, and extracellular matrix remodelling.
Cholesterol ester transfer protein	*CEPT*	Facilitates the transfer of cholesterol esters and triglycerides between lipoproteins. Modulation of this protein activity is being explored as a therapeutic target for cardiovascular diseases.
Clock circadian regulator	*CLOCK*	A core component of the circadian clock, which governs the daily rhythms of biological processes.
Coenzyme Q10	*COQ10A*	A lipid-soluble antioxidant and a critical component of the mitochondrial electron transport chain, where it plays a role in ATP production. conditions, including cardiovascular diseases.
C-reactive protein	*CRP*	An acute-phase protein produced by the liver in response to inflammation. It is widely used as a biomarker for systemic inflammation and is associated with an increased risk of cardiovascular disease when elevated.
E-selectin	*SELE*	An adhesion molecule expressed on endothelial cells activated by cytokines. It mediates the rolling of leukocytes on the endothelium during the early stages of inflammation.
Glutathione disulfide reductase	*GSR*	An enzyme that reduces glutathione disulfide to the sulfhydryl form glutathione, an important cellular antioxidant.
Glutathione peroxidase	*GPX1*	An enzyme that reduces lipid hydroperoxides to their corresponding alcohols and free hydrogen peroxide to water.
Growth differentiation factor 15	*GDF15*	A stress-responsive cytokine associated with inflammation, apoptosis, and tissue repair. Elevated levels are linked to cardiovascular diseases, cancer, and other chronic conditions.
Heme oxygenase- Heme oxygenase 1	*HMOX1*	An enzyme that catalyses the degradation of heme to biliverdin, free iron, and carbon monoxide and its induction is a key response to oxidative stress and tissue injury.
Hypoxia inducible factor 1 subunit alpha	*HIF1A*	A transcription factor that responds to low oxygen levels. It plays a key role in cellular adaptation to hypoxia and is implicated in cancer and ischemic diseases.

Table 3: **Continuation.**

Protein	Gene	Summary
Insulin	*INS*	A hormone produced by the pancreas that plays a crucial role in regulating blood glucose levels. Dysregulation of insulin function is a key factor in diabetes mellitus.
Insulin like growth factor 1	*IGF1*	A hormone with structural similarity to insulin, playing a crucial role in growth and development.
Integrin subunit alpha 4	*ITGA4*	A component of integrins, which are cell surface receptors that mediate cell adhesion to the extracellular matrix. It plays a crucial role in immune cell trafficking and inflammation.
Interleukin 10	*IL-10*	An anti-inflammatory cytokine that regulates immune responses by inhibiting the production of pro-inflammatory cytokines.
Interleukin 6	*IL-6*	A cytokine that plays a dual role in inflammation and immune associated with chronic inflammatory conditions and autoimmune diseases.
Lipoprotein associated phospholipase A2	*PLA2G2A*	An enzyme associated with low-density lipoprotein particles. It plays a role in the inflammatory processes of atherosclerosis and is considered a biomarker for cardiovascular risk.
MLX interacting protein like	*MLXIPL*	A transcription factor that regulates genes involved in glucose and lipid metabolism, contributing to the control of energy homeostasis.
Paraoxonases	*PON1, 2 and 3*	A group of enzymes associated with high-density lipoprotein that protect lipids from oxidation. They play a role in cardiovascular health by preventing the oxidation of low-density lipoproteins.
Peroxisome proliferator activated receptor γ2	*PPARG2*	A nuclear receptor that regulates the expression of genes involved in adipogenesis, lipid metabolism, and insulin sensitivity.
Pleiotrophin	*PTN*	A growth factor involved in angiogenesis, neural development, and tissue repair.
Serine/threonine-protein kinase mTOR	*MTOR*	A central regulator of cell growth, proliferation, and metabolism.
Soluble intercellular adhesion molecule	*ICAM*	Involved in leukocyte adhesion and transmigration across the endothelium. Elevated levels are associated with inflammation and have been linked to cardiovascular and autoimmune diseases.
Soluble Vascular Cell Adhesion Molecule-1	*VCAM-1*	A protein involved in the adhesion of leukocytes to the endothelium. Elevated levels are indicative of endothelial dysfunction and are associated with various inflammatory diseases.
Superoxide dismutase	*SOD1, 2 and 3*	A family of enzymes that catalyse the dismutation of superoxide radicals into oxygen and hydrogen peroxide, providing a critical defence against oxidative stress.
Transgelin	*TAGLN*	An actin-binding protein involved in the regulation of smooth muscle cell contraction and cytoskeletal organization and is implicated in vascular diseases.
Tumour necrosis factor	*TNF-alpha*	A pro-inflammatory cytokine that plays a central role in the immune response to infection and in the pathogenesis of chronic inflammatory diseases.

2.6.2 Challenges in Plasma Proteomics Research

The proteome refers to the complete set of proteins expressed by an organism at a specific time, representing the dynamic nature of protein expression under various physiological and environmental conditions [67, 68]. Essentially, proteomics is the large-scale study of proteins, encompassing their structures, functions, and interactions within an organism [68]. The human proteome is subject to modulation by numerous factors, including pharmacological interventions and lifestyle choices and it exhibits significant variation with age [67, 68].

The sequence of amino acids in a protein is determined by the nucleotide sequence of its corresponding gene, underscoring the direct link between genomics and proteomics [68]. After translation, proteins often undergo post-translational modifications, which chemically alter amino acid residues. These modifications significantly impact the protein's physical and chemical properties, including its folding, stability, and activity, ultimately influencing its biological function [67, 68]. The specific three-dimensional structure that results from these modifications dictates the protein's localization, regulatory mechanisms, and functional roles within the cell [68].

The plasma proteome is one of the richest and most complex proteomes in the human body [47, 68, 69]. The proteins within the plasma proteome can be classified into three distinct categories. One category includes abundant proteins with functional roles in the blood, such as human serum albumin, which constitutes approximately 55% of all plasma proteins [68, 69, 70]. This protein is crucial for maintaining osmotic pressure and facilitating the transport of insoluble molecules, as well as playing roles in the acute phase immune response and the coagulation cascade [69, 79]. Another category comprises tissue leakage proteins, which lack a specific function in plasma but reflect some health conditions. Examples include enzymes like aspartate aminotransferase and alanine aminotransferase, which are commonly used in the diagnosis of liver diseases [71]. The last category consists of signalling molecules, including small protein hormones like insulin and cytokines, which are typically present in low abundances under normal conditions but are upregulated when needed [71].

Plasma interacts directly with nearly all organs and tissues, and its protein composition is differentially regulated during disease progression. Consequently, plasma is a key source of potential pharmacodynamic biomarkers [72]. A thorough understanding of plasma proteins under both physiological and pathological conditions is crucial for the discovery of molecular markers and drug targets [72]. Figure 8 illustrates the percentage of plasma proteins that are established biomarkers, as well as those that may become biomarkers, with a proteomic depth of 1,500 proteins considered [48]. More abundant proteins are easier to study and among the 300 most abundant proteins, one in four is a biomarker. Among the other 1,200 proteins, only one in 25 are described to have this role [48]. This suggests that many potential biomarkers remain undiscovered [48]. Currently, 27% of plasma proteins are recognized as biomarkers, and for this study, we have selected specific biomarkers due to their clinical significance and widespread use [48].

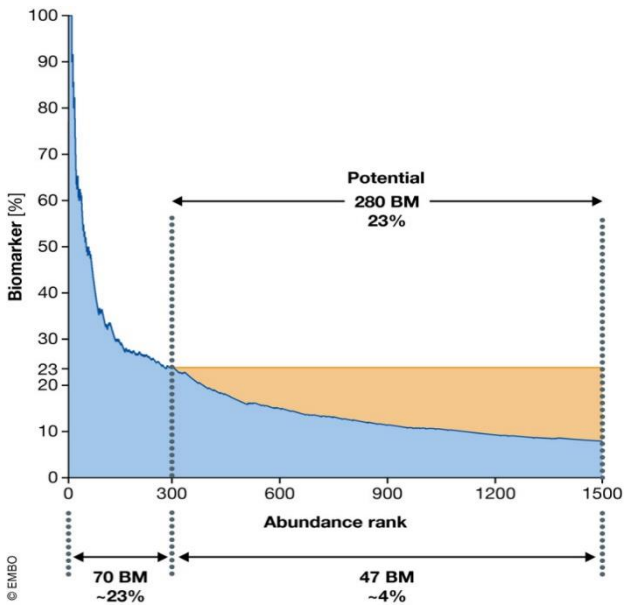


Figure 8: **Biomarker distribution across the abundance range.** The blue area illustrates the percentage of biomarker (BM) as a function of increasing depth of the plasma proteome. The top of the yellow region extrapolates this proportion to the remainder of the plasma proteome. If the portion of biomarkers remained as high as it is in the 300 most abundant proteins, there are at least 233 potential biomarkers to be discovered. Adapted from [48].

The wide dynamic range of protein concentrations in plasma makes proteome analysis particularly challenging [72]. Approximately 22 proteins, with concentrations reaching up to mg/mL, account for 99% of the total plasma protein content, with albumin, immunoglobulins, and fibrinogen making up 90% of this fraction [72]. In contrast, thousands of other proteins of potential interest exist in plasma at much lower concentrations, often in the ng/mL or even pg/mL range [72]. Although these low-abundance proteins may be critical indicators of physiological function and disease, the high-abundance proteins create a significant "masking" effect, complicating their detection [72]. Therefore, the development proteomics technologies are essential for achieving comprehensive profiles of the plasma proteome [71, 72].

2.7 Plasma characterization

2.7.1 Protein quantification

The initial step in plasma characterization is protein quantification, an essential step for proteomics studies. This quantification can be either targeted or untargeted. Untargeted, or total protein quantification, involves measuring the overall protein concentration in each sample [73]. Numerous methods exist for untargeted protein quantification, each with inherent biases. Common techniques include direct measurement via absorbance at 280 nm and indirect colorimetric assays [73].

A widely adopted method for protein quantification is the Bradford assay [74], which is based on the binding of Coomassie Blue G250 dye to proteins [74, 75]. Protein concentration is estimated by quantifying the dye in its blue ionic form, typically through absorbance measurement at 595 nm [75]. The dye exhibits a strong affinity for arginine and lysine residues, and to a lesser extent, for histidine and aromatic residues like tryptophan, tyrosine, and phenylalanine [75]. This specificity may cause variability in the response of different proteins, which is considered its primary limitation. Various modifications, such as sample acidification to enhance dye binding, have been introduced to mitigate this issue [74, 75].

Another prominent method is the Bicinchoninic acid protein assay (BCA), which operates on the principle that proteins reduce Cu^{2+} to Cu^{+} in an alkaline solution, resulting in the formation of a purple colour [76]. This reduction is predominantly driven by four amino acids: cysteine, tyrosine, and tryptophan [76, 77]. Unlike Coomassie dye-binding methods like Bradford, the BCA assay benefits from the contribution of the universal peptide backbone for colour formation, thereby reducing variability due to differences in protein composition [76, 77]. As Bradford, the BCA assay is vulnerable to interference of certain chemicals found in protein samples, such as reducing agents (e.g., β-mercaptoethanol), metal chelators (e.g., EDTA), and buffers [77].

Another method for protein quantification involves measuring the sample's absorbance, as initially described by Warburg and Christian [78]. Protein concentration can be inferred by directly assessing absorbance at 280 nm [77, 78], since the aromatic rings of specific amino acids, primarily tryptophan and tyrosine, and to a lesser extent phenylalanine, absorb ultraviolet light (UV) at 280 nm [77, 78]. Despite its utility, this method is biased as only certain amino acid residues contribute to absorbance, and various non-protein components can interfere with the measurements. The most common interfering substances include nucleic acids, nucleotides, heme-containing compounds like cytochromes, and reagents with sulfhydryl groups such as 2-mercaptoethanol, dithiothreitol, or glutathione [78]. Nucleic acids pose a significant challenge as they exhibit a maximum absorption peak at 260 nm, due to their nitrogenous bases, yet also absorb at 280 nm (Figure 9). In cell-free extracts, nucleic acids are typically the predominant interfering substances, an UV spectrophotometric assay must be carefully designed to minimize these interferences.

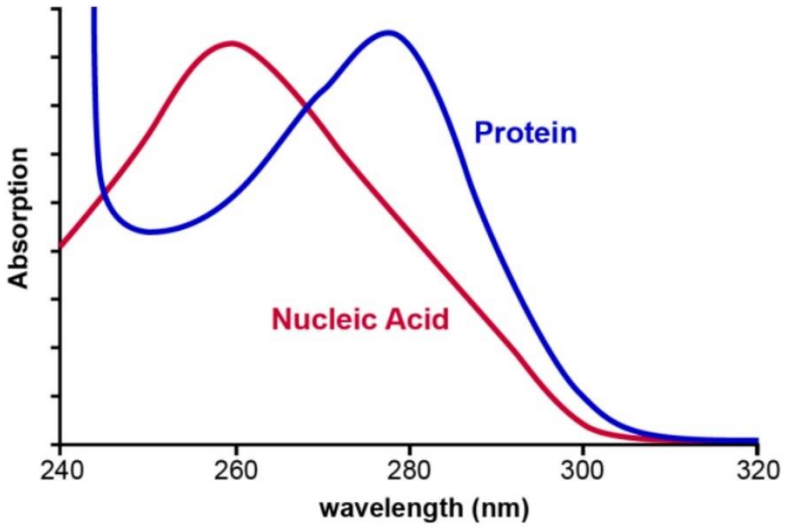


Figure 9: **Absorption spectrum for nucleic acids and proteins.** Nucleic acids have a maximum absorption peak at 260 nm, proteins at 280 nm.

Several authors have proposed equations to correct the nucleic acid interference in protein quantification. Warburg and Christian (WC), offer an equation applicable for a 1 cm pathlength and an extinction coefficient of 1, with protein concentration expressed in mg/mL [78, 79]. This method also stipulates that the absorbance at 320 nm should be below 0.02, corresponding to unsignificant impurity levels in the sample [78].

$$Protein\ concentration\ (\frac{\text{mg}}{\text{ml}}) = [(1.55 \times A280) - (0.76 \times A260)] \times diluition\ fator \quad (1)$$

Equation 1: **Warburg and Christian equation.**

The WC method can only accommodate low levels of nucleic acid interference, primarily because it relies on the natural absorbance of the amino acids tryptophan, tyrosine, and phenylalanine [78, 79]. However, the proportion of these amino acids varies significantly, resulting in considerable protein-to-protein variability [78, 79].

To address these limitations, Kalb and Bernlohr (KB) developed a more sensitive spectrophotometric method that also accounts for the presence of rRNA and tRNA [64]. This method also uses an equation applicable for a 1 cm pathlength and an extinction coefficient of 1 but utilizes a different wavelength (230 nm) that accounts for peptide bonds, providing a more accurate protein concentration measurement, typically expressed in µg/mL [79].

$$Protein\ Concentration\ \left(\frac{\mu\text{g}}{\text{ml}}\right) = [(183 \times A230) - (75.8\ \times A260)] \times dillution\ fator \quad (2)$$

Equation 2: **Kalb and Bernlohr equation.**

2.7.2 Plasma antioxidant capacity

Biochemical assays are essential tools in plasma characterization, enabling the detection and quantification of specific biomolecules, including enzymes, lipids, and metabolites [80]. Folin-Ciocalteu and Trolox Equivalent Antioxidant Capacity (TEAC) are willey used as proxies of antioxidant status, both having limitations like lack of specificity, dependence on reaction conditions or interference [80-83]. However, they provide an easy, fast and affordable way to characterized plasma samples.

The Folin-Ciocalteu method is widely used for the quantification of phenolic compounds as they are the main antioxidant detected by this method [81, 82]. This colorimetric assay involves the reduction of the Folin-Ciocalteu reagent by phenols and other reducing agents, leading to a measurable colour change [81, 82]. This method is particularly useful in assessing the total antioxidant capacity of plasma, providing valuable information on the oxidative stress status of the individual [81].

The TEAC assay measures the antioxidant capacity of plasma samples based on the ability of antioxidants to scavenge the 2,2'-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) ($ABTS^{+}$) radical cation [83]. This assay is crucial to understand the balance between oxidative stress and antioxidant defence mechanisms in various pathological conditions [83].

These two methods are willey used as proxies of antioxidant status, both having limitations like lack of specificity, dependence on reaction conditions or interference from coloured samples [80-83].

2.7.3 Plasma proteomics

2.7.3.1 Gel base approach

Gel-based methods are traditional yet powerful techniques in proteomics, with two-dimensional gel electrophoresis being the most prominent. In this approach, proteins are first separated based on their isoelectric point and then by their molecular weight through sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) [73]. Following separation and visualization, proteins can be excised from the gel and identified through mass spectrometry (MS) [73]. Although gel-based approaches are highly effective for resolving complex samples and detecting protein isoforms, they are less suited for high-throughput analysis and can struggle with detecting low-abundance proteins in plasma.

2.7.3.2 Gel free approach

Gel-free methods, such as liquid chromatography (LC) coupled with MS (LC-MS), have gained prominence due to their ability to provide high-throughput, quantitative, and comprehensive analysis of the plasma proteome [73, 84]. In gel-free approaches, proteins are digested into peptides, which are then separated by LC before being analysed by MS [73]. This technique enables theoretically the identification of thousands of proteins in a single experiment. Gel-free methods are particularly advantageous for the large scale study of complex proteomes and have become the method of choice for large-scale proteomic studies [73].

2.7.3.2.1 Nano LC-MS/MS

MS it is based on ionising chemical compounds to generate charged molecules or molecule fragments, which are then separated by their mass-to-charge ratio in an electric or magnetic field [48]. The resulting mass spectrum can provide information about the molecular weight and structure of the analyte. Over the past years, MS proteomics has seen remarkable advancements, becoming the mainstream approach due to its capability to provide highly accurate mass and fragmentation spectra of peptides derived from sequence-specific digestion of proteins, alongside its robust sensitivity [48].

Since the introduction of LC-MS in proteomics, the strategies for protein analysis have shifted significantly from gel-based methods to gel-free approaches. This transition is largely due to the ability of LC-MS to identify a vast number of proteins in a single run, as well as the timespan of the analysis process, which takes hours rather than days.

LC-MS-based proteomics holds the potential to analyse a big part of the proteome of a biological system [48]. However, achieving in-depth and reproducible coverage of complex proteomes remains challenging. This difficulty arises because the complexity of the digests subjected to LC-MS analysis often exceeds the analytical capacity of mass spectrometers, leading to data under-sampling [84].

Thus, even current proteomic techniques are not yet capable of analysing all the constituent proteins of a sample in a single analysis. To address this limitation, various strategies have been employed to fractionate samples and enrich specific proteins. Protein enrichment typically targets proteins based on their abundance [84]. Numerous methods have been developed to deplete high-abundance proteins in plasma or serum, thereby facilitating the detection of lower-abundance species [69].

The evolution of LC-MS has also brought about the development of advanced techniques such as tandem LC-MS and Nano LC-MS/MS (nLC-MS/MS), which offer greater sensitivity and precision in protein identification [85]. LC-MS/MS involves an additional stage of MS, allowing for more detailed fragmentation and thus more accurate protein identification and quantification [85]. nLC-MS/MS, on the other hand, utilizes nanoscale LC to enhance the separation of peptides before they enter the mass spectrometer, improving detection sensitivity and resolution [85]. These advancements have further

cemented the role of LC-MS-based techniques as the gold standard in proteomic research, enabling more comprehensive proteome analyses.

2.8 Ethical and Practical Considerations in Handling Human Biological Samples for Research

When working with biological samples derived from humans, certain precautions are necessary. For instance, safety conditions must be in place to prevent the transmission of diseases between the donor and the operator, as well as to ensure the maintenance of anonymity [86]. As such, approval from the relevant ethics committees is required, ensuring the study complies with applicable legislation.

In addition to legal parameters, particularly with blood samples, the biological stability of the sample must be maintained. Blood is highly sensitive to temperature fluctuations and handling, which can compromise the integrity of the sample [86]. To preserve the sample, it is recommended to aliquot it into different vials, so that only one is thawed at a time, with the remaining vials reserved for future analysis or result confirmation.

2.9 Potential of the Mediterranean Diet in increasing quality of life: + health + sustainability

The dietary habits of the Portuguese population are increasingly diverging from the traditional MD, particularly among younger individuals. This shift, influenced by age-related cultural factors, is largely due to a lack of awareness and society influences regarding the importance of healthy eating in promoting overall health and well-being [36, 42]. Among adults, over half of the population is overweighted, which significantly contributes to the high prevalence of major NCDs such as obesity and diabetes [36].

The DM4You project unites a consortium of several Portuguese partners with the aim to valorise local foods, highlighting its impacts at several levels, including health. Soups are recognised as a nutritious source of essential nutrients, including proteins, fibres, carbohydrates, vitamins, and minerals, as well as bioactive compounds and polyunsaturated fats (primarily from olive oil), which collectively support health and well-being, and appetite control. They also contribute to optimal intestinal function [87]. Additionally, soups are cost-effective, easy to prepare, and preserve, making them highly suitable for modern lifestyles. They also align with natural, ecological, and vegan dietary preferences (ecological footprint) [87]. As a low-calorie option that induces satiation, soup is an ideal meal choice for those pursuing weight loss strategies [88]. A single plate of soup contains 100 g of vegetables, meaning that consuming two plates of soup per day, in conjunction with three pieces of fruit, meets the World Health Organization recommendation of 400 g/day of fruits and vegetables [88, 89, 90].

To assess the impact of diet, the DM4You consortium proposed to monitor 80 healthy participants, both male and female, within two age groups (41-50 years and 65-75 years) over the course of one year, following a regime focused on soup and fruit consumption. The goal is to identify human protein biomarkers using a nanoscale LC coupled to tandem MS (nLC-MS/MS) in collaboration with the team at Luxembourg Institute of Science and Technology [38, 88]. Before proceeding with the project samples, it is crucial to optimise the sampling protocol, as well as the injection and subsequent data analysis processes. This optimisation includes selecting a suitable database that aligns with the objectives of this research.

3 Materials and Methods

3.1 Optimization of the protein quantification protocol

3.1.1 Samples

Plasma samples (Human Plasma Pooled Biowest S4180-100, from batch S1123454180) were used and identified by the letters A, B, C and D. These samples were used in previous works being stored in the laboratory at -20ºC.

3.1.2 Protein Quantification

3.1.2.1 Bradford Assay

The Bradford assay was conducted using 96-well microplates (astiK's). For each sample, 50 µL (diluted 1/2000) was mixed with 50 µL of a diluted HCl solution (0.1 mM, 1/8) and 150 µL of Bradford reagent (ROTH, Roti-Quant 5x Konzentrat) [50, 75]. The microplates were incubated for 5 minutes at room temperature, and absorbance was measured using a microplate reader (Spectra Max 190) at 595 nm. A Bovine Serum Albumin (BSA) solution was utilized to generate a calibration curve across a range of concentrations (0–20 µg) (Annex 1). Each sample and standard were analysed in triplicate.

3.1.2.2 BCA Assay

The BCA assay was performed using the Pierce BCA Protein Assay Kit (Thermo Scientific) following the manufacturer's protocol [76, 77]. For each sample, 25 µL (diluted 1/200) was mixed with 200 µL of the working reagent. In a 96-well plate, the samples were incubated at 37ºC for 30 minutes. After allowing the samples to cool to room temperature, absorbance was measured at 562 nm using a Spectra Max 190 microplate reader. A BSA solution served as the standard for generating a calibration curve across a concentration range of 0–30 µg (Annex 2). Each sample was analysed in triplicate.

3.1.2.3 Direct Absorbance

The absorbance of the diluted samples (1/2000) was measured at wavelengths of 230 nm, 260 nm, 280 nm, and 320 nm (NanoDrop Spectrophotometer ND-1000). Each sample was analysed in triplicate. Protein concentrations were subsequently calculated according to the equation (Equation 1 and 2, Section 2.7.1 Protein quantification) established by WC [78], and KB [79].

3.2 Optimization of the plasma extraction protocol

3.2.2 Samples

Blood samples were obtained from nine female volunteers, following approval from the NOVA-FCT Ethics Council. The participants were categorized into two age groups (20-30 years and 40-50 years) and their identities were anonymized throughout the study. A certified technician from SYNLAB Portugal performed the blood collection, using BD Vacutainer K2 EDTA tubes (containing 1.8 mg EDTA/mL of whole blood) as recommended by the International Society of Haematology [90].

3.2.1.1 Sample treatment

Following collection, the blood samples were kept on ice until processing. Each sample was initially divided into two aliquots (5 mL per tube), with protease and phosphatase inhibitors added to one set of tubes (as detailed in Table 4). Subsequently, all samples were further divided (2.5 mL per tube) and subjected to two different processing conditions: (A) centrifuged 10 minutes post-collection and (B) centrifuged 1-hour post-collection, while being maintained at 4°C. Centrifugation was conducted at 4°C for 15 minutes at 3000×*g* (3-16K, Sigma-Aldrich). Both the pellet and the supernatant (plasma) were flash-frozen in liquid nitrogen and stored at -80°C in different aliquots so that the same vial was only defrosted one time. The protocol is illustrated in Figure 10.

The effect of various protease and phosphatase inhibitors was also evaluated, including EDTA (an inhibitor of metalloproteases, which cleave peptide bonds in proteins), NaF and Na_3VO_4 (Sigma-Aldrich) (inhibitors of phosphatases, which remove phosphate groups from substrates by hydrolysing monophosphoric esters of phosphoric acid, affecting enzymes such as tyrosine phosphatases, alkaline phosphatases, and certain ATPases), and cOmplete EDTA-Free (Roche) (a protease inhibitor that prevents the cleavage of peptide bonds between amino acids in proteins) [91].

Table 4: **Protease and phosphatase inhibitors used in this work.**

Inhibitor	CAS number	Solution (M)	Volume added to each tube (µL)	Final volume (mL)
Complete® Mini EDTA-Free (protease inhibitor)	04693159001	One tablet diluted in 1.5 mL of 50 mM sodium phosphate pH 7 buffer	110	2.5
NaF (phosphatase inhibitor)	04693159001	0.1	5	2.5
Na_3VO_4 (phosphatase inhibitor)	13721-39-6	0.1	5	2.5

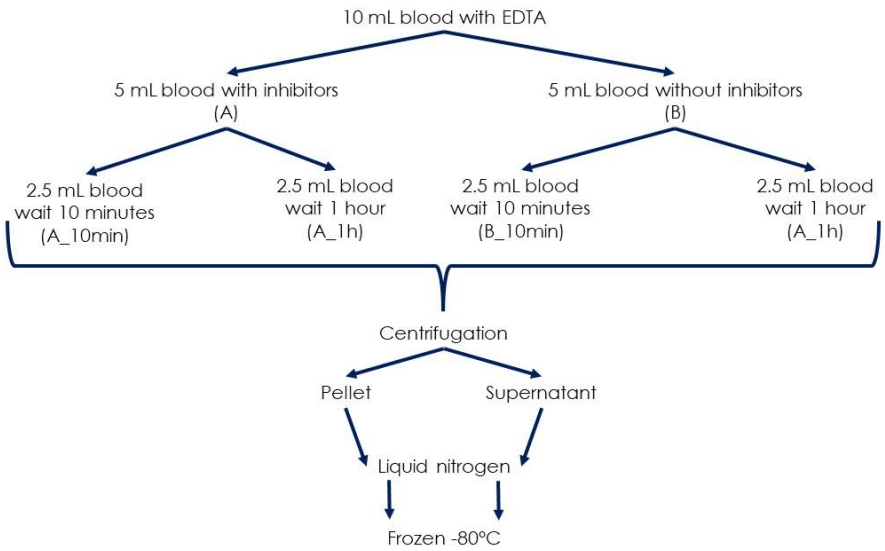


Figure 10: **Protocol used for collecting and processing blood in the pre-trial.** It is important to highlight that when starting that the sample is free of inhibitors, we are referring to the absence of protease and phosphatase inhibitors. All samples contain EDTA, which was present in the collection tube.

3.2.2 Protein quantification of samples

For the quantification of plasma samples, the Bradford and spectrophotometric quantification techniques were used following the protocols previously described in sections 3.1.2.1 of Materials and Methods after diluting the samples at 1/2500, as validated as suitable dilution based on the testing with commercial samples.

3.3 General analysis of the polypeptide electrophoresis pattern

Each sample (70 µg) was analysed by SDS-PAGE (in a continuous system (14×15 cm; 4% T, 3.3% C in the stacking gel; 10% T, 3.3% C in the running gel, acrylamide). The Precision Plus Protein™ Standard (BIO-RAD, 20 µL) marker was utilized, covering a molecular weight range of 10 to 250 kDa (Annex 3). Electrophoresis was conducted using the Hoefer SE600 Electrophoresis System at 190 V for 5 hours at 4°C. Following separation, the gel was initially stained with colloidal Coomassie G250 [92]. However, since only polypeptides in the 50 kDa region were clearly visible, the gel was subsequently stained with silver nitrate for enhanced detection [93].

3.4 Determination of total antioxidant capacities

3.4.1 Sample extraction

According to a protocol developed at the Thomas Roitschlad (Awby et. A. submitted) a single extraction procedure can be used for Folin-Ciocalteu and TEAC assay. In a sterile microtube, 250 µL of 100% methanol (Merck) was added to 100 µL of plasma in sterile tubes. The samples were incubated for 1 hour on a rotating shaker at 4°C. Subsequently, the samples were centrifuged at 3000×*g* for 5 minutes at 4°C (3-16K, Sigma-Aldrich), and the supernatant was collected and stored at -20°C. The procedure was repeated, pooled and the supernatant were stored at -20°C.

3.4.2 Folin-Ciocalteu

The reducing capacity of the samples was measured following Zarrouk et al. (2012) with modifications [82]. In a 96-well plate, 5 µL of the extract was added to 235 µL of water. Then, 15 µL of Folin-Ciocalteu reagent (Merck) and 45 µL of Na_2CO_3 solution (Sigma-Aldrich), 200 g/L, were added. The samples were incubated at 40°C for 30 minutes. The absorbance was measured at 765 nm using a microplate spectrophotometer (Spectra Max 190). A gallic acid (Sigma-Aldrich) solution served as the standard for

generating a calibration curve across a concentration range of 0–6 mM (Annex 4). Each sample was analysed in triplicate.

3.2.1 TEAC assay

The reducing capacity of the samples was measured also following the Wruss et al. (2015) with modifications [83]. A solution of ABTS (Sigma-Aldrich) with 8 mg/mL were prepared freshly by mixing with a solution of potassium persulfate (Sigma-Aldrich) 1.3 mg/mL to make final 1 mM ABTS stock solution prepared in darkness overnight (20ºC) and filtered through filter paper (Pore size of 20 µm from Fisherbrand™ Grade 122 Cellulose). It was diluted (1/10) in 50 mM sodium phosphate buffer solution pH 7.4 (Sigma-Aldrich) and considered the working solution. In each well of a 96-well plate, 2.5 µl sample extracts were mixed with 7.5 µl MeOH. Working solutions of 118 µl were added to each well. Plates were incubated at 6 min at room temperature and absorbance detected at 734 nm. An ascorbic acid (Sigma-Aldrich) solution served as the standard for generating a calibration curve across a concentration range of 0–5.5 mM. Each sample was analysed in triplicate.

3.3 Optimization of protein extraction

According to manufacturer's instructions, three PreOmics® preparation kits were utilised: iST kit (iST) designed for rapid and efficient protein sample preparation, streamlining the proteomics workflow, Enrich-iST kit specifically optimised for plasma and serum samples, enabling enrichment and improved detection of low-abundance proteins, (Enrich), and iST kit followed by iST Fractionation Add-on kit which allows for enhanced proteome depth through fractionation, enabling more comprehensive proteomic characterisation by increasing peptide coverage and resolution (PreOmics GmbH; Martinsried, Germany), for proteomic analysis by MS [94-96].

3.4 Optimization of running LC-MS/MS conditions

The extracted peptides (0.25 µg/µL) were first loaded onto a C18 pre-column (C18 PepMap, 5 µm, 5 mm × 300 µm, Thermo Scientific, Waltham, MA) for 10 minutes at a flow rate of 2 µL/min using a loading buffer consisting of 2% (v/v) acetonitrile and 0.05% (v/v) trifluoroacetic acid. Peptide separation was then performed on a C18 reverse-phase column (C18 PepMap 100, 3 µm, 100 Å, 75 µm × 15 cm, Thermo Scientific) using a linear gradient. Solvent A was 0.1% (v/v) formic acid, and solvent B was 0.1% (v/v) formic acid in acetonitrile. The flow rate was maintained at 300 nL/min. Peptides were eluted by increasing solvent B from 3% to 30% over 60 minutes, then raised to 40% over the next 10 minutes, and further to 80% for 5 minutes. The column was regenerated by washing at 80% B for 7 minutes and re-equilibrated to 3% B for 18 minutes.

LC-MS analysis was conducted using a NanoLC-425 Eksigent system coupled to a TripleTOF® 6600+ mass spectrometer. The acquired MS and MS/MS data were imported into the Progenesis QI for Proteomics software (version 4.2, Nonlinear Dynamics, Waters, Newcastle upon Tyne, UK). Protein and peptide identifications were performed by searching against Human plasma protein database on UniProtKB (616 731 sequences, downloaded on 29th May 2024) via Mascot Daemon (version 2.6.0, Matrix Science, London, UK). The Mascot search parameters included a peptide tolerance of 20 ppm, fragment mass tolerance of 0.02 Da, a maximum of two missed cleavages, carbamidomethylation of cysteine as a fixed modification, and oxidation of methionine, N-terminal protein acetylation, and tryptophan to kynurenine as variable modifications. Only proteins identified with a significance Mascot-calculated confidence of 95%, a minimum of two peptides per protein, at least one unique per protein, and an ANOVA p-value < 0.05 were retained.

3.5 Statistical analysis

Using the IBM SPSS Statistics 27, the kruskal.test (Kruskal-Wallis Rank Sum Test) from the stats package. When significant differences between groups were found, the post-hoc Dunn´s test was applied using IBM SPSS Statistics 27 [97].

Principal Component Analysis (PCA) in the results obtain in nLC-MS/MS was performed with IBM SPSS Statistics 27 [97].

4 Results and Discussion

4.1 Selection of the protein quantification protocol

For protein quantification, four methods were selected: the Bradford assay (Coomassie dye-binding method), the BCA assay, and two UV-Vis spectroscopy methods-WC and KB. Results were obtained with commercial plasma samples, and are depicted in Figure 11, highlighting that the value of protein concentration is method dependent.

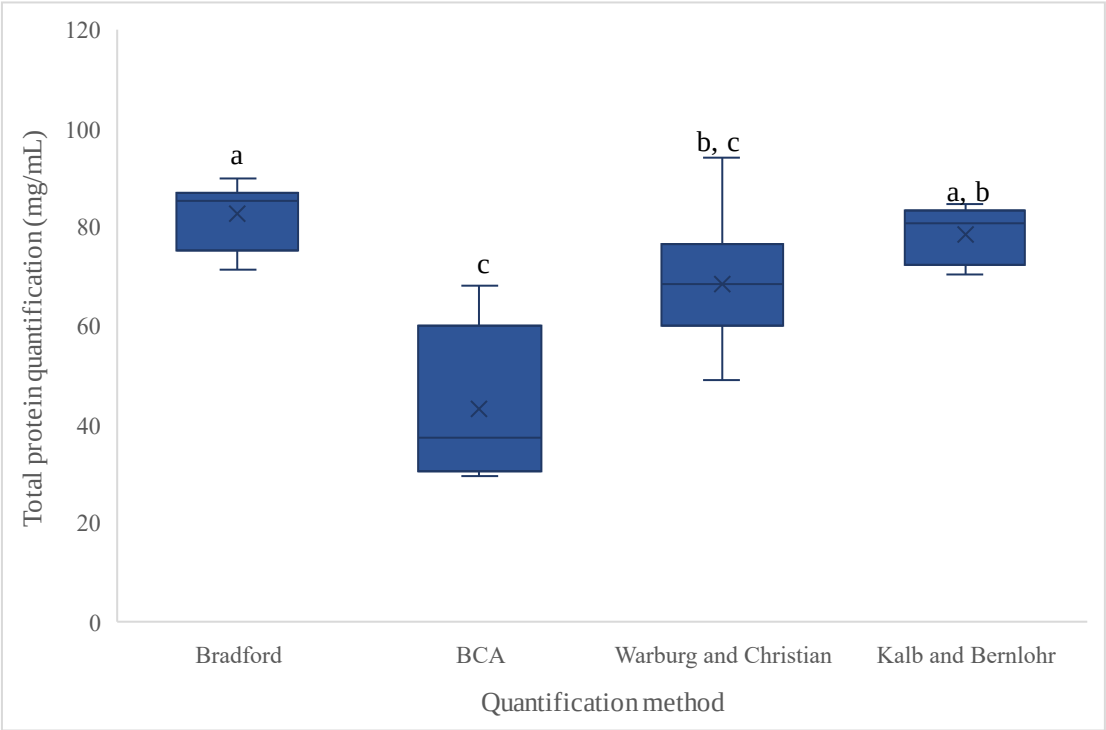


Figure 11: **Protein quantification of plasma samples using different quantification methods.** Significant results, as determined by Dunn's test, are denoted by lowercase letter ($p<0.05$). Values used to obtain this data are in Annex 6.

The total protein concentrations using the four different methods are overall low, comparing with the literature values of protein concentration in plasma (60-130 mg/mL). The commercial samples are in the lower range and information about protein concentration is not available from the manufacturer [98], particularly for the BCA method, which falls below the average protein levels typically observed in human plasma. This deviation can likely be attributed to the fact that these samples were stored at -20°C for more than 5 years, potentially leading to protein degradation/aggregation.

As anticipated, the quantification of the same plasma samples using different methods yielded different values. This discrepancy can be attributed to the inherent limitations of each method.

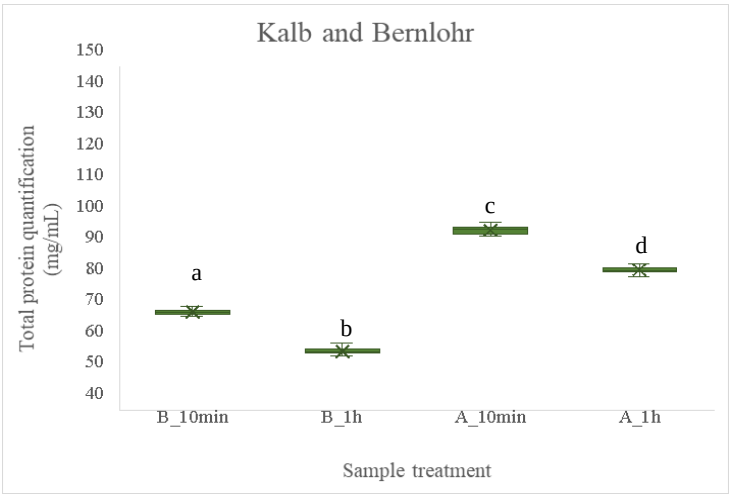
Regarding the reproducibility of these methods, the BCA assay exhibited the greatest variability within the same type of samples, indicating a lower consistency compared to the other methods. In contrast, the method described by KB demonstrated the highest reproducibility.

Given these findings, the Bradford method, along with the UV-Vis absorbance-based methods, were selected for further testing and use in subsequent assays. These methods were chosen based on their balance of sensitivity, reproducibility and for comparison with literature values as proteomics studies usually make use of Bradford as a protein quantification method.

4.2 Optimization of the plasma sampling protocol

The optimization test was conducted with nine volunteers divided into two age groups (20-30 years and 40-50 years). As the two age groups did not demonstrate statistically significant differences for the total soluble protein (Annexes 7-10), a single group was considered for the remaining comparisons.

As depicted in Figure 12, the protein values were within the typical range for total protein quantification in plasma (60-130 mg/mL). It's important to highlight that samples where not frozen for a long period of time and were only freeze-thawed one time, limiting degradation and denaturation [98]. Our data also show that processing time and the inclusion of several protease and phosphatase inhibitors allowed to maintain a higher protein concentration.



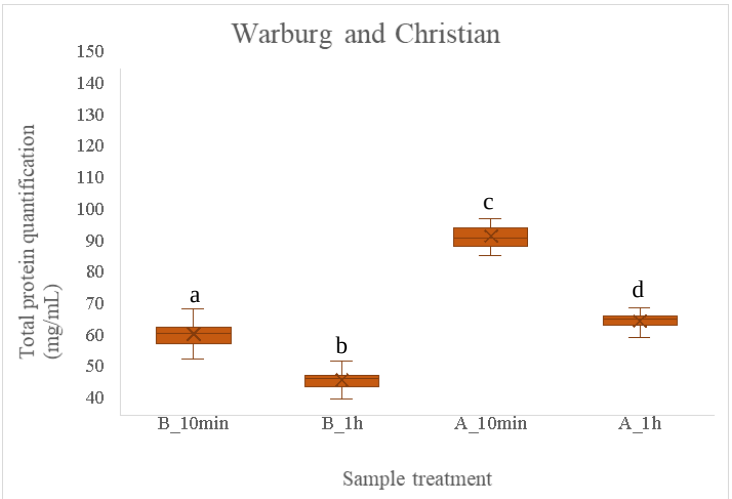


Figure 12: **Protein quantification of plasma samples using various methods**. The capital letter "A" indicates that the sample was treated with protease inhibitors, whereas "B" denotes that the samples were not. "10 min" refers to samples that were centrifuged 10 minutes after collection, while "1h" refers to those centrifuged 1-hour post-collection. Significant results, as determined by Dunn's test, are denoted by lowercase letter ($p<0.05$). The data underpinning these results are detailed in Annexes 6-10.

The findings revealed that concentration is method dependent. The Bradford method yielded higher total protein concentrations (69-180 mg/mL) compared to UV-Vis methods (59-101 mg/mL). Such variations may be attributed to the limitations of the quantification processes, as previously discussed. Across all methods, samples subjected to earlier centrifugation and treated with protease inhibitors showed higher protein concentrations. Conversely, samples that remained untreated for 1 hour and were not treated with protease/phosphatase inhibitors, exhibited lower protein concentration values. This indicates that both the handling time and the presence of protease inhibitors significantly impact the protein solubility of the samples, likely due to protein degradation/aggregation. It is also important to mention that, regardless of the method used, the relative impact of the treatments was maintained.

4.3 Protein plasma polypeptide profile

As different protein concentration values were observed for the sample collection procedure, extensive proteomic degradation may have occurred. To visualise the polypeptide profile of our samples, an SDS-PAGE was performed. The gel was silver stained as Coomassie staining only reveal the most abundant bands (40-60 kDa) (Figures 13 and 14). It is crucial to note that while silver nitrate staining provides enhanced sensitivity, allowing the detection of low-abundance proteins, it does not facilitate quantitative analysis of the samples [99]. This is a key disadvantage, as it limits the ability to compare polypeptide abundance between samples.

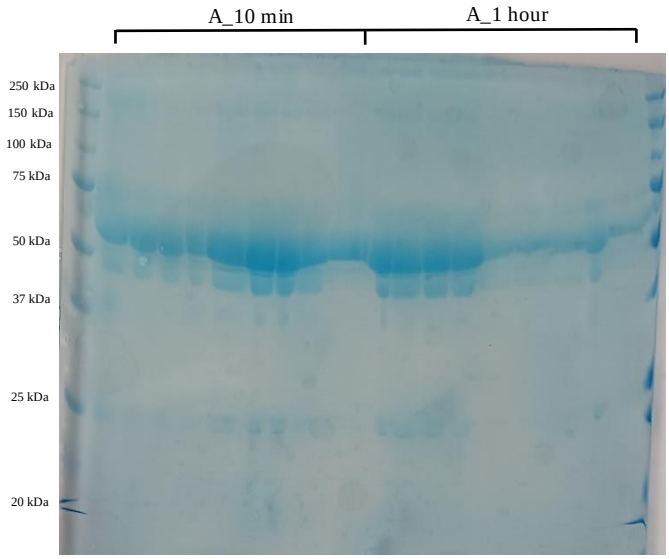


Figure 13: **Plasma protein samples analysis by SDS-PAGE stained with colloidal Coomassie G250**. 10% polyacrylamide gel. The Precision Plus ProteinTM Standards (BIO-RAD) marker was used with a range of 10 to 250 kDa. Samples were separated using the Electrophoresis System applying a current of 190 V for 5 hours. In this gel only represents samples with inhibitors, starting with sample 1 to 9 (form right to left). First the 10 minutes and then 1 hour.

In contrast, Coomassie staining, though less sensitive than silver nitrate, is widely used for quantitative purposes due to its ability to bind proteins stoichiometrically, thereby allowing the estimation of protein quantity [99].

Figure 14 represents the same gel stained with silver nitrate. Many protein bands are visible, particularly between 75 kDa and 25 kDa, where a large band can be observed. This highlights the dynamic range present in plasma, as various polypeptides are represented in this region. Consequently, this gel also demonstrates the complexity involved in analysing plasma.

The polypeptide band profile does not highlight evident protein degradation. Degraded proteins often appear as lower molecular weight bands or as smears at the bottom of the gel [99]. The absence of these features in the current analysis suggests that protein degradation did not significantly occur, or its molecular weight was not within the range of this gel. For example, fibrinogen is a protein present in plasma but is molecular weight is 340 kDa.

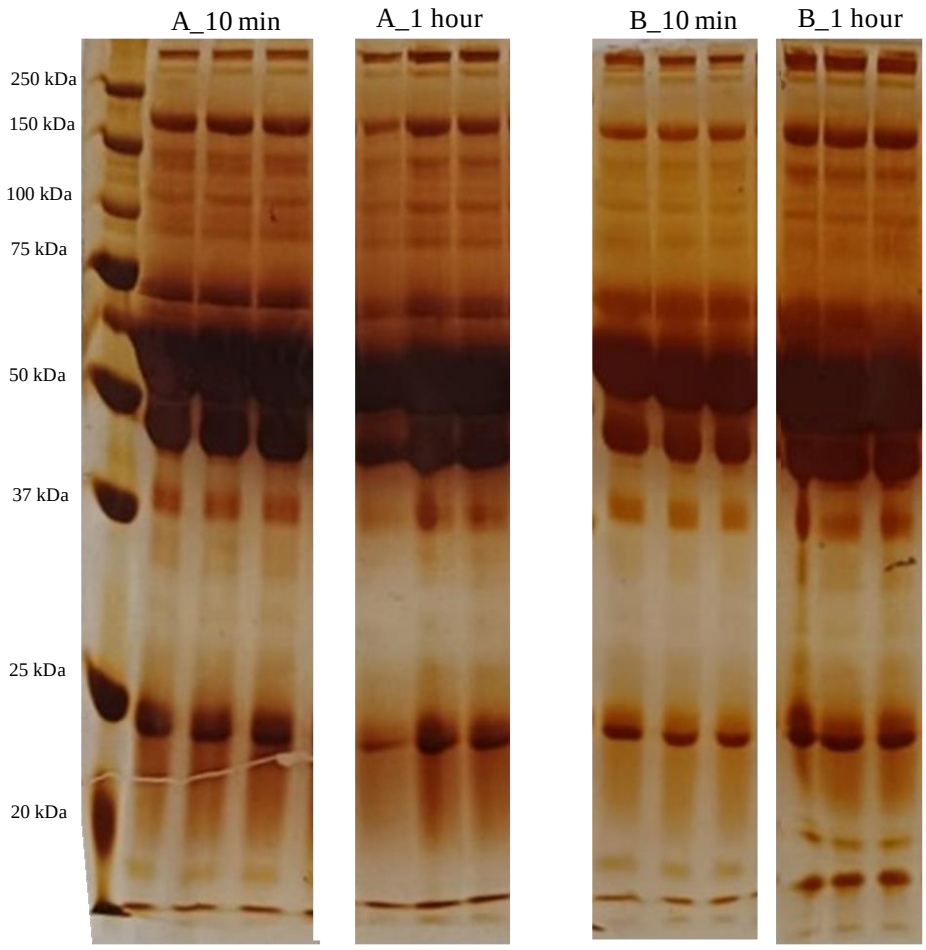


Figure 14: **Plasma protein samples analysis by SDS-PAGE stained with silver nitrate**. 10% polyacrylamide gel. The Precision Plus ProteinTM Standards (BIO-RAD) marker was used with a range of 10 to 250 kDa (Annex 3). Samples were separated applying a current of 190 V for 5 hours and the gel was stained with silver nitrate; 3 samples were subjected to different treatments: with inhibitors+10 minutes, with inhibitors+1 hour, without inhibitors+10 minutes, without inhibitors+1 hour (respectively).

4.4 Antioxidant capacity

In addition to proteomic analyses, and since one of the goals of DM4You is to compare age groups and study inflammation and aging, antioxidant capacities can serve as an indicator of these processes [80]. This is particularly relevant as certain dietary compounds possess properties that combat the inflammatory response [80].

To evaluate the antioxidant activity of the samples, a Folin-Ciocalteu assay and TEAC assay were conducted. It is important to note that this method provides an estimate of the sample's reducing capacity, serving as an approximation, much like the TEAC assay. In Figure 15, it can be observed that the concentration was higher when inhibitors were present in the samples. Conversely, this effect was not observed in the TEAC assay (Figure 16).

The antioxidant capacity of plasma is strongly influenced by an individual's diet. However, in this study, the volunteers' dietary intake was not controlled. As a result, we were unable to identify any diet-related differences and were limited to assessing the effects of the treatment. As shown in Figure 15, the samples treated with inhibitors displayed higher concentrations of gallic acid equivalents, which may suggest a protective effect of the inhibitors. Nonetheless, in the TEAC assay, no significant differences were observed between the various treatments. This can be explained by the fact that the Folin-Ciocalteu assay measures reducing compounds (primarily phenolics, but also other reducing agents), whereas the TEAC assay specifically assesses antioxidant activity in terms of radical scavenging capacity [80]. Different compounds within the sample may react differently to each assay, leading to variations in the results [80-83]. The discrepancy in methodology and sensitivity to specific compounds accounts for the differing outcomes often observed between these assays when applied to the same sample [80-83]. For instance, the Folin-Ciocalteu method does not detect vitamin E, which has antioxidant potential, while the TEAC assay can detect it [83].

In the upcoming DM4You study, it is anticipated that a controlled diet, particularly through the consumption of soups, will result in measurable differences in plasma antioxidant capacity throughout the duration of the study.

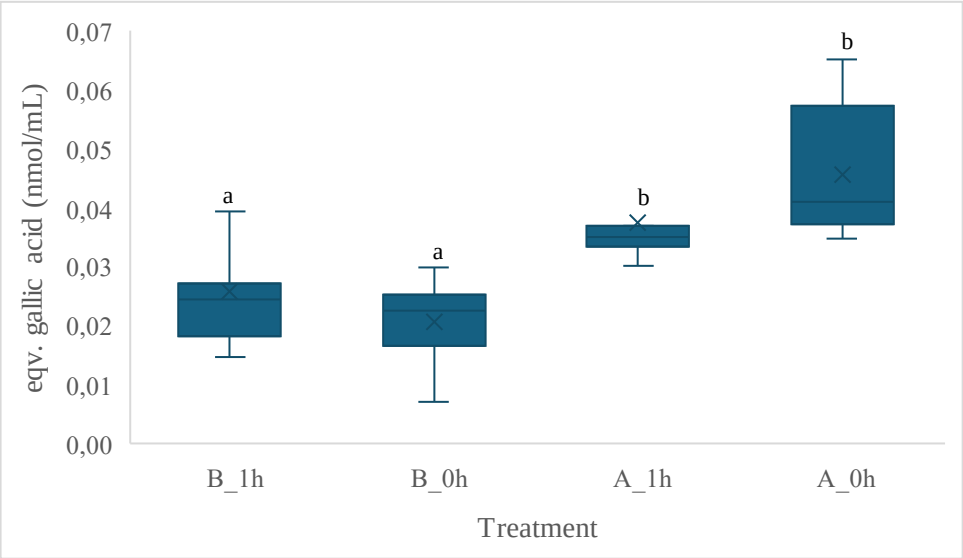


Figure 15: **Antioxidant capacity results by Folin-Ciocalteu.** Significant results, as determined by Dunn's test, are denoted by lowercase letter ($p<0.05$). The data supporting these results are provided in Annex 11.

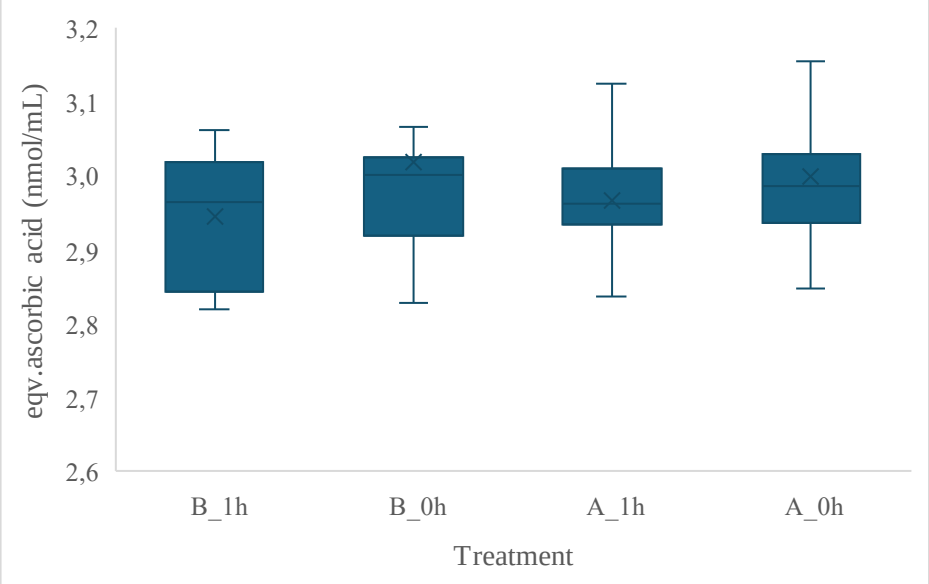


Figure 16: **Antioxidant capacity results by TEAC assay.** The data supporting these results are provided in Annex 14-12.

4.5 Characterization of the samples

For sample preparation, three kits from PreOmics, were selected. These protocols were chosen due to the reduction in sample handling time by four-fold, making the process faster and more cost-effective [94-96].

The protocol was optimised, regarding the amount of protein required for injection to obtain the most comprehensive spectrum and identification possible while minimising sample interference. This necessitated evaluating different dilutions to maximise protein detection.

At this stage, only two samples from individuals of different ages (individuals 7 and 8), were analysed, and a single run was conducted. This limited analysis was due to time and budget constraints.

4.5.1 Protein concentration

Sample protein quantity is a critical factor, as excessive protein, exceeding the maximum quality threshold of the equipment, can lead to distorted results. A non-diluted sample was injected, and in comparison, to diluted samples (0.05 µg/µL, 0.25 µg/µL, 0.5 µg/µL) (Figure 17 and 18), more peaks with higher intensities were obtained, but also more interference. This comparison clearly demonstrated the necessity to find an appropriate dilution to ensure column capacity and detection of proteins. Samples were quantified after the lysis step (Table 5) using the Bradford assay. Although we have demonstrated that the KB method provided more accurate and reproducible results, at Luxembourg Institute of Science and Technology (LIST), the routine procedure is to use Bradford method. The concentration values obtained were significantly lower compared to those measured previously (Figure 12) at FCT-NOVA after the collection. However, the proportions between treatments remain consistent.

Table 5: **Protein concentration after the lysis process carried out at LIST.** Quantification was performed using the Bradford method. Data represent the mean between the different preparations (n=3).

Sample	Concentration (µg/µL)	Standard deviation (±)
7B_10min	14.21	0.60
7B_1hour	8.26	0.05
7A_10min	28.75	1.17
7A_1hour	19.49	4.21
8B_10min	18.57	1.29
8B_1hour	8.47	0.06
8A_10min	27.19	0.75
8A_1hour	22.39	1.30

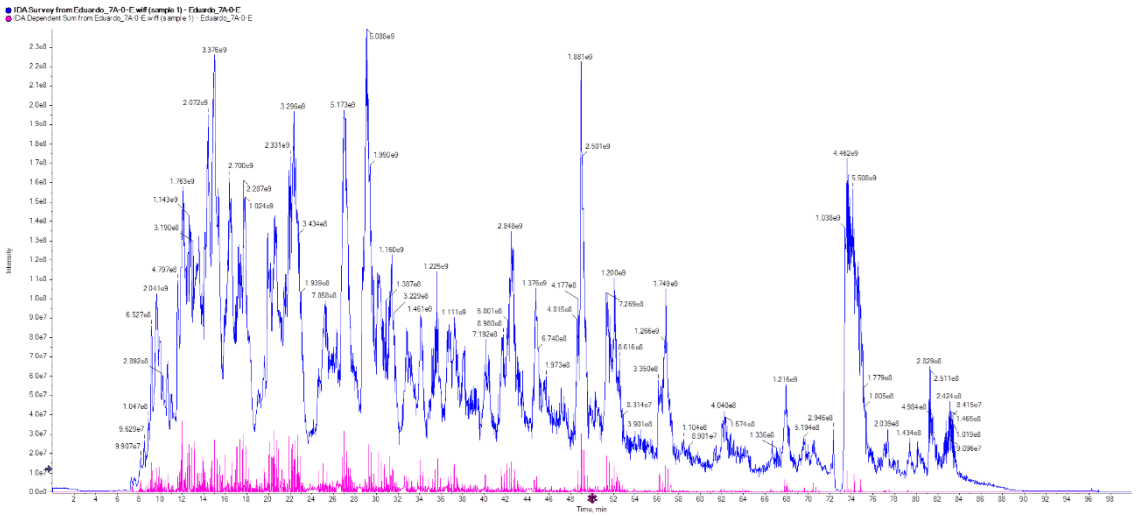


Figure 17: **Spectrum obtained from nLC-MS/MS analysis of a plasma sample.** Each peak represents a peptide group, which will be used to identify the corresponding proteins. In blue, the positive ionization is represented, and in pink, the negative ionization.

Sample dilution is crucial as it reduces matrix effects (the influence of co-eluting compounds from the sample matrix), leading to more accurate identification of the target analytes [99, 100]. This is especially important for plasma samples, that also contain other components (like lipids and metabolites) that can significantly influence the results.

Moreover, in nLC-MS/MS, columns are optimized for high sensitivity and separation efficiency but can easily become overloaded [101]. Overloading results in poor peak resolution, co-elution of analytes, and ultimately reduced sensitivity [85, 101]. Optimization of the protein concentration ensures that the sample concentration remains within the optimal range for the column capacity. Additionally, dilution helps to adjust the analyte concentration within the dynamic range of the mass spectrometer. If the concentration is too high, it can saturate the detector, resulting in inaccurate results [84, 85, 101].

Lastly, high concentrations of analytes or contaminants can lead to contamination or damage to the mass spectrometer, particularly the ion source and the detector [101]. The dilutions 0.05 µg/µL, 0.25 µg/µL, 0.5 µg/µL (Figure 18).

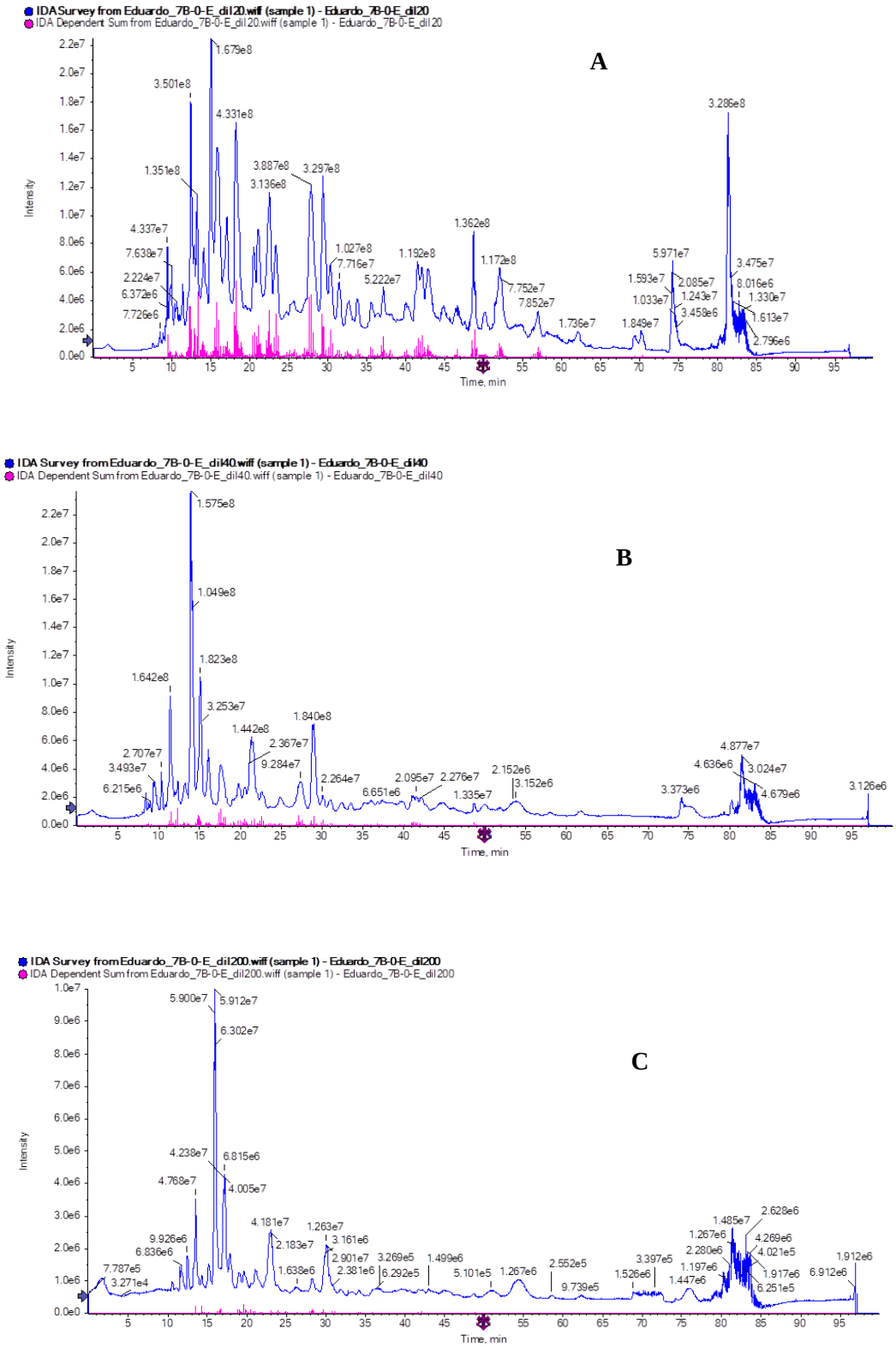


Figure 18: **Spectra obtained for the different dilutions used.** This figure displays the three sample dilutions in LC-Loading utilized for protocol optimization: 0.5 µg/µL (A), 0.25 µg/µL (B), 0.05 µg/µL (C). In blue, the positive ionization is represented, and in pink, the negative ionization.

As observed in the Figure 18, as the dilution increases, the intensity of the peaks diminishes. In Panel A, representing the least diluted sample (0.5 µg/µL), the intensity remains excessively high for the equipment, indicating potential overloading. Conversely, in Panel C (0.05 µg/µL), certain peaks disappear (notably in the 40–50 minutes range), suggesting that the dilution was leading to the loss of detectable peptide groups. Therefore, the dilution shown in Panel B (0.25 µg/µL) appears to be the most appropriate. Although some peptides may not be identifiable, this dilution allows for optimal sample analysis while avoiding the issues associated with the other dilutions.

4.5.2 Extraction method signature

Following the recommendation of LIST experts and to reduce technical variation, three kits from PreOmics were selected. The iST combined protein extraction, digestion, and peptide cleanup into a single protocol [94]. The ENRICH kit is specifically designed to enrich low-abundance proteins or specific protein classes from complex biological samples like plasma [95]. The kit includes steps for targeted enrichment, the specificities of the procedure not being disclosed [95]. The Add-on kit is used in conjunction with iST kit to enable the fractionation of peptide samples prior to MS analysis. Fractionation (procedure is not provided), but it is considering that enhances protein identification coverage by dividing the peptide mixture into multiple fractions, which are then analysed separately [96]. This step is particularly important for deep proteomic analysis as it increases the likelihood of detecting low-abundance peptides [96].

In our samples (all together), 405 features were detected with iST, 155 with ENRICH, and 396 with Add-on. A total of 350 proteins were identified across all three kits, after applying the filters (at least 2 peptides and at least 1 unique peptide). Complete list of proteins is listed in Annex 13 and 14. Figure 19 A illustrates the common protein identified by two or three methods, the different sample preparation kits significantly impacted protein identification. All three kits successfully identified 73 common proteins, including albumin, one of the most abundant proteins in plasma. The Add-on kit identified the highest number of proteins (253), followed by the iST kit (203), and the ENRICH kit (99).

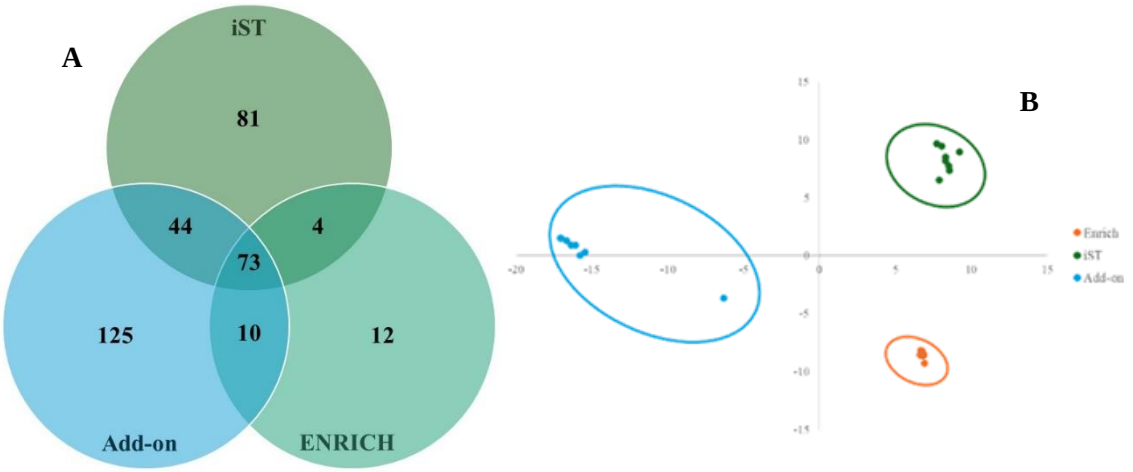


Figure 19: **Analysis of the identification results.** A*:* Venn diagram of proteins identified; data available in Annex 9. B: Statistical analysis using PCA for identified proteins. The average values from each kit and treatment were used in this analysis. Factor 1 accounts for 48% of the variance, while Factor 2 explains 20%. Each dot represents an individual with different treatment.

This outcome is particularly noteworthy, as a higher number of identified proteins was expected when compared with literature [85, 101], especially from the ENRICH kit, which was designed to enhance the detection of low-abundance proteins in the sample. Consequently, it would be advisable to repeat this procedure to validate these findings and determine if any errors during sample handling may have influenced the results.

The PCA (Figure 19 B) revealed a clear separation into three distinct clusters, corresponding to each sample preparation method. This separation underscored the distinct impact each kit has on the obtained protein profiles, further highlighting the importance of selecting an appropriate method for the specific objectives of the study.

The separation between the clusters of the three methods indicated that each method captures unique aspects of the protein composition in the samples. The distinct clustering highlights that these methods differ significantly, leading to the detection of different protein subsets [101]. This distinct separation implies that the choice of preparation method could be tailored based on the specific proteins or pathways of interest in the study [101]. Given that, the goal was to employ this procedure not only to detect new protein biomarkers but also to analyse those already documented in the literature, such as CRP. Thus, it is essential to examine the presence of these biomarkers in our results.

4.5.3 Protein Sources

In our dataset, 27% of the identified proteins were proteins from non-human organisms (Figure 20). Among these are microorganisms such as bacteria and viruses, as well as proteins from snakes and fish. A closer look on non-human protein identification, show that for most of them the confidence level of identification is relatively low (Annex 13-14). The confidence level indicates the likelihood that a particular peptide sequence corresponds to a specific ion detected in the MS/MS analysis, and it is typically expressed as a percentage or score reflecting the statistical reliability of the identification. At the highest confidence levels, predominantly human or mammalian proteins were observed, which aligns with the expected outcome, as the biological samples were derived from humans.

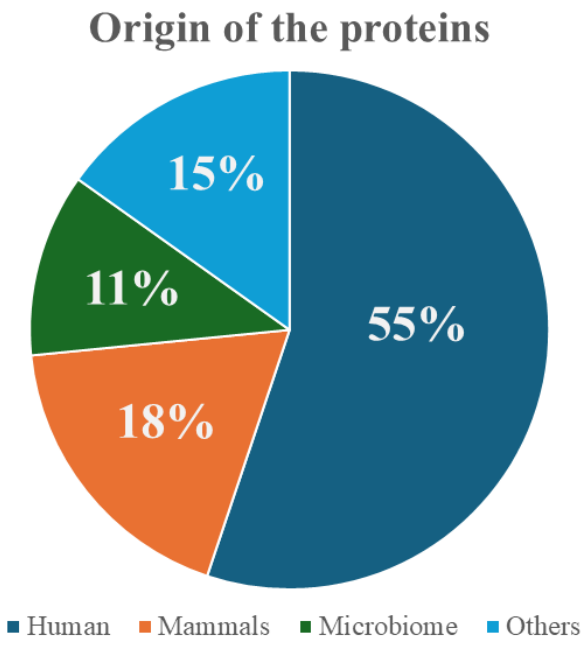


Figure 20: **Percentages of Identified Protein Sources.** Data supporting this chart can be found in Annex 9.

Given that the biological material analysed in this study is human plasma - a fluid responsible for the exchange and transport of molecules within the organism - and considering the limited number of samples analysed, it is plausible that the presence of non-human proteins may result from environmental exposure. Proteins from microorganisms or animals could enter the human bloodstream through bites, stings, vaccines or contact with environmental sources.

At intermediate confidence levels, proteins from microorganisms began to emerge. The presence of these microorganisms is referred to as the human microbiome, which encompasses the collective bacteria, fungi, viruses, and archaea residing in and on the human body. These microorganisms play essential roles in human health, influencing digestion, immune function, and metabolic processes [102, 103]. Proteins originating from the microbiome may enter the circulatory system via various mechanisms,

such as bacterial proteins translocation across the intestinal barrier, infection, or interaction with immune cells [102, 103]. In the group of identified proteins, 11% were found to originate from microorganisms within the human microbiome, with a notable presence of proteins from the digestive system microbiome, which were associated with intermediate confidence values.

Specific examples are, *Alistipes finegoldii* a member of the gut microbiota, associated with both health and disease. It participates in the fermentation of dietary fibres, producing short-chain fatty acids (SCFAs) like butyrate, which are essential for maintaining colon health [102, 103]. Similarly, *Alistipes onderdonkii* is involved in the fermentation of complex carbohydrates, contributing to SCFA production and supporting the integrity of the gut barrier [102, 103]. *Rubneribacter badeniensis* is a less well-characterised member of the gut microbiome, but it is thought to contribute to fibre degradation and SCFA production, although its precise metabolic functions are still under investigation [102, 103]. Members of the *Lachnospiraceae* family are prominent producers of SCFAs, such as butyrate, in the gut and play a crucial role in maintaining gut health, modulating immune responses, and reducing inflammation [102, 103].

Additionally, proteins from *Staphylococcus haemolyticus*, a commensal bacterium commonly found on human skin, were detected. *S. haemolyticus* is part of the skin microbiome but can act as an opportunistic pathogen, particularly in hospital settings [102, 103]. It is well-known for its ability to form biofilms and develop antibiotic resistance, making it a potential cause of infections, especially in immunocompromised individuals [102, 103]. The presence of protein of this bacterium in the bloodstream may be linked to its ability to translocate under certain conditions, particularly in cases of compromised immune barriers or invasive medical procedures, like collection of blood [102, 103].

4.5.4 Biomarkers analysis

An analysis was conducted to determine which of the identified proteins are already considered biomarkers, using data from the Human Protein Atlas [104] and the list of proteins previously highlighted as of interest for this project (Table 3), that is biomarkers related with MD. The outcomes of this analysis are summarized in Table 6. The Human Protein Atlas revealed that many of the identified proteins could indeed be considered biomarkers, particularly those related to inflammation and various pathologies.

Table 6: **Identified Biomarkers.** This table presents the total number of proteins identified by each kit, along with those classified as protein biomarkers according to the Human Protein Atlas. Additionally, it highlights how many of these proteins correspond to the list of diet-related biomarkers (Table 4). The data supporting this table can be found in Annex 13 and 14.

Kit	Total protein identification	Number of proteins describe as biomarkers (Human Protein Atlas)	In commom with Table 4
ENRICH	99	92	2
iST	203	65	4
Add-on	253	111	2

For the ENRICH kit, 92 proteins were classified as biomarkers (93% of the proteins identified), yet only two (Apolipoprotein E and Paraoxonase) were on the list of proteins of interest. This suggests that the enrichment process may have been effective, as only seven proteins identified were not biomarkers, which are typically less abundant in plasma. Therefore, this kit could be valuable for detecting new protein biomarkers. However, when compared to the specific list of proteins we aim to analyse, which are directly related to dietary factors, the results are not as favourable.

For the iST and Add-on kits, 65 and 111 proteins (32% and 43% of the proteins), respectively, were identified as biomarkers. Although the proportion is significantly lower than that observed with the ENRICH kit, these kits identified a greater number of proteins overall. Additionally, the iST kit identified four biomarkers of interest: Apolipoprotein E, Paraoxonase, Glutathione peroxidase, and CRP. Notably, for the Add-on kit, only two biomarkers from the list of protein of interest were identified, the same ones found using the ENRICH kit.

Apolipoprotein E is a protein primarily involved in the metabolism of fats in the body [51-56]. It plays a key role in the transport and clearance of cholesterol and triglycerides by binding to specific cell surface receptors, including the low-density lipoprotein receptor [51-56]. Apolipoprotein E is a biomarker for cardiovascular diseases and neurodegenerative conditions, particularly Alzheimer's disease [51-56].

Paraoxonase is an enzyme associated with high-density lipoproteins that protects low-density lipoprotein receptors from oxidative damage [51-56]. Paraoxonase activity is a biomarker for cardiovascular health and with increased oxidative stress and inflammation [51-56].

Glutathione peroxidase is an enzyme that protects cells from oxidative damage by reducing hydrogen peroxide and lipid peroxides to water and alcohol, respectively [51-56]. It is crucial for maintaining the oxidative balance within cells [51-56].

CRP is an acute-phase protein produced by the liver in response to inflammation. It binds to dead or dying cells and certain bacteria, activating the complement system, which plays a role in immune response and inflammation [51-56]. CRP is a widely used biomarker for systemic inflammation [51-56].

4.5.5 Proposed protocol for DM4You

Considering all the presented results, selecting the optimal protocol is not straightforward and involves balancing multiple factors. The ENRICH kit identified the highest proportion of biomarkers but simultaneously detected the fewest proteins overall among the kits. In contrast, while the Add-on kit yielded the largest number of identified proteins, it only associated two known biomarkers related to the diet. Additionally, it is important to note that using this kit, given the number of samples for the DM4You trial, would require 450 hours of MS analysis, a duration that is impractical considering the equipment's availability and the project timeline. In contrast, in the iST is only necessary 250 hours of MS analysis.

Therefore, the decision falls on the iST kit, which offers a medium-large protein abundance outcome in terms of protein identification. However, it stands out for identifying the most biomarkers relevant to this study (diet related), including CRP, a clinically well-known biomarker.

Furthermore, it was noted that this approach has the potential to support the detection and identification of proteins associated with the human microbiome, particularly the intestinal microbiome. By broadening the range of detectable proteins beyond those specific to *Homo sapiens*, this methodology may offer a means to explore, in the bloodstream, proteins absorbed via the mucosa as the ones derived from the gastric microbiome proteome. Consequently, this approach might provide insights into the microbiome itself, including potential alterations that could be influenced by dietary patterns observed during the DM4You project trial.

5. Conclusion

The DM4You project seeks to promote and safeguard the MD, with one of its main goals to provide data on how it enhances our quality of life. To this end, the project involves the analysis of plasma proteins, which provide comprehensive insights into the physiological state of the entire organism.

It was demonstrated that the methods for protein quantification with the highest reproducibility the method described by Karb and Bernlohr and Bradford. Considering plasma protein concentration, pre-test results further revealed no significant differences between the two age groups selected. However, it was observed that the method of blood sampling and subsequent sample treatment impacts the amount of protein that remains soluble. Hence, it is crucial to process samples quickly and maintain the sampling time as short as possible and to use protease and phosphate inhibitors to limit sample integrity alterations.

Furthermore, the samples exhibited varying antioxidant potential when quantified using two different methods. In the Folin-Ciocalteu assay, it was demonstrated that the introduction of inhibitors provided protection to our samples. However, this effect was not observed in the TEAC assay. This is due to the specific targets and specifications of each method. As these compounds are introduced through diet, in a controlled dietary study like DM4You, differences are to be expected.

Additionally, the study focused on optimizing three protein preparation methods for nLC-MS/MS. Although the number of proteins identified was lower than typically expected, the identified proteins were correlated with biomarkers relevant to this project. Consequently, the protocol utilizing the iST preparation kit, with a concentration of 0.25 µg/µL prior to injection into the equipment, will be followed.

The MD is distinguished by its abundance of fresh, high-quality foods, sourced primarily from the Mediterranean basin, as depicted in its dietary pyramid. The consumption of these foods is strongly linked to the prevention of numerous diseases, contributing to a lifespan that often surpasses the average, with significantly improved health outcomes. In conclusion, this work successfully optimized all the protocols intended for use in the DM4You project in the plasma analysis.

6. References

[1] Andrews, P., & Johnson, R. J. (2020). *Evolutionary basis for the human diet: consequences for human health*. Journal of Internal Medicine, 287(3), 226-237.
Doi: https://doi.org/10.1111/joim.13011

[2] Tattersall, I. (2009). *Becoming modern Homo sapiens. Evolution: Education and Outreach*, 2(4), 584-589.
Doi: https://doi.org/10.1007/s12052-009-0164-x

[3] Chen, P. J., & Antonelli, M. (2020). *Conceptual models of food choice: influential factors related to foods, individual differences, and society*. Foods, 9(12), 1898.
Doi: https://doi.org/10.3390/foods9121898

[4] Symons, M. (1994). *Simmel's gastronomic sociology: An overlooked essay*. Food and Foodways, 5(4), 333-351.
Doi: https://doi.org/10.1080/07409710.1994.9962016

[5] Hegarty, J. A., & O'Mahony, G. B. (2001). *Gastronomy: A phenomenon of cultural expressionism and an aesthetic for living*. International Journal of Hospitality Management, 20(1), 3-13.
Doi: https://doi.org/10.1016/S0278-4319(00)00028-1

[6] Carvache-Franco, M., Orden-Mejía, M., Carvache-Franco, W., Zambrano-Conforme, D., & Carvache-Franco, O. (2021). *Attributes of the service that influence and predict satisfaction in typical gastronomy*. International Journal of Gastronomy and Food Science, 24, 100356. Doi: https://doi.org/10.1016/j.ijgfs.2021.100356

[7] Smith, P., & Gregory, P. J. (2013). *Climate change and sustainable food production*. Proceedings of the nutrition society, 72(1), 21-28.
Doi: https://doi.org/10.1017/S0029665112002832

[8] Pérez-Rodrigo, C., & Aranceta-Bartrina, J. (2021*). Role of gastronomy and new technologies in shaping healthy diets*. In Gastronomy and food science (pp. 19-34). Academic Press.
Doi: https://doi.org/10.1016/B978-0-12-820057-5.00002-9

[9] Coveney, J., & Santich, B. (1997). *A question of balance: nutrition, health and gastronomy*. Appetite, 28(3), 267-277.
Doi: https://doi.org/10.1006/appe.1996.0083

[10] Gioia, C., Lucchino, B., Tarsitano, M. G., Iannuccelli, C., & Di Franco, M. (2020). *Dietary habits and nutrition in rheumatoid arthritis: can diet influence disease development and clinical manifestations?*. Nutrients, 12(5), 1456.
Doi: https://doi.org/10.3390/nu12051456

[11] Fresán U, Cvijanovic I and Chevance G. (2023*) You Can Help Fight Climate Change With Your Food Choices*. Frontiers Young Minds. 11:1004636.
Doi: 10.3389/frym.2023.1004636

[12] De Boer, J., De Witt, A., & Aiking, H. (2016*). Help the climate, change your diet: A cross-sectional study on how to involve consumers in a transition to a low-carbon society*. Appetite, 98, 19-27.
Doi: https://doi.org/10.1016/j.appet.2015.12.001

[13] Satia-Abouta, J., Kristal, A. R., Patterson, R. E., Littman, A. J., Stratton, K. L., & White, E. (2003). *Dietary supplement use and medical conditions: the VITAL study*. American Journal of Preventive Medicine, 24(1), 43-51.
Doi: https://doi.org/10.1016/S0749-3797(02)00571-8

[14] Caldeira, C., De Laurentiis, V., Ghose, A., Corrado, S., & Sala, S. (2021). *Grown and thrown: Exploring approaches to estimate food waste in EU countries*. Resources, Conservation and Recycling, 168, 105426.
Doi: https://doi.org/10.1016/j.envsci.2021.09.020

[15] *Bakırhan, H., Özkaya, V., & Pehlivan, M. (2023). Mediterranean diet is associated with better gastrointestinal health and quality of life, and less nutrient deficiency in children/adolescents with disabilities*. Frontiers in Public Health, 11.
Doi: https://doi.org/10.3389/fpubh.2023.1243513

[16] Lim, S. (2018). *Eating a balanced diet: a healthy life through a balanced diet in the age of longevity*. Journal of obesity & metabolic syndrome, 27(1), 39.
Doi: 10.7570/jomes.2018.27.1.3

[17] Bacon, S. L., Sherwood, A., Hinderliter, A., & Blumenthal, J. A. (2004). *Effects of exercise, diet and weight loss on high blood pressure*. Sports Medicine, 34, 307-316.
Doi: https://doi.org/10.2165/00007256-200434050-00003

[18] Haddy, F. J., Vanhoutte, P. M., & Feletou, M. (2006). *Role of potassium in regulating blood flow and blood pressure*. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 290(3), R546-R552.
Doi: https://doi.org/10.1152/ajpregu.00491.2005

[19] Azizi, S., Mahdavi, R., Vaghef-Mehrabany, E., Maleki, V., Karamzad, N., & Ebrahimi-Mameghani, M. (2020). *Potential roles of Citrulline and watermelon extract on metabolic and inflammatory variables in diabetes mellitus, current evidence and future directions: A systematic review*. Clinical and Experimental Pharmacology and Physiology, 47(2), 187-198.
Doi: https://doi.org/10.1111/1440-1681.13190

[20] Kennedy, E., Bowman, S. A., Spence; S., Freedman, M., & King, J. (2001). *Popular diets: correlation to health, nutrition, and obesity*. Journal of the American Dietetic Association, 101(4), 411-420.
Doi: https://doi.org/10.1016/S0002-8223(01)00108-0

[21] Costa-Camilo, E., Rovisco Pais, B., Duarte, I., Pinheiro, C., Carvalho, G. P. (2024). *Future-proof a mediterranean soup*. Proceedings of the 4th International Conference on Water Energy Food and Sustainability (ICoWEFS 2024), Springer Nature.
(unpublish)

[22] Graça, P. (2014). B*reve história do conceito de dieta Mediterrânica numa perspetiva de saúde*. Revista Factores de Risco, (31), 20-22.
Doi: http://hdl.handle.net/10400.26/10163

[23] Trichopoulou, A., & Lagiou, P. (1997). *Healthy traditional Mediterranean diet: an expression of culture, history, and lifestyle*. Nutrition Reviews, 55(11), 383-389.
Doi: https://doi.org/10.1111/j.1753-4887.1997.tb01578.x

[24] Hidayat, K., & Bimin, S. (2014). *Role of Mediterranean diet in prevention and management of type 2 diabetes*. Chinese Medical Journal, 127(20), 3651-3656.
Doi: 0.3760/cma.j.issn.0366-6999.20141358

[25] Costa-Camilo, E., Rovisco Pais, B., Duarte, I., Carvalho, G. P. (2024). *Dieta Mediterrânica: Uma dieta à prova dos desafios do futuro*. Livro de resumos do VII Simpósio "Produção e transformação de alimentos em ambiente sustentável, 119.
ISBN: 978-972-579-080-9

[26] Schröder, H. (2007). *Protective mechanisms of the Mediterranean diet in obesity and type 2 diabetes*. The Journal of Nutritional Biochemistry, 18(3), 149-160.
Doi: https://doi.org/10.1016/j.jnutbio.2006.05.006

[27] Menotti, A., & Puddu, P. E. (2015). *How the Seven Countries Study contributed to the definition and development of the Mediterranean diet concept: a 50-year journey*. Nutrition, metabolism and cardiovascular diseases, 25(3), 245-252.
Doi: https://doi.org/10.1016/j.numecd.2014.12.001

[28] Moro, E. (2016). *The Mediterranean diet from Ancel Keys to the UNESCO cultural heritage*. A pattern of sustainable development between myth and reality. Procedia-Social and Behavioral Sciences, 223, 655-661.
Doi: https://doi.org/10.1016/j.sbspro.2016.05.380

[29] Naska, A., & Trichopoulou, A. (2014). *Back to the future: the Mediterranean diet paradigm*. Nutrition, Metabolism and Cardiovascular Diseases, 24(3), 216-219.
Doi: https://doi.org/10.1016/j.numecd.2013.11.007

[30] Bach-Faig, A., Berry, E. M., Lairon, D., Reguant, J., Trichopoulou, A., Dernini, S., ... & Serra-Majem, L. (2011). *Mediterranean diet pyramid today*. Science and cultural updates. Public health nutrition, 14(12A), 2274-2284.
Doi: 10.1017/S1368980011002515

[31] Rico-Cabanas, L., García-Fernández, E. L., Estruchm, R., & Bach-Faig, A. (2014). *Mediterranean diet, the new pyramid and some insights on its cardiovascular preventive effect*. Rev Factores Risco, 31, 30-47.
ISSN: 1646-4834

[32] Willett, W. (2021). *Mediterranean dietary pyramid*. International journal of environmental research and public health, 18(9), 4568.
Doi: https://doi.org/10.3390/ijerph18094568

[33] Durão, C., Oliveira, J., & Almeida, M. D. V. D. (2008). *Portugal e o padrão alimentar mediterrânico*.
Doi: https://hdl.handle.net/10216/26328

[34] Davis, C., Bryan, J., Hodgson, J., & Murphy, K. (2015). *Definition of the Mediterranean diet: a literature review*. Nutrients, 7(11), 9139-9153.
Doi: https://doi.org/10.3390/nu7115459

[35] Dominguez, L. J., Di Bella, G., Veronese, N., & Barbagallo, M. (2021). *Impact of Mediterranean diet on chronic non-communicable diseases and longevity*. Nutrients, 13(6), 2028.
Doi: https://doi.org/10.3390/nu13062028

[36] Casas, R., Sacanella, E., & Estruch, R. (2014). *The immune protective effect of the Mediterranean diet against chronic low-grade inflammatory diseases*. Endocrine, Metabolic & Immune Disorders-Drug Targets (Formerly Current Drug Targets-Immune, Endocrine & Metabolic Disorders), 14(4), 245-254.
Doi: 2212-3873/14 $58.00+.00

[37] Panagiotakos, D. B., Pitsavos, C., Arvaniti, F., & Stefanadis, C. (2007). *Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore*. Preventive medicine, 44(4), 335-340.
Doi: https://doi.org/10.1016/j.ypmed.2006.12.009

[38] Martínez-González, M. Á., Corella, D., Salas-Salvadó, J., Ros, E., Covas, M. I., Fiol, M., ... & Estruch, R. (2012). *Cohort profile: design and methods of the PREDIMED study*. International journal of epidemiology, 41(2), 377-385.
Doi: https://doi.org/10.1093/ije/dyq250

[39] Ceriello, A., Esposito, K., La Sala, L., Pujadas, G., De Nigris, V., Testa, R., ... & Genovese, S. (2014). *The protective effect of the Mediterranean diet on endothelial resistance to GLP-1 in type 2 diabetes: a preliminary report*. Cardiovascular diabetology, 13(1), 1-9.
Doi: https://doi.org/10.1186/s12933-014-0140-9

[40] Bonaccio, M., Iacoviello, L., de Gaetano, G., & Moli-Sani Investigators. (2012). *The Mediterranean diet: the reasons for a success*. Thrombosis Research, 129(3), 401-404.
Doi: https://doi.org/10.1016/j.thromres.2011.10.018

[41] Vitale, M., Masulli, M., Calabrese, I., Rivellese, A. A., Bonora, E., Signorini, S., ... & TOSCA. IT Study Group. (2018). *Impact of a Mediterranean dietary pattern and its components on cardiovascular risk factors, glucose control, and body weight in people with type 2 diabetes: a real-life study*. Nutrients, 10(8), 1067.
Doi: https://doi.org/10.3390/nu10081067

[42] Castillo, D. J., Rifkin, R. F., Cowan, D. A., & Potgieter, M. (2019). *The healthy human blood microbiome: fact or fiction?*. Frontiers in cellular and infection microbiology, 9, 449041.
Doi: https://doi.org/10.3389/fcimb.2019.00148

[43] Ford, J. (2013). *Red blood cell morphology*. International journal of laboratory hematology, 35(3), 351-357.
Doi: https://doi.org/10.1111/ijlh.12082

[44] Frenette, P. S., & Wagner, D. D. (1996). *Adhesion molecules—Blood vessels and blood cells*. New England Journal of Medicine, 335(1), 43-45.
Doi: 10.1056/NEJM199607043350108

[45] Hartenstein, V. (2006). *Blood cells and blood cell development in the animal kingdom*. Annu. Rev. Cell Dev. Biol., 22, 677-712.
Doi: https://doi.org/10.1146/annurev.cellbio.22.010605.093317

[46] Li, J., Guasch-Ferré, M., Chung, W., Ruiz-Canela, M., Toledo, E., Corella, D., ... & Liang, L. (2020). *The Mediterranean diet, plasma metabolome, and cardiovascular disease risk*. European heart journal, 41(28), 2645-2656.
Doi: https://doi.org/10.1093/eurheartj/ehaa209

[47] García-Gutiérrez, M. S., Navarrete, F., Sala, F., Gasparyan, A., Austrich-Olivares, A., & Manzanares, J. (2020). *Biomarkers in psychiatry: concept, definition, types and relevance to the clinical reality*. Frontiers in psychiatry, 11, 432.
Doi: https://doi.org/10.3389/fpsyt.2020.00432

[48] Geyer, P. E., Holdt, L. M., Teupser, D., & Mann, M. (2017). *Revisiting biomarker discovery by plasma proteomics*. Molecular systems biology, 13(9), 942.
Doi: 10.15252/msb.20156297

[49] Clarivate Analytics. (2024). *Web of Science: Science Citation Index Expanded (Sci-EXPANDED)*.
URL: https://www.webofscience.com

[50] Lopez-Garcia, E., Schulze, M. B., Meigs, J. B., Manson, J. E., Rifai, N., Stampfer, M. J., ... & Hu, F. B. (2005). *Consumption of trans fatty acids is related to plasma biomarkers of inflammation and endothelial dysfunction*. The Journal of nutrition, 135(3), 562-566.
Doi: https://doi.org/10.1093/jn/135.3.562

[51] Lee, D. H., de Rezende, L. F. M., Eluf-Neto, J., Wu, K., Tabung, F. K., & Giovannucci, E. L. (2019). *Association of type and intensity of physical activity with plasma biomarkers of inflammation and insulin response*. International journal of cancer, 145(2), 360-369.
Doi: https://doi.org/10.1002/ijc.32111

[52] Lee, Bach-Faig, A., Geleva, D., Carrasco, J. L., Ribas-Barba, L., & Serra-Majem, L. (2006). *Evaluating associations between Mediterranean diet adherence indexes and biomarkers of diet and disease*. Public Health Nutrition, 9(8A), 1110-1117.
Doi: 10.1017/S1368980007668499

[53] Hidalgo-Liberona, N., Meroño, T., Zamora-Ros, R., Rabassa, M., Semba, R., Tanaka, T., ... & Cherubini, A. (2021). *Adherence to the Mediterranean diet assessed by a novel dietary biomarker score and mortality in older adults: the InCHIANTI cohort study*. BMC medicine, 19, 1-13. Doi: https://doi.org/10.1186/s12916-021-02154-7

[54] Tuttolomondo, A., Simonetta, I., Daidone, M., Mogavero, A., Ortello, A., & Pinto, A. (2019). *Metabolic and vascular effect of the Mediterranean diet*. International journal of molecular sciences, 20(19), 4716. Doi: https://doi.org/10.3390/ijms20194716

[55] Barbosa, K. B., Bressan, J., Zulet, M. A., & JA, M. H. (2008, September). *Influence of dietary intake on plasma biomarkers of oxidative stress in humans*. In Anales del sistema sanitario de Navarra (Vol. 31, No. 3, pp. 259-280). Doi: https://doi.org/10.4321/s1137-66272008000500006

[56] te Pas, M. F., Koopmans, S. J., Kruijt, L., Calus, M. P., & Smits, M. A. (2013). *Plasma proteome profiles associated with diet-induced metabolic syndrome and the early onset of metabolic syndrome in a pig model*. PLoS One, 8(9), e73087. Doi: https://doi.org/10.1371/journal.pone.0073087

[57] Monfort-Pires, M., & Ferreira, S. R. G. (2017). *Inflammatory and metabolic responses to dietary intervention differ among individuals at distinct cardiometabolic risk levels*. Nutrition, 33, 331-337. Doi: https://doi.org/10.1016/j.nut.2016.07.021

[58] Duggan, M. R., Butler, L., Peng, Z., Daya, G. N., Moghekar, A., An, Y., ... & Walker, K. A. (2023). *Plasma proteins related to inflammatory diet predict future cognitive impairment*. Molecular psychiatry, 28(4), 1599-1609. Doi: https://doi.org/10.1038/s41380-023-01975-7

[59] Dewell, A., Tsao, P., Rigdon, J., & Gardner, C. D. (2018). *Antioxidants from diet or supplements do not alter inflammatory markers in adults with cardiovascular disease risk*. A pilot randomized controlled trial. Nutrition Research, 50, 63-72. Doi: https://doi.org/10.1016/j.nutres.2017.10.017

[60] Ordovas, J. (2007). *Diet/genetic interactions and their effects on inflammatory markers*. Nutrition reviews, 65(suppl_3), S203-S207. Doi: https://doi.org/10.1111/j.1753-4887.2007.tb00363.x

[61] Hong, B. V., Agus, J. K., Tang, X., Zheng, J. J., Romo, E. Z., Lei, S., & Zivkovic, A. M. (2023). *Precision nutrition and cardiovascular disease risk reduction: The promise of high-density lipoproteins*. Current Atherosclerosis Reports, 25(10), 663-677. Doi: https://doi.org/10.1007/s11883-023-01148-5

[62] Whalen, K. A., McCullough, M. L., Flanders, W. D., Hartman, T. J., Judd, S., & Bostick, R. M. (2016). *Paleolithic and Mediterranean diet pattern scores are inversely associated with biomarkers of inflammation and oxidative balance in adults*. The Journal of nutrition, 146(6), 1217-1226. Doi: https://doi.org/10.3945/jn.115.224048

[63] Wang, F., Tessier, A. J., Liang, L., Wittenbecher, C., Haslam, D. E., Fernández-Duval, G., ... & Guasch-Ferré, M. (2023). *Plasma metabolomic profiles associated with mortality and longevity in a prospective analysis of 13,512 individuals*. Nature Communications, 14(1), 5744. Doi: https://doi.org/10.1038/s41467-023-41515-z

[64] Gaskins, A. J., Rovner, A. J., Mumford, S. L., Yeung, E., Browne, R. W., Trevisan, M., ... & BioCycle Study Group. (2010). *Adherence to a Mediterranean diet and plasma concentrations of lipid peroxidation in premenopausal women*. The American journal of clinical nutrition, 92(6), 1461-1467. Doi: https://doi.org/10.3945/ajcn.110.000026

[65] Tanaka, T., Talegawkar, S. A., Jin, Y., Candia, J., Tian, Q., Moaddel, R., ... & Ferrucci, L. (2022). *Metabolomic profile of different dietary patterns and their association with frailty index in community-dwelling older men and women*. Nutrients, 14(11), 2237.
Doi: https://doi.org/10.3390/nu14112237

[66] Carruba, G., Cocciadiferro, L., Di Cristina, A., Granata, O. M., Dolcemascolo, C., Campisi, I., ... & Traina, A. (2016). *Nutrition, aging and cancer: lessons from dietary intervention studies*. Immunity & Ageing, 13, 1-9.
Doi: https://doi.org/10.1186/s12979-016-0069-9

[67] Nedelkov, D., Kiernan, U. A., Niederkofler, E. E., Tubbs, K. A., & Nelson, R. W. (2005). *Investigating diversity in human plasma proteins*. Proceedings of the National Academy of Sciences, 102(31), 10852-10857.
Doi: https://doi.org/10.1073/pnas.050042610

[68] Anderson, N. L., & Anderson, N. G. (2002). *The human plasma proteome: history, character, and diagnostic prospects. Molecular & cellular proteomics*, 1(11), 845-867.
Doi: https://doi.org/10.1074/mcp.R200007-MCP200

[69] Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D., & Mann, M. (2016). *Plasma proteome profiling to assess human health and disease*. Cell systems, 2(3), 185-195.
Doi: https://doi.org/10.1016/j.cels.2016.02.015

[70] Foster, J. F. (1960). *Plasma albumin*. In The plasma proteins (pp. 179-239). Academic Press.
Doi: https://doi.org/10.1016/B978-1-4832-2810-5.50013-X

[71] Mosesson, M. W., & Umfleet, R. A. (1970). *The cold-insoluble globulin of human plasma: I. Purification, primary characterization, and relationship to fibrinogen and other cold-insoluble fraction components*. Journal of Biological Chemistry, 245(21), 5728-5736.
Doi: https://doi.org/10.1016/S0021-9258(18)62713-9

[72] Ma, C., Li, Y., Li, J., Song, L., Chen, L., Zhao, N., ... & Yuan, X. (2023). *Comprehensive and deep profiling of the plasma proteome with protein corona on zeolite NaY*. Journal of Pharmaceutical Analysis, 13(5), 503-513.
Doi: https://doi.org/10.1016/j.jpha.2023.04.002

[73] Pan, S., Aebersold, R., Chen, R., Rush, J., Goodlett, D. R., McIntosh, M. W., ... & Brentnall, T. A. (2009). *Mass spectrometry based targeted protein quantification: methods and applications*. Journal of proteome research, 8(2), 787-797.
Doi: https://doi.org/10.1021/pr800538n

[74] Bradford, M. M. (1976). *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Analytical biochemistry, 72(1-2), 248-254.
Doi: https://doi.org/10.1016/0003-2697(76)90527-3

[75] Kruger, N. J. (2009). *The Bradford method for protein quantitation*. The protein protocols handbook, 17-24.
Doi: 10.1007/978-1-59745-198-7_4

[76] Sapan, C. V., Lundblad, R. L., & Price, N. C. (1999). *Colorimetric protein assay techniques*. Biotechnology and applied Biochemistry, 29(2), 99-108.
Doi: https://doi.org/10.1111/j.1470-8744.1999.tb00538.x

[77] Smith, P. E., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M., ... & Klenk, D. C. (1985). *Measurement of protein using bicinchoninic acid*. Analytical biochemistry, 150(1), 76-85.
Doi: https://doi.org/10.1016/0003-2697(85)90442-7

[78] Warburg, 0.. and Christian, W. (1942) B&hem. 2. 310, 384-421.

[79] Kalb Jr, V. F., & Bernlohr, R. W. (1977). *A new spectrophotometric assay for protein in cell extracts*. Analytical biochemistry, 82(2), 362-371.
Doi: https://doi.org/10.1016/0003-2697(77)90173-7

[80] Sawka, A. M., Jaeschke, R., Singh, R. J., & Young Jr, W. F. (2003). *A comparison of biochemical tests for pheochromocytoma: measurement of fractionated plasma metanephrines compared with the combination of 24-hour urinary metanephrines and catecholamines*. The Journal of Clinical Endocrinology & Metabolism, 88(2), 553-558.
Doi: https://doi.org/10.1210/jc.2002-021251

[81] Okutucu, B., Dınçer, A., Habib, Ö., & Zıhnıoglu, F. (2007). *Comparison of five methods for determination of total plasma protein concentration*. Journal of biochemical and biophysical methods, 70(5), 709-711.
Doi: https://doi.org/10.1016/j.jbbm.2007.05.009

[82] Zarrouk, O., Francisco, R., Pinto-Marijuan, M., Brossa, R., Santos, R. R., Pinheiro, C., et al. (2012). *Impact of irrigation regime on berry development and flavonoids composition in Aragonez (Syn. Tempranillo) grapevine*. Agric. Water Manage. 114, 18–29.
Doi: https://doi.org/10.1016/j.agwat.2012.06.018

[83] Wang, C. C., Chu, C. Y., Chu, K. O., Choy, K. W., Khaw, K. S., Rogers, M. S., & Pang, C. P. (2004). *Trolox-equivalent antioxidant capacity assay versus oxygen radical absorbance capacity assay in plasma*. Clinical Chemistry, 50(5), 952-954.
Doi: https://doi.org/10.1373/clinchem.2004.031526

[84] Lai, X., & Witzmann, F. A. (2011). *Gel-based and gel-free sample preparation for LC-MS/MS analysis*. Sample Preparation in Biological Mass Spectrometry, 3-17.
Doi: https://doi.org/10.1007/978-94-007-0828-0_1

[85] Gaspari, M., & Cuda, G. (2011). *Nano LC–MS/MS: A robust setup for proteomic analysis*. Nanoproteomics: methods and protocols, 115-126.
Doi: https://doi.org/10.1007/978-1-61779-319-6_9

[86] Kapp, M. B. (2006). *Ethical and legal issues in research involving human subjects: do you want a piece of me?*. Journal of clinical pathology, 59(4), 335-339.
Doi: https://doi.org/10.1136/jcp.2005.030957

[87] Román, G. C., Jackson, R. E., Gadhia, R., Román, A. N., & Reis, J. (2019). *Mediterranean diet: The role of long-chain ω-3 fatty acids in fish; polyphenols in fruits, vegetables, cereals, coffee, tea, cacao and wine; probiotics and vitamins in prevention of stroke, age-related cognitive decline, and Alzheimer disease*. Revue neurologique, 175(10), 724-741.
Doi: https://doi.org/10.1016/j.neurol.2019.08.005

[88] Fernández-López, J., Botella-Martínez, C., Navarro-Rodríguez de Vera, C., Sayas-Barberá, M. E., Viuda-Martos, M., Sánchez-Zapata, E., & Pérez-Álvarez, J. A. (2020). *Vegetable soups and creams: Raw materials, processing, health benefits, and innovation trends*. Plants, 9(12), 1769.
Doi: https://doi.org/10.3390/plants9121769

[89] World Health Organization. 2013. Health 2020: A European policy framework and strategy for the 21st century. World Health Organization. Regional Office for Europe

[90] World Health Organization. 2021. Progress on the prevention and control of noncommunicable diseases in the Western Pacific Region: country capacity survey 2019.

[91] Rodeghiero, F., Michel, M., Gernsheimer, T., Ruggeri, M., Blanchette, V., Bussel, J. B., ... & Stasi, R. (2013). *Standardization of bleeding assessment in immune thrombocytopenia: report from the International Working Group*. Blood, The Journal of the American Society of Hematology, 121(14), 2596-2606.
Doi: https://doi.org/10.1182/blood-2012-07-442392

[92] Dyballa, N., & Metzger, S. (2009). *Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels*. Journal of Visualized Experiments, (30), e1431.
Doi: 10.3791/1431

[93] Poehling, H. M., & Neuhoff, V. (1981). *Visualization of proteins with a silver "stain": a critical analysis*. Electrophoresis, 2(3), 141-147.
Doi: https://doi.org/10.1002/elps.1150020304

[94] Geyer, P., Pichler, G., Treit, P., Doll, S., Kulak, N., & Transduction, S. (n.d.). *Application Note Clinical Research Automated Plasma Proteomic Sample Preparation Using PreOmics iST Technologies and the Agilent Bravo Automated Liquid Handling Platform 2*. Retrieved August 23, 2024 https://cdn.prod.websitefiles.com/6322f21c2bfa08926805e0fe/6322f21c2bfa080f4205e3b6_PreOmics_Agilent_Automated_plasma_sample_preparation.pdf

[95] Optimized biomarker discovery through advanced plasma proteomic pipeline. (n.d.). Retrieved May 23, 2024 https://cdn.prod.websitefiles.com/6322f21c2bfa08926805e0fe/6657170ce45da9cf7aec7bd7_ENRICH_Optimized%20biomarker%20discovery%20through%20advance%20plasma%20proteomic%20pipeline.pdf

[96] Peptide Fractionation Simplified (n.d.). Retrieved May 23, 2024 https://cdn.prod.websitefiles.com/6322f21c2bfa08926805e0fe/6322f21c2bfa080ffd05e6c2_Technical%20note_iST-Fractionation%20Add-on.pdf

[97] Marôco, J. (2018). *Análise Estatística com o SPSS Statistics*. 8ª edição. ReportNumber, Lda.
ISBN: 9789899676374

[98] Leeman, M., Choi, J., Hansson, S., Storm, M. U., & Nilsson, L. (2018). *Proteins and antibodies in serum, plasma, and whole blood—size characterization using asymmetrical flow field-flow fractionation (AF4)*. Analytical and bioanalytical chemistry, 410, 4867-4873.
Doi: 10.1007/s00216-018-1127-2

[99] Gallagher, S., & Sasse, J. (1998). *Protein analysis by SDS-PAGE and detection by Coomassie blue or silver staining*. Current Protocols in Pharmacology, 2(1), A-3B.
Doi: https://doi.org/10.1002/0471141755.pha03bs02

[100] Grebe, S. K., & Singh, R. J. (2011). *LC-MS/MS in the clinical laboratory–where to from here?*. *The Clinical biochemist reviews*, *32*(1), 5.
Doi: PMC3052391

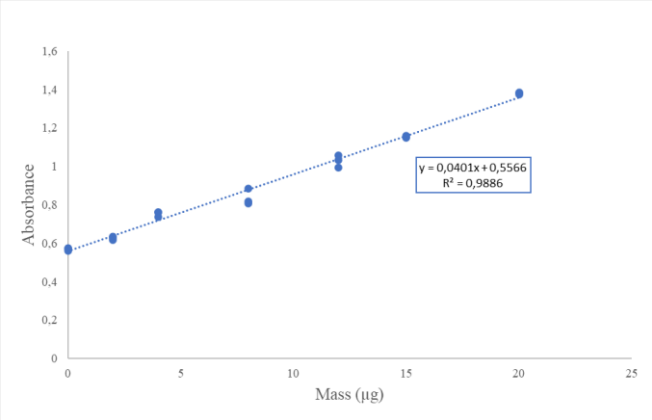
[101] Maćkiewicz, A., & Ratajczak, W. (1993). *Principal components analysis (PCA)*. Computers & Geosciences, 19(3), 303-342.
Doi: https://doi.org/10.1016/0098-3004(93)90090-R

[102] Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). *Current understanding of the human microbiome*. Nature medicine, 24(4), 392-400.
Doi: https://doi.org/10.1038/nm.4517

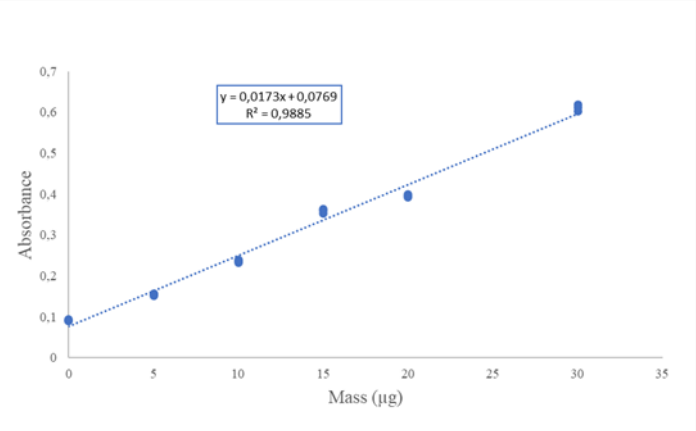
[103] Thomas, S. (Ed.1). (2022). *Human Microbiome: Clinical Implications and Therapeutic Interventions*. Springer Nature.
ISBN: 978-981-16-7672-7

[104] Pontén, F., Schwenk, J. M., Asplund, A., & Edqvist, P. H. (2011). *The Human Protein Atlas as a proteomic resource for biomarker discovery*. Journal of internal medicine, 270(5), 428-446.
Doi: https://doi.org/10.1111/j.1365-2796.2011.02427.x

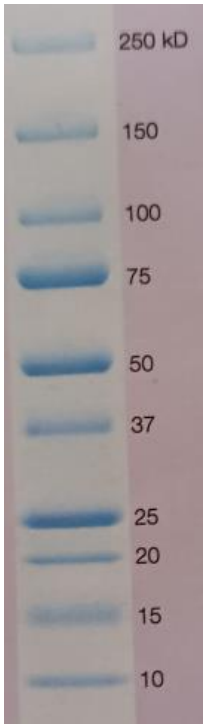
7. Annex



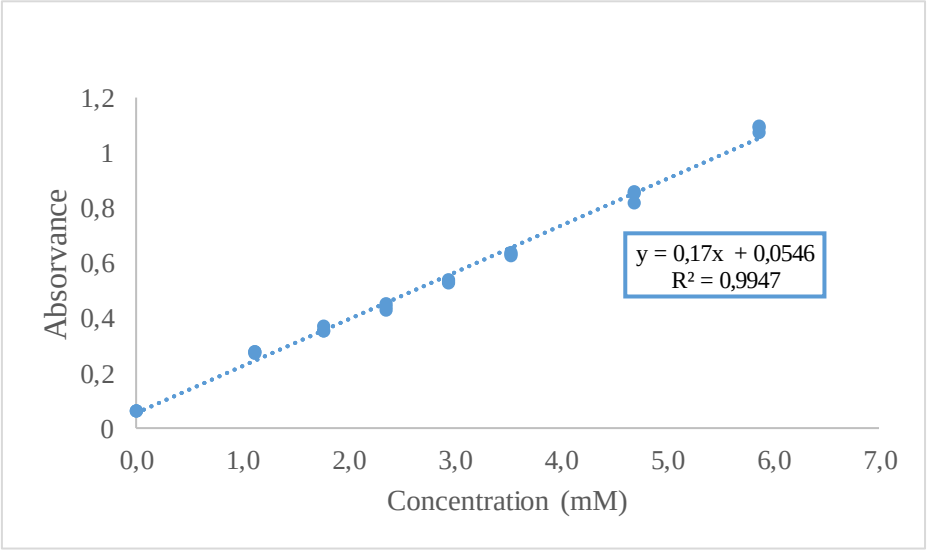
Annex 1: **Typical Calibration curve used for the Bradford method.** A 2 mg/mL BSA solution was used to make the several standard points.



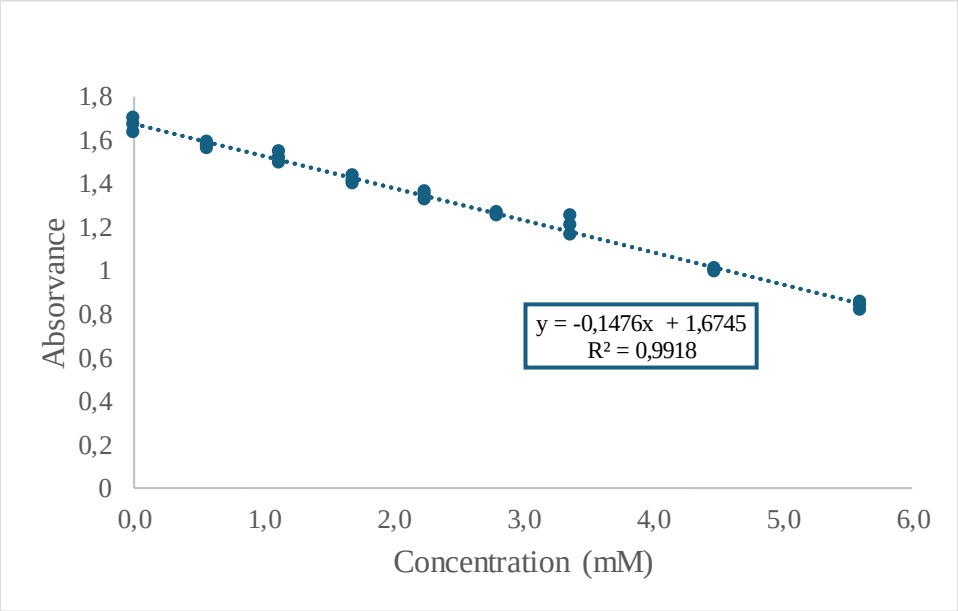
Annex 2: **Typical Calibration curve used for the BCA assay.** A 2 mg/mL BSA solution was used to make the several standard points.



Annex 3: **Marker used in the SDS-PAGE gel.** Precision Plus Protein™ All Blue Prestained Protein Standards (20 µL).



Annex 4: **Typical Calibration curve used for the Folin-Ciocaulteu assay.** A 1 mg/mL Gallic acid solution was used to make the several standard points.



Annex 5: **Typical Calibration curve used for the TEAC assay.** A 1 mg/mL Gallic acid solution was used to make the several standard points.

Annex 6: **Protein concentrations obtained using different quantification methods.** Data represents a mean±standard deviation (n=3)

Sample	Bradford (mg/mL)	BCA (mg/mL)	Warburg and Christian (mg/mL)	Kalb and Bernlohr (mg/mL)
A	88.6±1.2	30.1±1.1	73.6±8.7	84.0±1.1
B	86.7±0.3	32.1±1.7	55.7±4.7	82.1±0.7
C	83.1±0.6	67.1±1.3	68.9±14.3	72.1±0.8
D	71.8±0.6	44.2±1.7	72.4±3.8	76.2±2.7

Kruskal-Wallis Rank Sum Test

Bradford vs KB vs WC	$p<0.05$

Dunn's test

Bradford vs BCA	$p<0.05$
Bradford vs WC	$p<0.05$
Bradford vs KB	$p>0.05$
BCA vs WC	$p>0.05$
BCA vs KB	$p<0.05$
WC vs KB	$p>0.05$

Annex 7: **Protein concentrations obtained using different quantification methods for samples without the presence of inhibitors and which were handled for 10 minutes.** Data represents a mean±standard deviation (n=3)

Group (years)	Sample	Bradford (mg/mL)	Kalb and Bernlohr (mg/mL)	Warburg and Christian (mg/mL)
20-30	4	97.4±0.8	71.6±0.1	67.0±4.5
20-30	6	99.8±1.9	70.8±0.4	62.6±2.8
20-30	7	86.8±0.8	72.4±0.4	62.7±1.2
20-30	9	98.1±0.6	70.6±1.1	66.5±0.6
40-50	1	96.9±2.2	70.7±0.6	67.5±1.1
40-50	2	97.1±1.1	71.3±0.1	70.9±1.4
40-50	3	102.4±0.4	72.4±0.2	64.7±2.3
40-50	5	90.1±0.8	70.6±0.1	68.1±5.6
40-50	8	103.7±0.5	71.2±0.4	58.8±1.2
Kruskal-Wallis Rank Sum Test				
20-30 vs 40-50		p>0.05	p>0.05	p>0.05
Bradford vs KB vs WC			p<0.05	
Dunn's test				
Bradford vs KB			p<0.05	
Bradford vs WC			p<0.05	
KB vs WC			p<0.05	

Annex 8: **Protein concentrations obtained using different quantification methods for samples without the presence of inhibitors and which were handled for 1 hour.** Data represents a mean±standard deviation (n=3)

Group (years)	Sample	Bradford (mg/mL)	Kalb and Bernlohr (mg/mL)	Warburg and Christian (mg/mL)
20-30	4	77.8±1.7	58.1±1.2	48.1±0.4
20-30	6	84.8±0.2	59.9±0.9	52.7±0.4
20-30	7	78.6±0.8	58.5±0.7	49.1±0.9
20-30	9	68.9±1.1	58.5±0.4	50.6±2.1
40-50	1	86.9±0.3	57.6±0.3	52.1±1.5
40-50	2	77.8±1.1	59.8±0.9	54.5±5.1
40-50	3	92.2±0.1	58.4±0.3	50.6±0.7
40-50	5	88.8±1.3	58.4±0.2	52.5±1.1
40-50	8	69.2±1.1	58.1±0.8	53.5±0.6
Kruskal-Wallis Rank Sum Test				
20-30 vs 40-50		p>0.05	p>0.05	p>0.05
Bradford vs KB vs WC			p<0.05	
Dunn's test				
Bradford vs KB			p<0.05	
Bradford vs WC			p<0.05	
KB et all vs WC			p<0.05	

Annex 9: **Protein concentrations obtained using different quantification methods for samples with the presence of inhibitors and which were handled for 10 minutes.** Data represents a mean±standard deviation (n=3)

Group (years)	Sample	Bradford (mg/mL)	Kalb and Bernlohr (mg/mL)	Warburg and Christian 6mg/mL)
20-30	4	129.8±8.7	96.3±1.1	92.3±0.8
20-30	6	112.3±1.1	98.7±0.9	99.1±1.6
20-30	7	133.9±3.7	97.8±0.1	94.3±0.4
20-30	9	112.1±0.9	97.9±0.4	100.1±1.3
40-50	1	117.6±0.6	97.9±0.4	96.2±0.3
40-50	2	129.7±5.4	96.4±1.6	92.8±0.2
40-50	3	116.7±0.9	98.5±1.3	101.6±1.1
40-50	5	180.9±2.7	95.9±0.2	99.6±0.3
40-50	8	110.5±1.1	96.1±1.1	95.6±1.4
Kruskal-Wallis Rank Sum Test				
20-30 vs 40-50		p>0.05	p>0.05	p>0.05
Bradford vs KB vs WC			p<0.05	
Dunn's test				
Bradford vs KB			p<0.05	
Bradford vs WC			p<0.05	
KB vs WC			p<0.05	

Annex 10: **Protein concentrations obtained using different quantification methods for samples with the presence of inhibitors and which were handled for 1 hour.** Data represents a mean±standard deviation (n=3)

Group (years)	Sample	Bradford (mg/mL)	Kalb and Bernlohr (mg/mL)	Warburg and Christian (mg/mL)
20-30	4	107.5±0.1	84.1±0.3	70.9±0.6
20-30	6	106.5±0.4	86.2±0.4	71.4±0.2
20-30	7	105.9±1.4	84.5±0.1	67.1±1.2
20-30	9	102.1±0.7	84.5±01	70.5±0.9
40-50	1	105.2±3.1	85.0±0.3	67.6±2.1
40-50	2	105.2±0.2	84.1±0.6	70.1±4.3
40-50	3	106.8±0.3	85.6±1.4	69.5±0.6
40-50	5	109.5±0.5	85.2±0.2	73.3±0.8
40-50	8	102.1±0.9	86.1±0.2	70.1±0.8
Kruskal-Wallis Rank Sum Test				
20-30 vs 40-50		p>0.05	p>0.05	p>0.05
Bradford vs KB vs WC			p<0.05	
Dunn's test				
Bradford vs KB			p<0.05	
Bradford vs WC			p<0.05	
KB vs WC			p<0.05	

Annex 11: **Antioxidant potential when using the Folin-Ciocaulteu.** Data represents a mean±standard deviation (n=3).

Group (years)	Sample	B_1 hour (eqv. gallic acid mmol)	B_10 min (eqv. gallic acid mmol)	A_1 hour (eqv. gallic acid mmol)	A_10 min (eqv. gallic acid mmol)
20-30	4	0.05±0.01	0.05±0.02	0.09±0.01	0.14±0.01
20-30	6	0.04±0.01	0.04±0.01	0.09±0.01	0.10±0.01
20-30	7	0.06±0.01	0.07±0.03	0.09±0.01	0.16±0.01
20-30	9	0.06±0.01	0.05±0.01	0.09±0.01	0.10±0.01
40-50	1	0.17±0.01	0.02±0.01	0.19±0.01	0.09±0.01
40-50	2	0.06±0.01	0.06±0.01	0.14±0.01	0.09±0.02
40-50	3	0.06±0.02	0.07±0.01	0.07±0.01	0.09±0.01
40-50	5	0.09±0.03	0.03±0.01	0.08±0.01	0.15±0.03
40-50	8	0.04±0.01	0.05±0.01	0.09±0.01	0.09±0.02

Kruskal-Wallis Rank Sum Test

20-30 vs 40-50	p>0.05	p>0.05	p>0.05	p>0.05
B_1hour vs B_10min vs A_1hour vs A_10min			p<0.05	

Dunn's test

B_1hour vs B_10min	p>0.05
B_1hour vs A_1hour	p<0.05
B_1hour vs A_10 min	p<0.05
B_10min vs A_ 1hour	p<0.05
B_10min vs A_10 min	p<0.05
A_10min vs A_1hour	p>0.0

Annex 12: **Antioxidant potential when using the TEAC assay.** Data represents a mean±standard deviation (n=3).

Group (years)	Sample	B_1 hour (eqv. ascorbic acid mmol)	B_10 min (eqv. ascorbic acid mmol)	A_1 hour (eqv. ascorbic acid mmol)	A_10 min (eqv. ascorbic acid mmol)
20-30	4	2.5±0.1	2.9±0.2	2.4±0.2	2.3±0.1
20-30	6	2.7±0.1	2.8±0.2	2.7±0.2	3.7±0.2
20-30	7	2.9±0.1	2.7±0.2	2.8±0.2	2.6±0.4
20-30	9	2.4±0.4	2.5±0.3	2.4±0.2	3.4±0.1
40-50	1	2.2±0.3	2.8±0.1	3.3±0.1	2.7±0.1
40-50	2	2.5±0.5	2.6±0.2	2.5±0.6	2.6±0.3
40-50	3	2.6±0.1	2.7±1.1	2.5±0.2	2.6±0.1
40-50	5	3.1±0.1	3.6±0.1	2.4±0.2	2.3±0.1
40-50	8	2.7±0.3	2.2±0.1	2.5±0.1	2.7±0.1

Kruskal-Wallis Rank Sum Test

20-30 vs 40-50	p>0.05	p>0.05	p>0.05	p>0.05
B_1hour vs B_10min vs A_1hour vs A_10min			p>0.05	

Annex 13: **Identified proteins.** This table contains data on the 350 proteins identified across all the different kits. The "n" indicates absence/not present and the "s" signifies presence/yes.

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
2-hydroxyacylsphingosine 1-beta-galactosyltransferase	*Homo sapiens*	*UGT8*	61.63	54.29	n	n	n	s
ABC transporter	*Rubneribacter badeniensis*	*C2L80_04475*	67.60	34.04	n	s	n	n
ABC transporter ATP-binding protein	*Haematobacter massiliensis*	*CN97_12045*	31.28	15.65	n	n	n	s
ABC transporter ATP-binding protein	*Roseburia intestinalis*	*DW264_01450*	73.94	76.12	n	n	n	s
ABC transporter permease	*Coprococcus sp. OM04-5BH*	*DXB54_00675*	103.16	40.49	n	n	n	s
ABC transporter permease	*Lachnospiraceae bacterium TF09-5*	*DXC97_21875*	92.68	54.75	n	n	n	s
ABC3 transporter permease protein domain-containing protein	*Clostridiales bacterium*	*DBY05_04135*	84.10	31.82	n	n	n	s
ABC-type glutathione-S-conjugate transporter	*Myotis lucifugus*	*ABCC2*	174.37	24.86	s	n	n	s
ACX73 (Fragment)	*Homo sapiens*	-	27.90	185.86	-	s	n	s
Afamin	*Homo sapiens*	*AFM*	70.96	206.52	n	s	n	s
Albumin	*Homo sapiens*	*ALB*	70.96	241.09	s	s	s	s
Albumin	*Myotis lucifugus*	*ALB*	71.32	4011.62	s	s	n	s
Alpha-1-acid glycoprotein 1	*Homo sapiens*	*ORM1*	71.32	3567.44	s	s	s	s
Alpha-1-acid glycoprotein 2	*Homo sapiens*	*ORM2*	71.32	4181.67	s	s	s	s
Alpha-1-antichymotrypsin	*Homo sapiens*	*SERPINA3*	70.34	177.91	s	s	s	s
Alpha-1-antitrypsin	*Homo sapiens*	*SERPINA1*	70.34	221.95	s	s	s	s
Alpha-1-B glycoprotein	*Pan troglodytes*	*A1BG*	23.75	463.63	n	s	s	s
Alpha-1B-glycoprotein	*Homo sapiens*	*A1BG*	23.75	340.52	n	n	n	s
Alpha-2-antiplasmin	*Homo sapiens*	*SERPINF2*	23.75	617.20	s	s	s	s
Alpha-2-glycoprotein 1, zinc-binding	*Pan troglodytes*	*AZGP1*	23.87	500.14	s	s	n	s
Alpha-2-HS-glycoprotein	*Homo sapiens*	*AHSG*	23.87	224.49	s	s	s	s
Alpha-2-macroglobulin	*Homo sapiens*	*A2M*	23.87	486.53	s	s	s	s
Amino acid transporter (Fragment)	*Homo sapiens*	*SLC1A3*	47.79	862.89	s	n	n	s
Angiotensinogen	*Homo sapiens*	*AGT*	47.79	436.51	s	s	s	s
Anti-FactorVIII scFv (Fragment)	*Homo sapiens*	-	47.79	1006.28	-	n	s	s
Antithrombin-III	*Homo sapiens*	*SERPINC1*	46.88	1741.29	s	s	s	s
Apolipoprotein A-I	*Homo sapiens*	*APOA1*	46.88	1235.49	s	s	s	s
Apolipoprotein A-I	*Macaca fascicularis*	*APOA1*	46.88	1816.42	s	s	n	s
Apolipoprotein A-II	*Homo sapiens*	*APOA2*	54.88	429.82	s	s	s	s
Apolipoprotein A-IV	*Homo sapiens*	*APOA4*	54.88	313.75	s	s	s	s
Apolipoprotein A-IV	*Pan paniscus*	*APOA4*	54.88	178.17	s	n	n	s
Apolipoprotein B-100	*Homo sapiens*	*APOB*	54.79	253.54	s	s	s	s
Apolipoprotein C-I	*Homo sapiens*	*APOC1*	54.87	228.04	s	s	s	s
Apolipoprotein C-II	*Homo sapiens*	*APOC2*	54.87	130.47	s	s	s	n
Apolipoprotein C-II	*Pan paniscus*	*APOC2*	54.87	237.86	s	n	n	s
Apolipoprotein C-III	*Homo sapiens*	*APOC3*	34.53	176.25	s	s	s	s
Apolipoprotein D	*Homo sapiens*	*APOD*	34.53	330.15	s	s	s	s
Apolipoprotein E	*Homo sapiens*	*APOE*	40.11	339.52	s	s	s	s
Apolipoprotein E	*Pan paniscus*	*APOE*	40.11	315.90	s	s	s	s
Apolipoprotein E	*Tursiops truncatus*	*APOE*	40.11	243.84	s	n	n	s
Apolipoprotein L1	*Homo sapiens*	*APOL1*	164.61	3946.35	s	n	s	s
Apolipoprotein M	*Homo sapiens*	*APOM*	164.61	2346.36	s	s	n	s
ATP synthase subunit b	*Selenomonadales bacterium*	*atpF*	164.61	3573.08	n	s	n	n
ATP synthase subunit delta	*Limosilactobacillus mucosae*	*atpH*	10.23	51.52	n	n	n	s

Annex 13: **Continuation.**

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
ATP-binding cassette domain-containing protein	*Hungatella hathewayi*	*DXC39_16800*	52.32	363.00	n	s	n	s
ATP-dependent zinc metalloprotease FtsH	*Acetivibrio ethanolgignens*	*ftsH*	52.32	225.16	n	n	s	n
ATP-dependent zinc metalloprotease FtsH	*Candidatus Melainabacteria bacterium MEL.A1*	*ftsH*	52.32	489.94	n	s	n	n
ATP-dependent zinc metalloprotease FtsH	*Clostridium sp. AM22-11AC*	*ftsH*	25.18	175.07	n	s	n	n
ATP-dependent zinc metalloprotease FtsH	*Enterocloster aldenensis*	*ftsH*	25.18	351.08	n	s	n	s
ATP-dependent zinc metalloprotease FtsH	*Oscillospiraceae bacterium*	*ftsH*	53.03	669.80	n	n	n	s
Beta-2-glycoprotein 1	*Homo sapiens*	*APOH*	53.03	341.86	s	s	s	s
BMP family ABC transporter substrate-binding protein	*Clostridium innocuum*	*ADH65_05660*	53.03	999.20	n	n	n	s
C2 domain-containing protein	*Paracoccidioides lutzii (strain ATCC MYA-826 / Pb01)*	*PAAG_02650*	30.76	2107.97	n	s	n	n
C4b-binding protein alpha chain	*Homo sapiens*	*C4BPA*	30.76	1736.22	s	s	s	s
Cadherin domain protein	*Necator americanus*	*NECAME_00460*	30.76	2556.97	n	s	n	n
Cadherin domain-containing protein	*SR1 bacterium human oral taxon HOT-345*	*BSK20_01290*	30.72	1383.12	n	s	n	n
Calcium-transporting ATPase	*Nocardia cerradoensis*	*pacL*	30.72	1720.17	n	n	n	s
Carboxypeptidase N catalytic chain	*Homo sapiens*	*CPN1*	11.28	252.85	s	n	n	s
Carboxypeptidase N subunit 2	*Homo sapiens*	*CPN2*	11.28	187.15	s	s	n	s
Carboxypeptidase Q (Fragment)	*uncultured bacterium*	*CPQ*	11.28	292.78	n	n	n	s
Cd(2+)-exporting ATPase	*Clostridium sp. AF34-13*	*DWZ63_14940*	45.34	793.50	n	s	n	n
CD5 antigen-like	*Homo sapiens*	*CD5L*	45.34	746.58	n	n	n	s
CD5 molecule like	*Pan troglodytes*	*CD5L*	45.34	696.90	n	s	n	n
CDK5 regulatory subunit-associated protein 2	*Rattus norvegicus*	*Cdk5rap2*	45.46	1066.80	s	n	n	s
cDNA FLJ25298 fis, clone STM07683, highly similar to Protein Tro alpha1 H,myeloma	*Homo sapiens*	-	516.65	5692.49	-	s	n	n
Centrosomal protein of 83 kDa	*Danio rerio*	*cep83*	516.65	2215.21	s	s	n	n
Ceruloplasmin	*Homo sapiens*	*CP*	516.65	7790.83	s	s	s	s
chitin synthase	*Sporothrix brasiliensis 5110*	*SPBR_02173*	9.33	147.45	n	s	n	n
Clusterin	*Homo sapiens*	*CLU*	9.33	152.66	s	s	s	s
Clusterin	*Pan paniscus*	*CLU*	9.33	220.14	s	n	n	s
Coagulation factor X	*Homo sapiens*	*F10*	11.28	135.65	s	n	n	s
Coagulation factor XII	*Homo sapiens*	*F12*	11.28	191.06	s	s	s	s
Complement C1r subcomponent	*Homo sapiens*	*C1R*	20.12	166.35	s	s	n	s
Complement C1r subcomponent-like protein	*Homo sapiens*	*C1RL*	10.85	290.26	n	n	n	s
Complement C1s subcomponent	*Homo sapiens*	*C1S*	10.85	305.91	s	s	s	s
Complement C2	*Homo sapiens*	*C2*	10.85	183.49	s	s	n	s
Complement C3	*Homo sapiens*	*C3*	21.55	126.56	s	s	s	s
Complement C4-A	*Homo sapiens*	*C4A*	21.55	83.28	s	s	s	s
Complement C4-B	*Homo sapiens*	*C4B*	21.55	321.84	s	n	s	s
Complement C4-B	*Mus musculus*	*C4b*	36.25	572.40	s	n	n	s
Complement C5	*Homo sapiens*	*C5*	36.25	590.85	s	s	n	s
Complement component 4 binding protein alpha	*Pan troglodytes*	*C4BPA*	36.25	708.38	s	n	n	s
Complement component C6	*Homo sapiens*	*C6*	36.14	402.25	s	n	n	s
Complement component C6	*Pan troglodytes*	*C6*	36.14	529.03	s	s	n	s
Complement component C7	*Homo sapiens*	*C7*	36.14	616.41	s	s	n	s
Complement component C8 alpha chain	*Homo sapiens*	*C8A*	36.33	249.21	s	n	n	s
Complement component C8 beta chain	*Homo sapiens*	*C8B*	44.00	101.24	s	n	n	s
Complement component C8 beta chain	*Pan paniscus*	*C8B*	44.00	364.96	s	s	n	n

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
Complement component C9	*Equus caballus*	*C9*	21.58	130.26	s	n	n	s
Complement component C9	*Homo sapiens*	*C9*	21.58	53.09	s	s	s	s
Complement factor B	*Homo sapiens*	*CFB*	18.30	29.50	s	s	s	s
Complement factor H	*Homo sapiens*	*CFH*	20.18	29.61	s	s	s	s
Complement factor H-related protein 1	*Homo sapiens*	*CFHR1*	92.05	39.21	s	n	s	n
Complement factor H-related protein 2	*Homo sapiens*	*CFHR2*	68.99	49.62	s	n	n	s
Complement factor I	*Homo sapiens*	*CFI*	70.03	52.20	s	s	n	s
Copper-transporting ATPase	*Plasmodium falciparum (isolate 3D7)*	*PF3D7_0904900*	68.71	30.09	n	n	n	s
Coronin	*Pan paniscus*	*CORO1B*	68.84	32.25	n	n	s	n
Corticosteroid-binding globulin	*Homo sapiens*	*SERPINA6*	68.84	27.60	s	s	s	s
C-reactive protein	*Homo sapiens*	*CRP*	66.66	49.38	s	s	n	n
Cytochrome bd-I ubiquinol oxidase subunit CydA	*Salmonella typhimurium*	*cydA_3*	39.58	157.73	n	n	n	s
Diadenylate cyclase	*Coprococcus sp. OM04-5BH*	*dacA*	39.58	345.30	n	s	n	n
diguanylate cyclase	*Pseudomonas sp. S10E 269*	*CWC48_23560*	39.58	155.26	n	s	n	n
DIRAS family GTPase 3	*Oryctolagus cuniculus*	*DIRAS3*	39.87	23.71	n	n	n	s
DNA topoisomerase 4 subunit A	*Staphylococcus haemolyticus (strain JCSC1435)*	*parC*	142.10	29.07	n	s	n	s
DNA translocase FtsK	*Alistipes finegoldii*	*F2A26_05465*	69.04	205.75	n	n	n	s
DNA translocase FtsK	*Alistipes onderdonkii*	*F2Y10_09840*	69.04	260.73	n	s	s	n
Dystonin	*Homo sapiens*	*DST*	69.04	179.34	s	s	n	n
Dystonin	*Oryctolagus cuniculus*	*DST*	281.03	41.31	s	s	n	n
Dystrophin-related protein 2	*Mus musculus*	*Drp2*	160.37	36.89	n	n	n	s
Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	*Mus musculus*	*Enpp1*	95.16	44.12	s	s	n	n
Efflux pump membrane transporter	*Derxia gummosa DSM 723*	-	52.54	81.58	-	s	n	n
Env polyprotein (Fragment)	*Human immunodeficiency virus 1*	*env*	61.37	107.48	n	s	n	n
Envelope glycoprotein gp160	*Human immunodeficiency virus 1*	*env*	61.37	124.43	n	s	n	s
Envelope glycoprotein gp160 (Fragment)	*Human immunodeficiency virus 1*	*env*	23.62	20.33	n	s	n	n
Envelope glycoprotein gp160 (Fragment)	*Human immunodeficiency virus 1*	*env*	95.01	51.96	n	s	n	n
Erythrocyte membrane protein 1, PfEMP1	*Plasmodium falciparum (isolate 3D7)*	*PF3D7_0425800*	39.60	83.20	n	n	n	s
ESX-3 secretion system protein EccB3	*Mycobacterium tuberculosis*	*eccB3*	39.70	89.01	n	s	s	s
Exocyst complex component 8	*Danio rerio*	*exoc8*	216.95	16.22	s	s	n	n
Exocyst complex component SEC5	*Aspergillus thermomutatus*	*CDV56_101051*	54.47	972.01	n	n	n	s
Ferric siderophore transport system protein ExbB	*Aeromonas salmonicida*	*C5B77_08420*	84.09	74.71	n	s	n	n
ferroxidase	*Pan troglodytes*	*CP*	123.00	1209.97	s	s	s	s
Fetuin-B	*Homo sapiens*	*FETUB*	123.00	624.26	n	s	n	s
Fibrinogen alpha chain	*Homo sapiens*	*FGA*	123.00	1515.00	s	s	s	s
Fibrinogen beta chain	*Homo sapiens*	*FGB*	210.80	46.35	s	s	s	s
Fibrinogen beta chain	*Myotis lucifugus*	*FGB*	53.03	476.81	s	s	n	n
Fibrinogen beta chain	*Oryctolagus cuniculus*	*FGB*	53.03	564.98	s	s	n	s
Fibrinogen beta chain	*Pan paniscus*	*FGB*	53.03	492.16	s	n	n	s
Fibrinogen gamma chain	*Cavia porcellus*	*FGG*	52.93	363.36	s	s	n	n
Fibrinogen gamma chain	*Homo sapiens*	*FGG*	56.07	47.33	s	s	s	s
Fibrinogen gamma chain	*Myotis lucifugus*	*FGG*	70.03	76.74	s	n	n	s

Annex 13: **Continuation.**

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
Fibronectin	*Homo sapiens*	*FN1*	70.03	81.84	s	s	n	s
Fibronectin	*Myotis lucifugus*	*FN1*	70.03	81.95	s	s	n	s
Ficolin-3	*Homo sapiens*	*FCN3*	81.61	108.01	s	s	n	s
Galectin-3-binding protein	*Homo sapiens*	*LGALS3BP*	81.61	37.00	s	n	n	s
GCT-A4 heavy chain variable region (Fragment)	*Homo sapiens*	-	54.21	61.61	-	n	n	s
Gelsolin	*Pan paniscus*	*GSN*	78.17	132.68	s	s	s	s
Glutathione peroxidase 3	*Homo sapiens*	*GPX3*	78.17	62.82	n	s	n	n
Glyceraldehyde-3-phosphate dehydrogenase	*Pan troglodytes*	*GAPDH*	78.17	381.40	n	n	s	n
Glycosyl/glycerophosphate transferase	*Leucobacter massiliensis*	*B4915_08365*	84.58	105.74	n	n	n	s
Golgin-45	*Anopheles funestus*	-	84.58	207.57	-	s	n	n
HAD family hydrolase	*Blautia sp. AM23-13AC*	*DW642_10905*	188.57	4424.60	n	s	n	n
HAD family hydrolase	*Eubacterium sp. AF17-7*	*DWW96_13075*	188.57	2487.04	n	n	n	s
HAD-IC family P-type ATPase	*Candidatus Saccharibacteria bacterium*	*HG451_000945*	188.57	4574.26	n	s	n	n
Haptoglobin	*Homo sapiens*	*HP*	194.26	1846.15	s	s	s	s
Haptoglobin	*Myotis lucifugus*	*HP*	194.26	650.46	s	s	n	s
Haptoglobin-related protein	*Homo sapiens*	*HPR*	194.26	1941.47	n	s	s	s
HAUS augmin-like complex subunit 3	*Homo sapiens*	*HAUS3*	194.17	649.77	n	s	n	n
HAUS augmin-like complex subunit 8	*Homo sapiens*	*HAUS8*	194.17	1841.65	n	s	n	n
Heme ABC transporter ATP-binding protein	*Oscillospiraceae bacterium*	*DBY17_01615*	194.45	439.13	n	s	n	n
Hemoglobin subunit alpha	*Homo sapiens*	*HBA1*	189.90	504.73	s	s	s	s
Hemoglobin subunit beta	*Homo sapiens*	*HBB*	189.90	724.73	s	s	s	s
Hemoglobin, beta adult 2	*Danio rerio*	*hbba2*	69.17	137.75	n	s	n	n
Hemopexin	*Homo sapiens*	*HPX*	108.37	224.94	s	s	s	s
Heparin cofactor 2	*Homo sapiens*	*SERPIND1*	108.50	154.34	s	s	s	s
histidine kinase	*Treponema sp. Marseille-Q3903*	*H9I37_09985*	108.50	116.12	n	s	n	s
Histidine-rich glycoprotein	*Homo sapiens*	*HRG*	96.65	98.19	s	s	s	s
Hyaluronan-binding protein 2	*Homo sapiens*	*HABP2*	96.65	116.33	s	n	n	s
IBM-B2 heavy chain variable region (Fragment)	*Homo sapiens*	-	66.83	111.00	-	n	n	s
IG c102_heavy_IGHV3-48_IGHD1-26_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	63.88	49.49	n	n	n	s
IG c1037_heavy_IGHV4-34_IGHD3-22_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	68.78	54.43	n	n	n	s
IG c1064_heavy_IGHV3-73_IGHD6-6_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	63.40	31.68	n	n	n	s
IG c1247_heavy_IGHV1-69_IGHD1-1_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	64.62	317.67	n	n	n	s
IG c1355_heavy_IGHV4-4_IGHD2-2_IGHJ1 (Fragment)	*Homo sapiens*	*multiple*	64.62	216.01	n	s	n	n
IG c1405_heavy_IGHV3-7_IGHD3-3_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	64.62	296.56	n	n	n	s
IG c1447_heavy_IGHV3-15_IGHD3-3_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	86.85	637.19	n	s	n	n
IG c1609_heavy_IGHV3-21_IGHD3-10_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	86.85	534.34	n	n	s	s
IG c1654_heavy_IGHV3-30_IGHD3-10_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	86.85	720.55	n	s	n	n
IG c198_heavy_IGHV5-51_IGHD7-27_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	143.68	441.56	n	s	n	n
IG c2_heavy_IGHV6-1_IGHD2-2_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	143.68	479.49	n	n	n	s
IG c255_heavy_IGHV3-15_IGHD3-10_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	143.68	707.02	n	n	n	s
IG c366_heavy_IGHV3-7_IGHD1-26_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	38.77	250.01	n	s	n	s
IG c396_heavy_IGHV4-34_IGHD2-2_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	31.54	73.85	n	n	n	s
IG c488_heavy_IGHV3-49_IGHD3-10_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	68.10	92.17	n	n	n	s
IG c507_heavy_IGHV3-7_IGHD3-3_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	68.10	218.89	n	n	n	s
IG c543_heavy_IGHV3-9_IGHD5-24_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	301.60	52.07	n	n	n	s

Annex 13: **Continuation.**

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
IG c55_heavy_IGHV3-48_IGHD3-10_IGHJ3 (Fragment)	*Homo sapiens*	*multiple*	54.92	89.95	n	n	n	s
IG c574_heavy_IGHV4-61_IGHD3-16_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	45.28	137.56	n	n	n	s
IG c720_heavy_IGHV3-49_IGHD3-3_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	45.28	36.40	n	s	n	n
IG c733_heavy_IGHV3-7_IGHD3-9_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	45.28	105.36	n	s	n	n
IG c828_heavy_IGHV4-61_IGHD3-22_IGHJ3 (Fragment)	*Homo sapiens*	*multiple*	25.19	74.95	n	s	n	n
IG c855_heavy_IGHV3-23_IGHD1-26_IGHJ1 (Fragment)	*Homo sapiens*	*multiple*	58.05	32.42	n	n	n	s
Ig heavy chain variable region (Fragment)	*Homo sapiens*	*multiple*	33.18	27.33	n	n	n	s
IgG H chain	*Homo sapiens*	-	59.34	41.51	-	n	n	s
IGH + IGL c20_heavy_IGHV3-21_IGHD3-22_IGHJ3 (Fragment)	*Homo sapiens*	*multiple*	31.64	20.78	n	n	n	s
IGH + IGL c326_heavy_IGHV1-69_IGHD2-2_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	90.51	70.71	n	n	n	s
IGH + IGL c336_heavy_IGHV3-23_IGHD3-22_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	106.51	150.24	n	n	n	s
IGH + IGL c38_heavy_IGHV3-15_IGHD1-14_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	105.94	143.77	n	n	n	s
IGH + IGL c392_heavy_IGHV1-18_IGHD3-9_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	105.94	123.12	n	n	n	s
IGH + IGL c517_heavy_IGHV2-70_IGHD4-17_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	865.26	94.35	n	n	n	s
IGH + IGL c534_heavy_IGHV4-39_IGHD3-22_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	292.47	85.70	n	s	n	n
IGH + IGL c62_heavy_IGHV3-7_IGHD3-22_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	109.24	28.02	n	s	n	n
IGH c1126_heavy_IGHV3-11_IGHD6-13_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	105.16	48.32	n	n	n	s
IGH c1340_heavy_IGHV3-7_IGHD2-2_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	112.61	40.93	n	s	n	s
IGH c1503_heavy_IGHV3-49_IGHD1-1_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	20.48	24.13	n	n	s	n
IGH c1758_heavy_IGHV3-33_IGHD3-16_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	97.78	46.75	n	s	n	n
IGH c1796_heavy_IGHV4-39_IGHD3-10_IGHJ3 (Fragment)	*Homo sapiens*	*multiple*	97.78	20.63	n	s	n	n
IGH c2097_heavy_IGHV3-7_IGHD6-19_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	72.29	80.42	n	n	n	s
IGH c2207_heavy_IGHV4-4_IGHD5-24_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	99.52	31.25	n	n	n	s
IGH c235_heavy__IGHV3-23_IGHD3-10_IGHJ5 (Fragment)	*Homo sapiens*	*multiple*	97.92	37.17	n	n	n	s
IGH c2826_heavy_IGHV5-51_IGHD5-24_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	98.50	62.77	n	n	n	s
IGH c283_heavy__IGHV3-30_IGHD6-19_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	99.47	38.31	n	n	n	s
IGH c3844_heavy_IGHV3-74_IGHD6-19_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	411.97	45.73	n	n	n	s
IGH c3886_heavy_IGHV3-15_IGHD2-15_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	56.39	48.05	n	n	n	s
IGH c395_heavy__IGHV3-7_IGHD1-14_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	56.39	22.46	n	n	n	s
IGH c4042_heavy_IGHV4-34_IGHD5-12_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	56.39	50.95	n	n	n	s
IGH c4137_heavy_IGHV5-51_IGHD5-12_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	79.33	27.44	n	n	n	s
IGH c557_heavy__IGHV4-34_IGHD3-9_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	115.03	29.17	n	n	n	s
IGH c776_heavy_IGHV3-7_IGHD3-22_IGHJ6 (Fragment)	*Homo sapiens*	*multiple*	19.53	19.53	n	n	n	s
IGH c817_heavy_IGHV3-7_IGHD7-27_IGHJ4 (Fragment)	*Homo sapiens*	*multiple*	123.35	1151.33	n	s	n	n
IGH c905_heavy_IGHV1-18_IGHD3-22_IGHJ3 (Fragment)	*Homo sapiens*	*multiple*	123.35	576.19	n	n	n	s
IGH@ protein	*Homo sapiens*	*IGH@*	123.35	1517.73	n	n	n	s
IGHG3 (Fragment)	*Homo sapiens*	*IGHG3*	42.88	47.75	n	s	s	s
Ig-like domain-containing protein	*Homo sapiens*	*DKFZp686K03196*	42.88	84.40	-	s	s	s
Ig-like domain-containing protein	*Pan troglodytes*	-	95.66	1650.94	-	n	n	s
Immunoglobulin alpha-2 heavy chain	*Homo sapiens*	-	95.66	1249.44	-	n	n	s
Immunoglobulin delta heavy chain	*Homo sapiens*	-	95.66	1972.96	-	s	n	s
Immunoglobulin heavy constant gamma 2	*Homo sapiens*	*IGHG2*	56.58	1405.98	n	s	n	s
Immunoglobulin heavy constant mu	*Homo sapiens*	*IGHM*	56.58	1323.94	s	s	s	s
Immunoglobulin heavy variable 3-72	*Homo sapiens*	*IGHV3-72*	56.58	1231.70	n	s	n	n
Immunoglobulin J chain	*Homo sapiens*	*JCHAIN*	55.18	324.70	n	s	s	s

Annex 13: **Continuation.**

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
Immunoglobulin kappa constant	*Homo sapiens*	*IGKC*	57.16	285.29	n	s	s	s
Immunoglobulin kappa light chain	*Homo sapiens*	*IGKV*	57.16	334.56	n	s	s	s
Immunoglobulin kappa variable 2-24	*Homo sapiens*	*IGKV2-24*	56.56	1707.15	n	s	s	s
Immunoglobulin kappa variable 2-28	*Homo sapiens*	*IGKV2-28*	51.46	525.79	n	s	s	s
Immunoglobulin kappa variable 2-30	*Homo sapiens*	*IGKV2-30*	52.11	1368.78	n	n	n	s
Immunoglobulin kappa variable 3D-11	*Homo sapiens*	*IGKV3D-11*	52.11	1168.54	n	n	n	s
Immunoglobulin kappa variable 3D-20	*Homo sapiens*	*IGKV3D-20*	52.11	1298.14	n	s	s	s
Immunoglobulin lambda constant 2	*Homo sapiens*	*IGLC2*	52.60	157.41	n	s	s	s
Immunoglobulin lambda constant 7	*Homo sapiens*	*IGLC7*	275.74	742.53	n	s	s	s
Immunoglobulin lambda variable 2-8	*Homo sapiens*	*IGLV2-8*	275.74	736.10	n	s	n	n
Immunoglobulin lambda-1 light chain	*Homo sapiens*	*IGLL1*	275.10	343.50	s	s	s	s
Immunoglobulin mu heavy chain	*Homo sapiens*	-	275.10	312.99	-	s	n	n
Insulin-like growth factor-binding protein complex acid labile subunit	*Homo sapiens*	*IGFALS*	33.40	26.21	s	n	n	s
Inter-alpha (Globulin) inhibitor H4 (Plasma Kallikrein-sensitive glycoprotein)	*Homo sapiens*	*ITIH4*	33.40	22.77	s	s	s	n
Inter-alpha (Globulin) inhibitor H4 (Plasma Kallikrein-sensitive glycoprotein) variant (Fragment)	*Homo sapiens*	-	66.20	73.91	-	s	n	n
Inter-alpha-trypsin inhibitor heavy chain 3	*Pan troglodytes*	*ITIH3*	14.41	156.89	n	n	n	s
Inter-alpha-trypsin inhibitor heavy chain 4	*Pan troglodytes*	*ITIH4*	86.77	629.04	n	n	n	s
Inter-alpha-trypsin inhibitor heavy chain H1	*Homo sapiens*	*ITIH1*	86.77	307.96	n	s	s	s
Inter-alpha-trypsin inhibitor heavy chain H2	*Homo sapiens*	*ITIH2*	86.77	552.88	n	s	s	s
Inter-alpha-trypsin inhibitor heavy chain H3	*Homo sapiens*	*ITIH3*	25.77	25.20	n	s	n	s
Inter-alpha-trypsin inhibitor heavy chain H4	*Oryctolagus cuniculus*	*ITIH4*	36.20	83.23	n	s	n	s
Ion-translocating oxidoreductase complex subunit C	*Clostridiales bacterium*	*rnfC*	86.40	49.97	n	s	n	n
IQGAP1 protein	*Danio rerio*	*IQGAP1*	48.91	32.09	n	n	n	s
Iron ABC transporter permease	*Mitsuokella sp. AF21-1AC*	*DWX75_03725*	90.42	54.96	n	s	n	n
Junction plakoglobin	*Homo sapiens*	*JUP*	89.96	51.34	s	n	n	s
Kallistatin	*Homo sapiens*	*SERPINA4*	90.91	50.43	n	n	n	s
Keratin 1	*Pan troglodytes*	*KRT1*	45.86	1652.86	s	n	n	s
Keratin, type I cytoskeletal 18	*Mus musculus*	*Krt18*	45.86	1387.95	s	s	n	n
Keratin, type I cytoskeletal 19	*Cavia porcellus*	*KRT19*	45.86	1600.40	s	n	s	s
Keratin, type II cytoskeletal 1	*Homo sapiens*	*KRT1*	39.20	134.46	s	n	s	s
Keratin, type II cytoskeletal 1	*Pan troglodytes*	*KRT1*	39.20	50.45	s	s	n	s
Keratin, type II cytoskeletal 5	*Mus musculus*	*Krt5*	39.52	817.81	s	n	s	n
Keratin, type II cytoskeletal 7	*Homo sapiens*	*KRT7*	39.52	706.82	s	n	s	n
Kininogen-1	*Homo sapiens*	*KNG1*	39.52	668.62	s	s	s	s
L-amino acid oxidase	*Cerastes cerastes*	-	70.01	43.44	n	n	n	s
Large-conductance mechanosensitive channel	*Firmicutes bacterium AM55-24TS*	*mscL*	44.94	16.46	n	s	n	s
Lauroyl acyltransferase	*Desulfobulbus oralis*	*CAY53_11640*	63.44	23.16	n	s	n	n
Leucine-rich alpha-2-glycoprotein	*Homo sapiens*	*LRG1*	15.31	251.01	n	s	n	s
Lumican	*Homo sapiens*	*LUM*	15.31	84.86	s	s	n	s
Mechanosensitive channel MscK	*Pseudomonas sp. 02C 26*	*CXQ80_00965*	15.31	183.17	n	n	n	s
Melatonin receptor type 1A	*Oryctolagus cuniculus*	*MTNR1A*	16.10	422.55	n	s	n	n
Membrane metalloprotease	*Intestinimonas butyriciproducens*	*IB211_01597*	16.10	136.89	n	n	n	s
Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 1	*Homo sapiens*	*MAGI1*	16.10	400.21	s	s	n	n
MFS transporter	*Bacillota bacterium*	*DBX60_06090*	16.46	39.03	n	s	n	n

Annex 13: **Continuation.**

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
Miniconductance mechanosensitive channel MscM	*Aeromonas salmonicida*	*C5B77_12590*	52.39	678.33	n	n	n	s
MMPL family transporter (Fragment)	*Limosilactobacillus mucosae*	*JTJ10_03210*	52.39	566.68	n	s	n	n
Multifunctional fusion protein	*Blautia obeum*	*secD*	52.39	610.95	n	s	n	n
Na/Pi cotransporter	*Clostridiales bacterium*	*DBX53_06585*	57.21	458.01	n	n	n	s
Na+/H+ antiporter subunit A	*Auritidibacter sp. NML120779*	*DCC26_07245*	57.21	123.22	n	s	n	n
N-acetylmuramoyl-L-alanine amidase	*Homo sapiens*	*PGLYRP2*	57.21	473.46	n	n	n	s
Neuraminidase	*Influenza A virus*	*NA*	48.93	59.14	n	s	n	n
Nuclear transcription factor, X-box binding 1	*Pan paniscus*	*NFX1*	48.93	39.06	n	s	n	n
Obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	*Homo sapiens*	*OBSCN*	60.51	407.43	s	s	s	s
Oligosaccharide flippase family protein	*Bifidobacterium ramosum*	*DSM100688_0224*	60.51	369.03	n	n	n	s
Oligosaccharide flippase family protein	*Clostridiales bacterium*	*HG454_000635*	60.51	592.63	n	s	n	n
Otoferlin isoform X3	*Danio rerio*	*otofb*	64.74	60.90	n	s	n	n
Outer capsid protein VP4 (Fragment)	*Human rotavirus A*	*VP4*	13.68	172.21	n	n	n	s
Phenylalanine-specific permease	*Escherichia coli O1:K1 / APEC*	*pheP*	12.71	77.87	n	n	n	s
Phosphatidate cytidylyltransferase	*Bifidobacterium avesanii*	*GFD22_00510*	15.62	141.76	n	s	n	n
Phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta	*Cavia porcellus*	*PIK3C2B*	13.91	63.37	n	n	n	s
Phosphatidylinositol-glycan-specific phospholipase D	*Homo sapiens*	*GPLD1*	13.93	89.52	n	n	n	s
Phosphoethanolamine transferase	*Haemophilus seminalis*	*F2S80_01850*	12.95	123.89	n	n	n	s
Phosphorylase b kinase regulatory subunit	*Clonorchis sinensis*	*CLF_100451*	12.94	81.34	n	n	n	s
Phosphotransferase RcsD	*Enterobacter sp. AM17-18*	*rcsD*	13.75	110.19	n	n	n	s
Pigment epithelium-derived factor	*Homo sapiens*	*SERPINF1*	13.15	127.47	s	s	s	s
Plasma kallikrein	*Homo sapiens*	*KLKB1*	13.15	143.43	s	s	n	s
Plasminogen	*Homo sapiens*	*PLG*	13.20	76.07	s	s	s	s
Plasminogen	*Myotis lucifugus*	*PLG*	14.79	123.25	s	n	n	s
Plexin A4	*Callorhinchus milii*	*plxna4*	14.02	86.77	n	s	n	n
Podocin	*Danio rerio*	*nphs2*	13.20	148.82	s	s	n	s
Pol protein (Fragment)	*Human immunodeficiency virus 1*	*pol*	13.52	71.95	n	n	n	s
Polyprotein P1234	*Barmah forest virus*	-	13.52	132.81	-	n	n	s
Polysaccharide biosynthesis protein	*Eisenbergiella massiliensis*	*DXC51_15365*	13.05	81.01	n	n	n	s
Polysaccharide biosynthesis protein	*Lachnospiraceae bacterium AM26-1LB*	*DW698_08170*	13.54	119.58	n	s	s	s
Polysaccharide biosynthesis protein C-terminal domain-containing protein	*Clostridiales bacterium*	*DBY36_06110*	14.02	73.98	n	n	n	s
Pregnancy zone protein	*Homo sapiens*	*PZP*	13.43	89.67	n	s	n	s
Pregnancy-associated plasma protein A, pappalysin 1b isoform X1	*Danio rerio*	*pappab*	13.72	121.64	n	n	n	s
Probable non-functional immunoglobulin kappa variable 3-7	*Homo sapiens*	*IGKV3-7*	13.74	116.01	n	s	s	n
Protein AMBP	*Homo sapiens*	*AMBP*	14.68	88.43	n	s	n	n
Protein AMBP	*Pan paniscus*	*AMBP*	13.38	117.25	n	n	n	s
Protein AMBP	*Pan troglodytes*	*AMBP*	13.88	120.25	n	n	s	s
Protein S100	*Pan paniscus*	*S100A9*	13.16	57.26	n	n	s	n
Protein S100-A8	*Homo sapiens*	*S100A8*	11.19	70.52	n	n	s	n
Protein translocase subunit SecA	*Eggerthella sp. (strain YY7918)*	*SecA*	32.60	215.53	n	s	n	n
Protein translocase subunit SecA	*Faecalibacterium prausnitzii*	*secA*	32.67	175.08	n	s	n	n
Protein translocase subunit SecY	*Bifidobacterium bifidum*	*secY*	13.86	56.45	n	n	s	n
Protein UL87	*Human cytomegalovirus*	*UL87*	13.58	76.67	n	s	n	n

Annex 13: **Continuation.**

Name	*Organism*	*Gene*	Mass (kDa)	Confidence score	Biomarker	iST	Enrich	Add-on
Prothrombin	*Homo sapiens*	*F2*	14.09	58.23	s	s	n	n
Prothrombin	*Myotis lucifugus*	*F2*	12.93	121.05	s	s	s	s
Prothrombin	*Pan paniscus*	*F2*	14.35	23.43	s	n	s	s
PTS fructose transporter subunit IIC	*Dorea formicigenerans*	*DXD84_00040*	13.34	85.58	n	n	n	s
PTS fructose transporter subunit IIC	*Firmicutes bacterium AF36-19BH*	*DWZ97_03050*	13.43	136.11	n	s	n	n
PTS trehalose transporter subunit IIBC	*Megamonas rupellensis*	*DWZ11_06425*	13.03	87.68	n	n	n	s
Rap guanine nucleotide exchange factor 2	*Callorhinchus milii*	*RAPGEF2*	13.87	93.47	s	n	n	s
Reticulocyte-binding protein homolog 1	*Plasmodium falciparum (isolate 3D7)*	*RH1*	13.39	77.25	n	n	s	n
Retinol-binding protein	*Pan troglodytes*	*RBP4*	14.28	142.31	s	n	n	s
Reverse transcriptase (Fragment)	*Human immunodeficiency virus 1*	*pol*	13.80	81.31	n	n	n	s
Selenoprotein P	*Homo sapiens*	*SELENOP*	13.43	81.12	n	n	s	s
Serotransferrin	*Homo sapiens*	*TF*	14.53	108.59	s	s	s	s
Serotransferrin	*Mus musculus*	*TF*	12.66	57.84	s	s	n	n
Serpin family G member 1	*Pan paniscus*	*SERPING1*	14.01	67.65	s	s	s	s
Serum amyloid A-4 protein	*Homo sapiens*	*SAA4*	13.22	83.08	n	s	s	s
Serum amyloid P-component	*Homo sapiens*	*APCS*	13.73	50.87	s	s	s	s
Serum paraoxonase/arylesterase 1	*Homo sapiens*	*PON1*	13.60	96.92	s	s	s	s
Spectrin beta, non-erythrocytic 4	*Pan troglodytes*	*SPTBN4*	13.51	123.07	s	n	n	s
SPG11 vesicle trafficking associated, spatacsin	*Pan troglodytes*	*SPG11*	12.84	67.36	s	s	n	s
Sugar transporter	*Bacteroides sp. OM05-12*	*DXB63_12185*	13.79	85.50	n	s	n	n
Symplekin scaffold protein	*Homo sapiens*	*SYMPK*	13.80	34.21	n	s	n	s
Talin 2a	*Danio rerio*	*tln2a*	14.24	60.44	n	s	n	n
T-cell surface glycoprotein CD4-like isoform X1	*Pantherophis guttatus*	*LOC117661814*	13.36	48.17	n	s	n	n
Teneurin transmembrane protein 4	*Pan troglodytes*	*TENM4*	12.77	131.41	s	n	n	s
Tetranectin	*Homo sapiens*	*CLEC3B*	13.67	30.45	s	n	n	s
Thyroxine-binding globulin	*Homo sapiens*	*SERPINA7*	51.68	1144.58	n	s	n	s
Titin	*Danio rerio*	*ttn.2*	39.33	727.15	n	n	n	s
Transient receptor potential cation channel subfamily M member 3	*Homo sapiens*	*TRPM3*	39.33	693.18	n	s	n	n
Translocation/assembly module TamB	*Parabacteroides sp*	*HP046_01095*	39.33	604.99	n	s	n	n
Transthyretin	*Homo sapiens*	*TTR*	53.01	1071.92	s	s	s	s
TRAP transporter permease DctM/Q	*Desulfovibrionaceae bacterium*	*DBY37_02230*	53.01	999.34	n	s	n	n
UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	*Flavobacteriales bacterium*	*murG*	53.01	873.67	n	n	n	s
Uncharacterized protein DKFZp686C02220 (Fragment)	*Homo sapiens*	*DKFZp686C02220*	14.38	143.96	-	s	n	n
Uncharacterized protein DKFZp686K18196 (Fragment)	*Homo sapiens*	*DKFZp686K18196*	49.82	591.09	-	n	s	s
V1-7 protein (Fragment)	*Homo sapiens*	*V1-7*	56.76	184.46	-	s	n	s
V2-17 protein (Fragment)	*Homo sapiens*	*V2-17*	56.76	299.29	-	s	n	s
VH6DJ protein (Fragment)	*Homo sapiens*	*VH6DJ*	44.52	705.34	n	n	n	s
Virion infectivity factor	*Human immunodeficiency virus 1*	*vif*	44.52	721.48	n	s	n	s
Vitamin D-binding protein	*Homo sapiens*	*GC*	52.52	1161.79	n	s	s	s
Vitamin K-dependent protein S (Fragment)	*Homo sapiens*	*PROS1*	52.52	743.60	s	s	n	s
Vitronectin	*Homo sapiens*	*VTN*	52.52	817.16	s	s	s	s
YihY family inner membrane protein	*Parabacteroides distasonis*	*yihY*	11.27	164.96	n	n	s	n
YihY/virulence factor BrkB family protein	*Streptomyces sp. CB04723*	*HXS80_04895*	18.54	159.99	n	n	n	s

Annex 14: **Theorical data.** This table contains data on the 350 proteins identified across all the different kits, the information was extracted from UNIPROT and Expasy (ProtParam). The "-" indicates absence of data.

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
2-hydroxyacylsphingosine 1-beta-galactosyltransferase	Q16880	59.19	9.51
ABC transporter	A0A2K2U6D1	20.50	8.47
ABC transporter ATP-binding protein	A0A086Y7S0	-	-
ABC transporter ATP-binding protein	A0A3R6IAG7	26.26	6.58
ABC transporter permease	A0A417T6M6	13.45	7.89
ABC3 transporter permease protein domain-containing protein	A0A316R183	12.88	9.52
ABC-type glutathione-S-conjugate transporter	G1PLU4	24.52	6.09
ACX73 (Fragment)	A0A679K5L3	10.54	4.78
Afamin	P43652	66.57	5.58
Albumin	P02768	66.47	5.67
Albumin	G1P5R0	66.25	5.82
Alpha-1-acid glycoprotein 1	P02763	21.58	5.11
Alpha-1-acid glycoprotein 2	P19652	21.65	5.12
Alpha-1-antichymotrypsin	P01011	45.26	5.32
Alpha-1-antitrypsin	P01009	44.32	5.37
Alpha-1B-glycoprotein	P04217	51.92	5.63
Alpha-2-antiplasmin	P08697	50.45	5.87
Alpha-2-glycoprotein 1. zinc-binding	H2QV12	32.21	5.71
Alpha-2-HS-glycoprotein	P02765	30.23	4.53
Alpha-2-macroglobulin	A8K2U0	159.38	5.5
Amino acid transporter (Fragment)	A0A087WT87	2.387	8.59
Angiotensinogen	P01019	49.76	5.6
Anti-FactorVIII scFv (Fragment)	A2KBC6	9.73	6.33
Antithrombin-III	P01008	49.03	5.95
Apolipoprotein A-I	P02647	28.96	5.45
Apolipoprotein A-I	P68292	28.05	5.43
Apolipoprotein A-II	P02652	8.70	5.05
Apolipoprotein A-IV	P06727	43.37	5.18
Apolipoprotein A-IV	A0A2R9BIY1	43.45	5.18
Apolipoprotein B-100	P04114	512.79	6.57
Apolipoprotein C-I	P02654	6.63	7.93
Apolipoprotein C-II	P02655	8.91	4.58
Apolipoprotein C-II	A0A2R8ZTC4	50.99	8.51
Apolipoprotein C-III	P02656	8.76	4.72
Apolipoprotein D	P05090	19.30	5.2
Apolipoprotein E	P02649	34.23	5.52
Apolipoprotein E	A0A2R9B3N2	34.18	5.46
Apolipoprotein E	P0DMM1	34.38	5.31
Apolipoprotein L1	O14791	41.12	5.49
Apolipoprotein M	O95445	21.25	5.66
ATP synthase subunit b	A0A316QJ11	2.77	8.75
ATP synthase subunit delta	A0A7L9VNH2	-	-
ATP-binding cassette domain-containing protein	A0A3E4U6D3	2.27	5.49

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
ATP-dependent zinc metalloprotease FtsH	A0A0V8QC33	2.23	5.52
ATP-dependent zinc metalloprotease FtsH	A0A1D7YTZ1	2.14	5.57
ATP-dependent zinc metalloprotease FtsH	A0A396K6J1	2.67	5.52
ATP-dependent zinc metalloprotease FtsH	-	-	-
ATP-dependent zinc metalloprotease FtsH	A0A3C1LWI7	2.53	5.52
Beta-2-glycoprotein 1	P02749	36.25	8.37
BMP family ABC transporter substrate-binding protein	-	-	-
C2 domain-containing protein	C1GVV5	1.88	5.52
C4b-binding protein alpha chain	P04003	61.67	6.24
Cadherin domain protein	W2T4I0	274.51	5.43
Cadherin domain-containing protein	A0A328EI09	156.47	4.98
Calcium-transporting ATPase	A0A231HF60	3.25	4
Carboxypeptidase N catalytic chain	P15169	50.03	6.88
Carboxypeptidase N subunit 2	P22792	58.22	5.54
Carboxypeptidase Q (Fragment)	A0A1B2Z7C9	20.36	5.31
Cd(2+)-exporting ATPase	A0A3R6LZR9	2.05	4
CD5 antigen-like	O43866	36.05	5.21
CD5 molecule like	H2Q0B2	36.17	5.34
CDK5 regulatory subunit-associated protein 2	Q9JLH5	215.48	5.31
cDNA FLJ25298 fis. clone STM07683. highly similar to Protein Tro alpha1 H.myeloma	Q96DK0	51.33	6.22
Centrosomal protein of 83 kDa	F1R4Y7	83.91	5.66
Ceruloplasmin	P00450	120.09	5.41
chitin synthase	A0A0C2J7L4	86.29	6.14
Clusterin	E7ETB4	29.02	5.19
Clusterin	A0A2R9AC29	49.96	5.88
Coagulation factor X	P00742	50.33	5.38
Coagulation factor XII	P00748	39.64	9
Complement C1r subcomponent	P00736	78.21	5.76
Complement C1r subcomponent-like protein	Q9NZP	-	-
Complement C1s subcomponent	P09871	74.88	4.85
Complement C2	P06681	81.08	7.57
Complement C3	P01024	184.95	6
Complement C4-A	P0C0L4	71.67	8.69
Complement C4-B	P0C0L5	71.67	8.69
Complement C4-B	P01029	72.07	8.7
Complement C5	P01031	73.29	5.01
Complement component 4 binding protein alpha	H2Q114	6.96	6.99
Complement component C6	P13671	102.41	6.17
Complement component C6	P61134	102.57	6.16
Complement component C7	P10643	91.11	6.09
Complement component C8 alpha chain	P07357	61.71	5.74
Complement component C8 beta chain	P07358	60.94	7.65
Complement component C8 beta chain	A0A2R9C373	63.62	8.25
Complement component C9	P48770	59.83	5.32
Complement component C9	P02748	60.97	5.42

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
Complement factor B	P00751	83.01	6.66
Complement factor H	P08603	137.05	6.12
Complement factor H-related protein 1	Q03591	35.73	7.1
VComplement factor H-related protein 2	P36980	28.73	5.8
Complement factor I	P05156	63.48	7.38
Copper-transporting ATPase	Q8I3A0	2.59	8.5
Coronin	A0A2R9A938	7.19	9.51
Corticosteroid-binding globulin	P08185	42.63	5.64
C-reactive protein	P02741	23.04	5.28
Cytochrome bd-I ubiquinol oxidase subunit CydA	A0A0F7J976	2.61	4
Diadenylate cyclase	A0A3R7AWH3	18.03	5.07
diguanylate cyclase	A0A2M9GKG0	2.61	8.43
DIRAS family GTPase 3	G1U8U1	-	-
DNA topoisomerase 4 subunit A	Q4L663	51.46	5.91
DNA translocase FtsK	-	-	-
DNA translocase FtsK	A0A5B3GXK5	22.93	9.41
Dystonin	Q03001	860.66	5.14
Dystonin	A0A5F9C4B5	4.33	9.82
Dystrophin-related protein 2	Q05AA6	108.05	5.84
Ectonucleotide pyrophosphatase/phosphodiesterase family member 1	P06802	103.17	6.13
Efflux pump membrane transporter	A0A8B6X3C9	2.26	5.97
Env polyprotein (Fragment)	D2ECD4	22.39	7.74
Envelope glycoprotein gp160	P05877	93.45	8.87
Envelope glycoprotein gp160 (Fragment)	P05877	93.45	8.87
Envelope glycoprotein gp160 (Fragment)	P05877	93.45	8.87
Erythrocyte membrane protein 1. PfEMP1	Q8I098	3.78	8.13
ESX-3 secretion system protein EccB3	P9WNQ3	47.94	9.03
Exocyst complex component 8	A4IG64	12.15	6.79
Exocyst complex component SEC5	A0A397GMB0	104.66	5.98
Ferric siderophore transport system protein ExbB	-	-	-
ferroxidase	H2QNJ9	120.45	5.36
Fetuin-B	Q9UGM5	40.48	6.52
Fibrinogen alpha chain	P02671	91.35	5.79
Fibrinogen beta chain	P02675	50.76	7.95
Fibrinogen beta chain	G1P9V0	52.75	6.03
Fibrinogen beta chain	P14478	-	-
Fibrinogen beta chain	A0A2R9AN20	52.29	7.6
Fibrinogen gamma chain	H0VBM5	47.91	5.37
Fibrinogen gamma chain	P02679	48.48	5.24
Fibrinogen gamma chain	G1P505	50.76	7.95
Fibronectin	P02751	269.25	5.25
Fibronectin	G1NWI0	269.81	5.45
Ficolin-3	O75636	30.35	6.22
Galectin-3-binding protein	Q08380	63.27	5.07
GCT-A4 heavy chain variable region (Fragment)	A0A0X9USM3	10.65	8.04

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
Gelsolin	A0A2R9AWE7	8.31	5.61
Glutathione peroxidase 3	P22352	23.46	7.85
Glyceraldehyde-3-phosphate dehydrogenase	A0A2J8JEW8	16.14	8.68
Glycosyl/glycerophosphate transferase	A0A2S9QMS4	-	-
Golgin-45	A0A4Y0BUS8	4.50	5.05
HAD family hydrolase	A0A373GT13	2.20	5.49
HAD family hydrolase	A0A373VN31	2.11	5.49
HAD-IC family P-type ATPase	A0A7W4E4M9	2.26	5.24
Haptoglobin	P00738	43.34	6.13
Haptoglobin	G1PZC5	36.77	5.7
Haptoglobin-related protein	P00739	39.02	6.63
HAUS augmin-like complex subunit 3	Q68CZ6	69.51	5.39
HAUS augmin-like complex subunit 8	Q9BT25	44.72	6.67
Heme ABC transporter ATP-binding protein	A0A316PNG9	26.42	5.91
Hemoglobin subunit alpha	P69905	15.12	8.73
Hemoglobin subunit beta	P68871	15.86	6.81
Hemoglobin. beta adult 2	Q6ZM12	15.60	8.77
Hemopexin	P02790	49.29	6.43
Heparin cofactor 2	P05546	54.96	6.26
histidine kinase	A0A923NV46	11.92	5.61
Histidine-rich glycoprotein	P04196	57.65	7.03
Hyaluronan-binding protein 2	Q14520	32.90	5.2
IBM-B2 heavy chain variable region (Fragment)	A0A125QYY9	11.71	8.7
IG c102_heavy_IGHV3-48_IGHD1-26_IGHJ5 (Fragment)	A0A5C2GET1	12.51	7
IG c1037_heavy_IGHV4-34_IGHD3-22_IGHJ5 (Fragment)	A0A5C2GWT2	-	-
IG c1064_heavy_IGHV3-73_IGHD6-6_IGHJ6 (Fragment)	A0A5C2GF55	-	-
G c1247_heavy_IGHV1-69_IGHD1-1_IGHJ6 (Fragment)	A0A5C2GXC2	-	-
IG c1355_heavy_IGHV4-4_IGHD2-2_IGHJ1 (Fragment)	A0A5C2GPN1	-	-
IG c1405_heavy_IGHV3-7_IGHD3-3_IGHJ4 (Fragment)	A0A5C2GPS8	-	-
IG c1447_heavy_IGHV3-15_IGHD3-3_IGHJ5 (Fragment)	A0A5C2GXV5	-	-
IG c1609_heavy_IGHV3-21_IGHD3-10_IGHJ4 (Fragment)	A0A5C2GUY1	-	-
IG c1654_heavy_IGHV3-30_IGHD3-10_IGHJ4 (Fragment)	A0A5C2GSI8	-	-
IG c198_heavy_IGHV5-51_IGHD7-27_IGHJ6 (Fragment)	A0A5C2GIF2	-	-
IG c2_heavy_IGHV6-1_IGHD2-2_IGHJ4 (Fragment)	A0A5C2GHR6	-	-
IG c255_heavy_IGHV3-15_IGHD3-10_IGHJ4 (Fragment)	A0A5C2GDV4	-	-
IG c366_heavy_IGHV3-7_IGHD1-26_IGHJ4 (Fragment)	A0A5C2GLY	-	-
IG c396_heavy_IGHV4-34_IGHD2-2_IGHJ5 (Fragment)	A0A5C2GM13	-	-
IG c488_heavy_IGHV3-49_IGHD3-10_IGHJ6 (Fragment)	A0A5C2GJP1	-	-
IG c507_heavy_IGHV3-7_IGHD3-3_IGHJ4 (Fragment)	A0A5C2GHW9	-	-
IG c543_heavy_IGHV3-9_IGHD5-24_IGHJ4 (Fragment)	A0A5C2GI29	-	-
IG c55_heavy_IGHV3-48_IGHD3-10_IGHJ3 (Fragment)	A0A5C2GGM0	-	-
IG c574_heavy_IGHV4-61_IGHD3-16_IGHJ6 (Fragment)	A0A5C2GJ11	10.44	9.04
IG c720_heavy_IGHV3-49_IGHD3-3_IGHJ6 (Fragment)	A0A5C2GKE3	10.72	7
IG c733_heavy_IGHV3-7_IGHD3-9_IGHJ5 (Fragment)	A0A5C2GPG2	10.68	9.14
IG c828_heavy_IGHV4-61_IGHD3-22_IGHJ3 (Fragment)	A0A5C2GGM6	12.53	8.91

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
IG c855_heavy_IGHV3-23_IGHD1-26_IGHJ1 (Fragment)	A0A5C2GMZ5	10.50	5.21
Ig heavy chain variable region (Fragment)	-	-	-
IgG H chain	-	-	-
IGH + IGL c20_heavy_IGHV3-21_IGHD3-22_IGHJ3 (Fragment)	A0A5C2GB38	10.19	5.19
IGH + IGL c326_heavy_IGHV1-69_IGHD2-2_IGHJ5 (Fragment)	A0A5C2GAP1	8.88	9.07
IGH + IGL c336_heavy_IGHV3-23_IGHD3-22_IGHJ5 (Fragment)	A0A5C2GAQ1	10.52	8.72
IGH + IGL c38_heavy_IGHV3-15_IGHD1-14_IGHJ4 (Fragment)	A0A5C2GBE5	12.74	8.67
IGH + IGL c392_heavy_IGHV1-18_IGHD3-9_IGHJ6 (Fragment)	A0A5C2GBI3	10.67	9.2
IGH + IGL c517_heavy_IGHV2-70_IGHD4-17_IGHJ4 (Fragment)	A0A5C2GEN2	10.99	9.33
IGH + IGL c534_heavy_IGHV4-39_IGHD3-22_IGHJ4 (Fragment)	A0A5C2G9Q	0.44	5.56
IGH + IGL c62_heavy_IGHV3-7_IGHD3-22_IGHJ5 (Fragment)	A0A5C2GAB7	10.61	5.24
IGH c1126_heavy_IGHV3-11_IGHD6-13_IGHJ5 (Fragment)	A0A7S5ETU8	11.81	9.44
IGH c1340_heavy_IGHV3-7_IGHD2-2_IGHJ6 (Fragment)	A0A7S5C11	-	-
IGH c1503_heavy_IGHV3-49_IGHD1-1_IGHJ6 (Fragment)	A0A7S5EXD1	10.69	6.86
IGH c1758_heavy_IGHV3-33_IGHD3-16_IGHJ6 (Fragment)	A0A7S5EXJ8	10.77	8.98
IGH c1796_heavy_IGHV4-39_IGHD3-10_IGHJ3 (Fragment)	A0A7S5C0F0	10.41	8.62
IGH c2097_heavy_IGHV3-7_IGHD6-19_IGHJ6 (Fragment)	A0A7S5C1D9	10.85	6.86
IGH c2207_heavy_IGHV4-4_IGHD5-24_IGHJ5 (Fragment)	A0A7S5EWK6	10.27	9.07
IGH c235_heavy__IGHV3-23_IGHD3-10_IGHJ5 (Fragment)	A0A5C2GAD3	10.46	6.43
IGH c2826_heavy_IGHV5-51_IGHD5-24_IGHJ4 (Fragment)	A0A7S5C2F8	11.02	6.85
IGH c283_heavy__IGHV3-30_IGHD6-19_IGHJ4 (Fragment)	A0A5C2GAG8	10.61	8.66
IGH c3844_heavy_IGHV3-74_IGHD6-19_IGHJ4 (Fragment)	A0A7S5C661	10.55	9.11
IGH c3886_heavy_IGHV3-15_IGHD2-15_IGHJ4 (Fragment)	A0A7S5EVS2	10.82	6.46
IGH c395_heavy__IGHV3-7_IGHD1-14_IGHJ4 (Fragment)	A0A5C2GAY9	10.79	4.4
IGH c4042_heavy_IGHV4-34_IGHD5-12_IGHJ6 (Fragment)	A0A7S5C546	13.59	8.91
IGH c4137_heavy_IGHV5-51_IGHD5-12_IGHJ4 (Fragment)	A0A7S5C5N4	10.61	6.44
IGH c557_heavy__IGHV4-34_IGHD3-9_IGHJ4 (Fragment)	A0A5C2G7R8	9.09	7.74
IGH c776_heavy_IGHV3-7_IGHD3-22_IGHJ6 (Fragment)	A0A7S5BYP6	10.74	6.86
IGH c817_heavy_IGHV3-7_IGHD7-27_IGHJ4 (Fragment)	A0A7S5BZ26	10.56	6.47
IGH c905_heavy_IGHV1-18_IGHD3-22_IGHJ3 (Fragment)	A0A7S5C040	10.77	8.98
IGH@ protein	Q99856	62.88	4.84
IGHG3 (Fragment)	P01860	-	-
Ig-like domain-containing protein	Q6N095	50.17	8.87
Ig-like domain-containing protein	-	-	-
Immunoglobulin alpha-2 heavy chain	P01877	-	-
Immunoglobulin delta heavy chain	P01859	-	-
Immunoglobulin heavy constant gamma 2	P01859	-	-
Immunoglobulin heavy constant mu	P01871	-	-
Immunoglobulin heavy variable 3-72	A0A0B4J1Y9	11.05	8.66
Immunoglobulin J chain	P01591	15.59	4.59
Immunoglobulin kappa constant	P01834	-	-
Immunoglobulin kappa light chain	A2NJV5	10.93	6.06
Immunoglobulin kappa variable 2-24	A0A0C4DH68	10.99	8.01
Immunoglobulin kappa variable 2-28	A0A075B6P5	10.90	5.04
Immunoglobulin kappa variable 2-30	P06310	11.06	8

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
Immunoglobulin kappa variable 3D-11	A0A0A0MRZ8	10.40	6.85
Immunoglobulin kappa variable 3D-20	A0A0C4DH25	10.26	4.77
Immunoglobulin lambda constant 2	P0DOY2	-	-
Immunoglobulin lambda constant 7	A0M8Q6	10.21	8.75
Immunoglobulin lambda variable 2-8	P01709	10.35	5.61
Immunoglobulin lambda-1 light chain	P15814	19.13	9.97
Immunoglobulin mu heavy chain	P0DOX6	63.48	8.12
Insulin-like growth factor-binding protein complex acid labile subunit	P35858	63.24	6.13
Inter-alpha (Globulin) inhibitor H4 (Plasma Kallikrein-sensitive glycoprotein)	Q14624	70.58	5.92
Inter-alpha (Globulin) inhibitor H4 (Plasma Kallikrein-sensitive glycoprotein) variant (Fragment)	Q59FS1	14.20	8.27
Inter-alpha-trypsin inhibitor heavy chain 3	A0A2I3TUE8	96.18	6.05
Inter-alpha-trypsin inhibitor heavy chain 4	A0A2J8P849	100.73	6.29
Inter-alpha-trypsin inhibitor heavy chain H1	P19827	71.41	6.33
Inter-alpha-trypsin inhibitor heavy chain H2	P19823	72.45	5.75
Inter-alpha-trypsin inhibitor heavy chain H3	Q06033	69.36	5.01
Inter-alpha-trypsin inhibitor heavy chain H4	G1TY29	97.74	6.52
Ion-translocating oxidoreductase complex subunit C	A0A7C6S4J6	3.30	6.13
IQGAP1 protein	A0A0F7C9F3	13.64	7.07
Iron ABC transporter permease	A0A412KYP3	22.46	6.91
Junction plakoglobin	P14923	81.74	5.75
Kallistatin	P29622	46.35	7.88
Keratin 1	A5A6M6	65.48	7.61
Keratin. type I cytoskeletal 18	P05784	47.40	5.22
Keratin. type I cytoskeletal 19	A0A286XP79	35.51	4.81
Keratin. type II cytoskeletal 1	P04264	65.90	8.15
Keratin. type II cytoskeletal 1	A5A6M6	65.48	7.61
Keratin. type II cytoskeletal 5	Q922U2	61.76	7.59
Keratin. type II cytoskeletal 7	P08729	51.25	5.39
Kininogen-1	P01042	69.89	6.23
L-amino acid oxidase	P0DQH9	-	-
Large-conductance mechanosensitive channel	A0A373QAU5	2.45	8.22
Lauroyl acyltransferase	A0A2L1GQW2	2.13	5.57
Leucine-rich alpha-2-glycoprotein	P02750	34.34	5.66
Lumican	P51884	36.66	6.17
Mechanosensitive channel MscK	A0A2K9C763	119.79	8.37
Melatonin receptor type 1A	G1U149	26.69	9.1
Membrane metalloprotease	A0A0S2W3Q5	15.68	7.99
Membrane-associated guanylate kinase. WW and PDZ domain-containing protein 1	Q96QZ7	164.58	7.3
MFS transporter	A0A2V2ESD5	2.91	5.52
Miniconductance mechanosensitive channel MscM	-	-	-
MMPL family transporter (Fragment)	-	-	-
Multifunctional fusion protein	A0A415LAD1	17.59	4.76
Na/Pi cotransporter	A0A2V2CQG3	9.47	5.88
Na+/H+ antiporter subunit A	-	-	-
N-acetylmuramoyl-L-alanine amidase	Q96PD5	59.98	7.64

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
Neuraminidase	Q9IGQ6	51.40	5.53
Nuclear transcription factor. X-box binding 1	A0A2R8ZXQ4	6.47	8.51
Obscurin. cytoskeletal calmodulin and titin-interacting RhoGEF	A6NGQ3	9.79	8.37
Oligosaccharide flippase family protein	A0A6L4X2X5	2.87	5.49
Oligosaccharide flippase family protein	A0A8T3U2I9	2.23	8.76
Otoferlin isoform X3	A0A8M9P0R6	15.63	5.47
Outer capsid protein VP4 (Fragment)	P11193	87.69	6.12
Phenylalanine-specific permease	A0A0H2YWP8	1.82	5.52
Phosphatidate cytidylyltransferase	A0A7K3TFS2	2.72	5.88
Phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta	H0V6Q3	18.27	8.37
Phosphatidylinositol-glycan-specific phospholipase D	P80108	89.81	5.78
Phosphoethanolamine transferase	-	-	-
Phosphorylase b kinase regulatory subunit	G7Y3H2	14.45	8.74
Phosphotransferase RcsD	A0A8B2XSM9	23.62	5.47
Pigment epithelium-derived factor	P36955	44.38	5.9
Plasma kallikrein	P03952	41.39	8.66
Plasminogen	P00747	88.43	7.08
Plasminogen	G1PGD5	88.28	7.72
Plexin A4	A0A4W3IPZ4	205.37	6.14
Podocin	A4FVJ0	18.39	7.8
Pol protein (Fragment)	P04585	161.91	8.88
Polyprotein P1234	A0A7D5G7K3	25.66	8.66
Polysaccharide biosynthesis protein	A0A3E3I2A2	15.31	9.8
Polysaccharide biosynthesis protein	A0A417JRY1	2.52	5.27
Polysaccharide biosynthesis protein C-terminal domain-containing protein	A0A316LWF2	14.90	9.42
Pregnancy zone protein	P20742	161.05	5.91
Pregnancy-associated plasma protein A. pappalysin 1b isoform X1	A0A8M3ASU7	180.43	6.2
Probable non-functional immunoglobulin kappa variable 3-7	A0A075B6H7	10.33	7.99
Protein AMBP	P02760	20.84	6.13
Protein AMBP	A0A2R8Z833	38.44	8.28
Protein AMBP	A0A2I3RER0	37.15	5.84
Protein S100	A0A2R9A6I5	4.31	4.55
Protein S100-A8	P05109	10.83	6.5
Protein translocase subunit SecA	F7UV48	68.68	5.01
Protein translocase subunit SecA	A0A291T8Q6	73.71	5.48
Protein translocase subunit SecY	A0A415C5S3	2.11	9.99
Protein UL87	P16730	104.80	9.39
Prothrombin	P00734	65.30	5.23
Prothrombin	G1PXB6	67.87	6
Prothrombin	A0A2R9C6W7	67.57	5.48
PTS fructose transporter subunit IIC	A0A3E4F9Y5	35.94	8.96
PTS fructose transporter subunit IIC	A0A416CSC7	15.62	4.55
PTS trehalose transporter subunit IIBC	A0A411ZRF1	39.68	8.23
Rap guanine nucleotide exchange factor 2	A0A4W3HUB2	11.20	4.53
Reticulocyte-binding protein homolog 1	P86148	355.16	8.01

Annex 14: **Continuation.**

Name	Code UNIPROT	Theoretical mass (kDa)	p.I
Retinol-binding protein	P61641	21.07	5.27
Reverse transcriptase (Fragment)	P12497	64.28	8.54
Selenoprotein P	P49908	41.23	7.95
Serotransferrin	P02787	75.18	6.7
Serotransferrin	Q921I1	74.88	6.81
Serpin family G member 1	A0A2R9BXU2	52.47	6.4
Serum amyloid A-4 protein	P35542	12.80	9.07
Serum amyloid P-component	P02743	23.25	6.12
Serum paraoxonase/arylesterase 1	P27169	39.60	5.08
Spectrin beta. non-erythrocytic 4	A0A2J8QFZ1	12.29	7.09
SPG11 vesicle trafficking associated. spatacsin	A0A2I3T9P4	33.76	7.72
Sugar transporter	A0A374WAE9	2.36	10.84
Symplekin scaffold protein	Q92797	141.14	5.82
Talin 2a	A0A0R4IIA7	36.95	8.66
T-cell surface glycoprotein CD4-like isoform X1	A0A6P9BFZ3	48.69	8.96
Teneurin transmembrane protein 4	H2Q4H4	37.47	7.24
Tetranectin	P05452	20.13	5.8
Thyroxine-binding globulin	P05543	44.10	5.76
Titin	A0A8M9QKG2	9.45	5.01
Transient receptor potential cation channel subfamily M member 3	Q9HCF6	2.55	5.95
Translocation/assembly module TamB	A0A942AG48	-	-
Transthyretin	P02766	13.76	5.31
TRAP transporter permease DctM/Q	A0A316N0F4	44.29	8.5
UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	A5FIY3	39.93	9.18
Uncharacterized protein DKFZp686C02220 (Fragment)	Q6N091	50.11	5.78
Uncharacterized protein DKFZp686K18196 (Fragment)	Q6N092	9.89	9897.14
V1-7 protein (Fragment)	-	-	-
V2-17 protein (Fragment)	-	-	-
VH6DJ protein (Fragment)	A2N0S6	10.24	8.32
Virion infectivity factor	P12504	22.69	9.93
Vitamin D-binding protein	P02774	51.19	5.16
Vitamin K-dependent protein S (Fragment)	P07225	70.64	5.17
Vitronectin	P04004	52.27	5.47
YihY family inner membrane protein	A0A069SE16	2.49	5.52
YihY/virulence factor BrkB family protein	A0A7H8ZK65	2.40	5.57