UNIVERSITY OF PADOVA

MASTER THESIS

Quantifying immune contexture of tumors using imaging and sequencing data

Author:
Antonio COLLESEI

Advisor:
Prof. Barbara DI CAMILLO
Co-advisor(s):
Dipl.-Ing. Dr.techn.
Univ.-Prof. Zlatko TRAJANOSKI
Dr. Francesca FINOTELLO

Master Degree in Bioengineering

Department of Information Engineering

December 4, 2017 ~ A.Y. 2016-17

"*Many things cause terror and wonder, yet nothing is more terrifying and wonderful than man. [...] Possessing means of invention, a skillfulness beyond expectation, now toward evil he moves, now toward good. By integrating the laws of the earth and justice under oath sworn to the gods, he is lofty of city. Citiless is the man with whom ignobility because of his daring dwells. May he never reside at my hearth or think like me, whoever does such things.*"

Sophocles, *Antigone*, vv. 332-333, 365-375

Abstract

Antonio COLLESEI

Quantifying immune contexture of tumors using imaging and sequencing data

Cancer treatment has always been a main issue in medicine, though, thanks to the recent discovery and involvement of novel technologies able to amplify the research spectrum of action, the efforts have been increased towards precision oncology. Cancer immunology is a promising approach, supposed to reduce the invasiveness for the patient due to its peculiarity of eliciting the human immune response to act selectively against the tumor in a personalized way for each single patient, depending on his genetic tumor environment. The tumor immune contexture, defined as the type and density of tumor-infiltrating immune cells, has a strong impact on patients' prognosis and response to therapy. In particular, high infiltration of cytotoxic CD8+ T cells, which has the capability of recognize and kill tumor cells, has been associated with a good prognosis. Thus, the quantification of CD8+ T cells is of paramount importance for the stratification of cancer patients and for the development of effective combination therapies.

In this thesis, two techniques are described and combined experimentally to quantify CD8+ T cells and their subtypes: RNA sequencing (RNA-Seq) and immunohistochemistry (IHC) imaging. First, we developed a deconvolution algorithm based on support-vector regression and on a novel signature matrix, which computes the relative cell fractions of naïve, central memory, and effector memory CD8+ T cells from

bulk tumor RNA-seq data. After validation with simulated data, the deconvolution method was applied to published data from glioblastoma tumors, proving its applicability to data from both fresh-frozen and archived tumor samples. Second, we optimized a bioinformatics pipeline for the quantification of total CD8+ T cells from immunohistochemistry images of human tumors and we applied it to the analysis of 29 melanoma samples. Given the availability of large collections of archived tumor samples exploited for image pathology and, more recently, for RNA-seq, these methods represent valuable tools for the efficient extraction of immunological features from tumor data and may provide mechanistic insights for the optimization of anti-cancer treatments.

Acknowledgements

I would first like to thank my thesis advisor Prof. *Barbara Di Camillo* of the Department of Information Engineering at the University of Padova. She inspired my preferences in the variegated Bioengineering universe throughout her genomics lectures and gave me the possibility to achieve a higher level of knowledge during my Erasmus period at the Medizinische Universität in Innsbruck.

I would also like to thank Prof. *Zlatko Trajanoski*, head of the Division of Bioinformatics at the Biocenter in Innsbruck, who welcomed and treated me as a colleague, and not as a simple Master student, and Dr. *Francesca Finotello*, who, despite a few (or more) misunderstandings, introduced me to the real fast and competitive research world and offered me a complementary perspective on handling the scientific production.

I would like to thank as well *Clemens Mayer* and *Christina Plattner*, who were able to lighten the daily work with the historical temperature war, during our months of office sharing.

Finally, I must express my very profound gratitude to my parents, *Italo* and *Cinzia*, who always kept my confidence up even when I was under pressure, and to my brother *Eugenio* that, due to my thesis semester abroad, was afraid of a possible decision of mine to bring GDP to a country different from Italy, forever.

Thank you all for easing my growth. Not physical, of course. That is encoded in my genes and I guess 1.93 metres should be enough.

Antonio Collesei

Contents

List of Figures

List of Tables

To Governing Dynamics.

Chapter 1

Introduction

1.1 Cancer immunotherapy

The immune system is a collection of organs, cells, and substances that help protecting the human body from infections and other general diseases.

It keeps track of all the substances normally found in the body. Any new substance that the immune cell, for instance the T lymphocyte, does not recognize as part of the so-called *self-environment* raises an alarm, urging the body to attack it. For example, bacteria produce substances, such as certain proteins, that are not normally found in the human body. The immune system recognizes these substances as *foreign* and attacks steadily with the intention of destroying them. The immune response is not exclusively targeted to the foreign substance, but has a wider action that leads to the destruction of the whole intruder.

Nonetheless, how can the immune system avoid attacking the body itself? Recalling a study of the Pennsylvania School of Medicine (Olenchock et al. 2006), small fatty acids called diacylglycerols (DAGs), and the enzymes that metabolize them, are critical players in the molecular pathway that leads to activation of the immune system. In fact, understanding under which conditions the immune cells get activated is crucial to fight either cancer, or autoimmune diseases, or moreover organ transplant rejection. They found out that when DAGs are chemically modified by enzymes called

diacylglycerol kinases (DGKs), T cells become unresponsive to foreign substances and to self-environment. The discovery was made by studying mice that had been engineered to lack DGKs. When DAGs could not be chemically altered because the DGKs were absent, the T cells were hyperreactive to foreign antigens and could not be made tolerant to host cells.

This preamble leads to the fact that, potentially, the immune system has the ability to target and therefore destroy cancer cells (Pardoll 1999). Clearly there are limits on the immune system's ability to fight cancer on its own, otherwise immunotherapy would have no matter of existance. Sometimes the immune system cannot detect the cancer cells as foreign because the difference between them and normal cells is not so striking. Sometimes the immune system is even able to recognize the cancer cells, but the response might not be strong enough to get rid of the invasion. Cancer cells themselves could as well contribute giving off substances that inhibit the immune system response.

Accordingly, the premise that has led to the development of cancer immunotherapy is that the immune system owns the instruments to fight cancer, but it needs a push. As it was mentioned above, the main characters in this action are T-cells: they have receptors on their surface acting as gas pedals or brakes of activation and inhibition.

There are three types of cancer immunotherapy:

- Monoclonal antibodies, useful in treating cancer because they can be designed to attack a very specific part of a cancer cell;
- Immune checkpoint inhibitors, which block the ability of certain proteins, called *immune checkpoint proteins*, to limit the strength and duration of immune responses; researchers learnt that tumors can pilot these proteins and use them to suppress immune responses, so blocking the activity of immune checkpoint proteins releases the *brakes* on the immune system, increasing its ability to destroy cancer cells (Pardoll 2012);

- Cancer vaccines, substances put into the body to start an immune response against certain diseases, for example helping prevent or treat cancer.

Especially drugs that act as checkpoint inhibitors have been recently approved by the Food and Drug Administration (FDA), as *ipilimumab* for receptor CTLA-4 (with important results, see Prieto et al. 2012) and *nivolumab* or *pembrolizumab* for receptor PD-1 (Ansell et al. 2015). These treatments concentrate on one side of the immunity, called *adaptive*, which is meant to provide a durable response relying on the memory capacity: this immune action, once elicited, remembers the intruder and performs a continuous filter.

1.2 T-cell subtypes

The importance of lymphocytes has already been outlined in relation to the immune system. Memory T cells constitute the most abundant lymphocyte population in the body for the majority of a person's lifetime; however, our knowledge of memory T cells derives mostly from mouse studies. Mice have a short lifespan, which cannot sum up the wider exposure to pathogens that occurs inexorably over many decades in humans. Thus, it is important to focus on human studies with human data.

In the human body, after the encounter with a specific antigen, which is a molecule capable of inducing an immune response (production of an *antibody*), the immune response outbreaks, leading the single naïve T cell to the development of multiple subsets of memory T cells. Each subset has different and personal properties and expression profiles. Due to heterogeneity of the memory T cell compartment, a classification is required to identify the various subsets based on phenotypic traits and properties.

According to Golubovskaya and Wu 2016, T cells express TCR (T cell receptor) and can express either CD8 glycoprotein on their surface, therefore being called CD8+ T cells (cytotoxic), or CD4 glycoprotein, being accordingly called CD4 cells (helper T cells, or Th). CD4+ cells differentiate into different subsets: Th1, Th2, Th9, Th17, Th22, Treg (regulatory T cells), and Tfh (follicular helper T cells). These different CD4+ subsets play a critical role in the immune and effector response functions of T cells.

The mechanisms of subdivision into subsets is similar comparing CD4+ and CD8+ T cells, thus our intention is to focus selectively on CD8+ T cells due to their cytotoxicity, that is the quality of inducing cell death, and therefore their relevance to the immunotherapy application. The epitope, the part of the antigen recognized by the immune system, elicits the response of the thymus, which releases mature naïve T cells (T_N). T_N cells are the starters of the immune response after the encounter with an antigen. They are in a quiescent and non-dividing state and are easily recognized

because of the over-expression of the lymph-node homing receptor CD62L and lack of the most common activation markers (HLA-DR, CD69). The T_N, due to antigen encounter, are stimulated to proliferation as effector cells. According to Obar and Lefrançois 2010, 95% of them have a short life, since they are terminally differentiated effector cells and die by apoptosis, being still able of immediate cyto-toxicity ex vivo; the other 5% is constituted by memory precursor effector cells, which are bound to form the memory T-cell population. This last type is recognized by the expression of CD27 (member of the tumor necrosis factor family) and is characterized by a lack in the immediate killing activity.

The different subsets of memory T cells are detected via exploitation of specific markers, like CCR7, a mediator of the homing to the lymph-nodes (LNs). CCR7 expression is related to a lack of production of IL-4, IL-5 effector cytokines and leads to a potential homing to secondary lymphoid tissues, whereas the $CCR7^-$ counterpart has opposite production outcomes and the ability to move to perypheral lymphoid tissues. The cells that express this mediator are defined as central memory (T_{CM}), the ones that don't express it are called effector memory (T_{EM}). A distinguishable phenotypical difference between these two memory types is the fact that T_{CM} have longer telomeres compared to T_{EM}, probably due to a faster replication cycle of the latter one.

Between the T_N and the T_{CM}, the stem cell memory cells (T_{SCM}) have been identified: they have stem-cell like properties, retaining a largely naïve phenotype, but also showing high expression of CD95 antigen, which is characteristic for all memory cells. They preserve genes typical of the naïve with the same functional capacity of memory cells and are also able to self-renew in vitro: the capacity of maintaining the original phenotype after differentiation induction has been demonstrated to be higher with respect to T_{CM} (Gattinoni et al. 2011). Another important property is an enhanced multipotency, as they are able to generate all the memory subsets, whereas it is not possible for them to be generated by other memory subtypes. Additionally, this type is characterized by a superior survival potential.

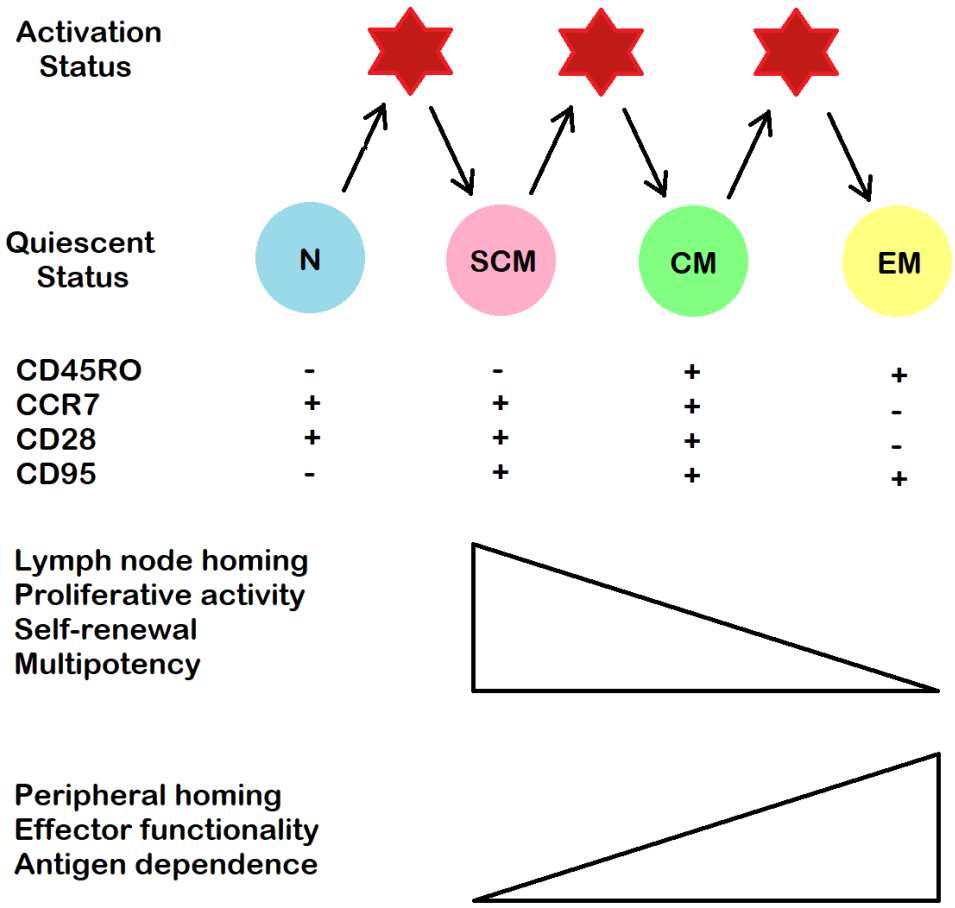

FIGURE 1.1: The main features of T cell subsets are summed up in terms of markers (on/off switch) and likely evolution from naïve (sequence and marker features extracted from Mahnke et al. 2013).

It has been made clear that the quickest way to discriminate between different cell types is offered by the expression of specific markers in different combinations. Given that the expression of a marker is binary (on/off), four markers have been selected as fundamental signals of expression to differentiate the subsets: CD45RO, CCR7, CD28 and CD95 (Mahnke et al. 2013).

Besides the already presented model in which effector cells are the ancestors of central and effector memory, the following (Figure 1.2) has been proposed by Farber, Yudanin, and Restifo 2014. All the memory subsets descend directly from the naïve and

the final stage is the irreversible compartment of T_{Eff} that leads quickly to apoptosis (however it is still unclear, and therefore possibile, whether T_{Eff} and T_{EM} have a reversible link). The markers involved in this alternative scheme have not changed from the ones represented in Figure 1.1. However, the situation is fluid, since numerous paths are being followed, that could potentially amplify our knowledge on the topic: a brief example can be the examination of the differentiation of memory cells after multiple antigenic encounters.

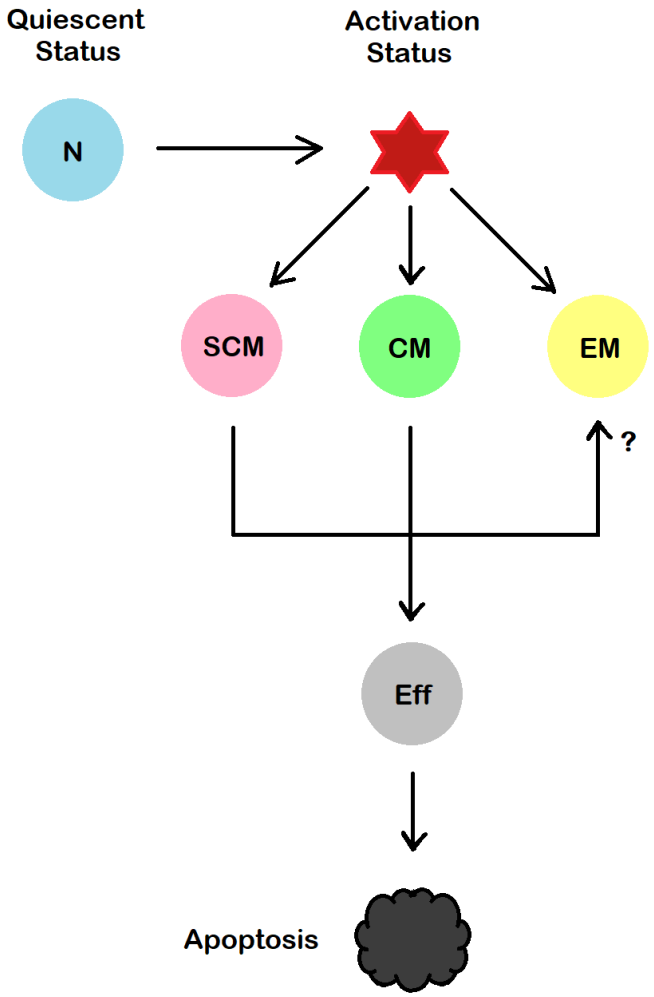

FIGURE 1.2: Graphical illustration of an alternative model for development of T cell subsets from a naïve T cell (Farber, Yudanin, and Restifo 2014).

1.3 Why quantifying the immune contexture

The following subsection is a step towards the understanding of the relationship between tumor and the immune landscape in humans. The main accepted ascertainment is that tumors grow developing an elaborate network of cells, vessels and releasing signaling molecules such as cytokines and chemokines; besides this, it can be infiltrated by immune cells. It is worth concentrating on this last mixture in the cancer contexture, because different combinations of infiltrating immune cells lead to different effects on clinical prognosis.

To give a hint about the variability of these possible combinations, in the tumor all immune cells can be present, like macrophages, mast cells, dendritic cells, natural killer (NK) cells, B cells, naïve, memory lymphocytes and whatever T cell subset. The variability and the variations are triggered by inflammatory factors (Fernández-Figueras et al. 2007) and, as well, different regulatory T cells (Treg) subpopulations seem to explain discrepancies among various types of tumor (Conrad et al. 2012). Moreover, the immune cells can be situated in different locations, either in the core, in the invasive margin or in the adjacent tertiary lymphoid structures (TLS), ectopic lymphoid formations found in inflamed, infected, or tumoral tissues exhibiting all the characteristics of structures in the lymph nodes (LN) associated with the generation of an adaptive immune response (Dieu-Nosjean et al. 2014): a correlation has been found between high densities of TLS and prolonged patient's survival in more than 10 different types of cancer (Sautès-Fridman et al. 2016).

For instance, CD8+ T cells can be visible in both the invasive margin and the core of the tumor, while the TLS seem to lack these cells. In addition, the mixture of immune cells can vary differently in relation to tumor types. Some components of the immune contexture, more than others, are helpful in terms of good prognosis: this fact is shared by multiple papers, such as Dave et al. 2004, which paved the way in the early years of the XXI century, while in Parker et al. 2008 and Parker et al. 2009

the importance of the subtypes' population was outlined. In general, the correlation between infiltration of the cancer contexture by immune cells and the clinical outcome has been researched in various tumor types and a strong lymphocytic infiltration (TILs) has always been stated as beneficial for the patient medical condition. According to Fridman et al. 2012, the infiltration of especially CD3+ T cells, CD8+ T cells and CD45RO+ memory T cells is associated with a longer disease-free survival after therapy. Surprisingly, as stated by Hadrup, Donia, and Straten 2013, as little as approximately 30% of the cells in the tumor are represented by pure cancerous cells: this implies that tumor cells are not able to survive alone, but need a guest environment to grow. This assumption is shared among the scientific community, leading to a massive production of papers on the topic: Afanasiev et al. 2013 was one of the first to propose the therapeutic mobilization of anti-cancer T cells. Indeed, the previously stated variability in the tumor landscape is a considerable obstacle which needs to be taken into account. The approach followed in this dissertation will focus on two complementary paths: on one hand, RNA-Sequencing (RNA-Seq) data processing, on the other, immunohistochemistry (IHC) imaging analysis (developed firstly by Galon et al. 2012), both meant to estimate the numerosity of CD8+ T cell subtypes in the cancer contexture.

1.4 RNA sequencing

RNA-Sequencing (*RNA-Seq*) is a technique useful to study the transcriptome, which is the full range of messenger RNA, or *mRNA*, molecules expressed by an organism. The particular importance of this has to be found in the difference with the genome. In fact, in contrast with it, that is characterized by relative stability, the transcriptome changes, varying due to the different conditions of the organism. Technically, the starting RNA is sequenced in random positions and the subsequently produced reads, which are fractionary copies of the native RNA, are amplified and mapped on a reference genome in the corresponding position (*alignment*): the number of reads abundance (so called *counts*) mapping within a particular gene is proportional to its mRNA (process schematized in Figure 1.3).

Indeed, for large transcriptomes, alignment is more complicated by the fact that a wider section of sequence reads match multiple locations in the genome. One solution is to assign these multi-matched reads by proportionally assigning them based on the number of reads mapped to their neighbouring unique sequences. Alternatively, a paired-end sequencing strategy, in which short sequences are determined from both ends of a DNA fragment, extends the mapped fragment length (Wang, Gerstein, and Snyder 2009).

Despite these obstacles, RNA-Seq has the advantages to become, and is indeed developing, as the gold standard for transcriptome analysis: in fact, this tool

- offers a sensitive and accurate measurement of gene expression;
- is not biased by *a priori* knowledge;
- can be applied to any species;
- is less expensive compared to other techniques (microarray is still more advantageous under these terms, but the costs of RNA-Seq are rapidly decreasing over the years).

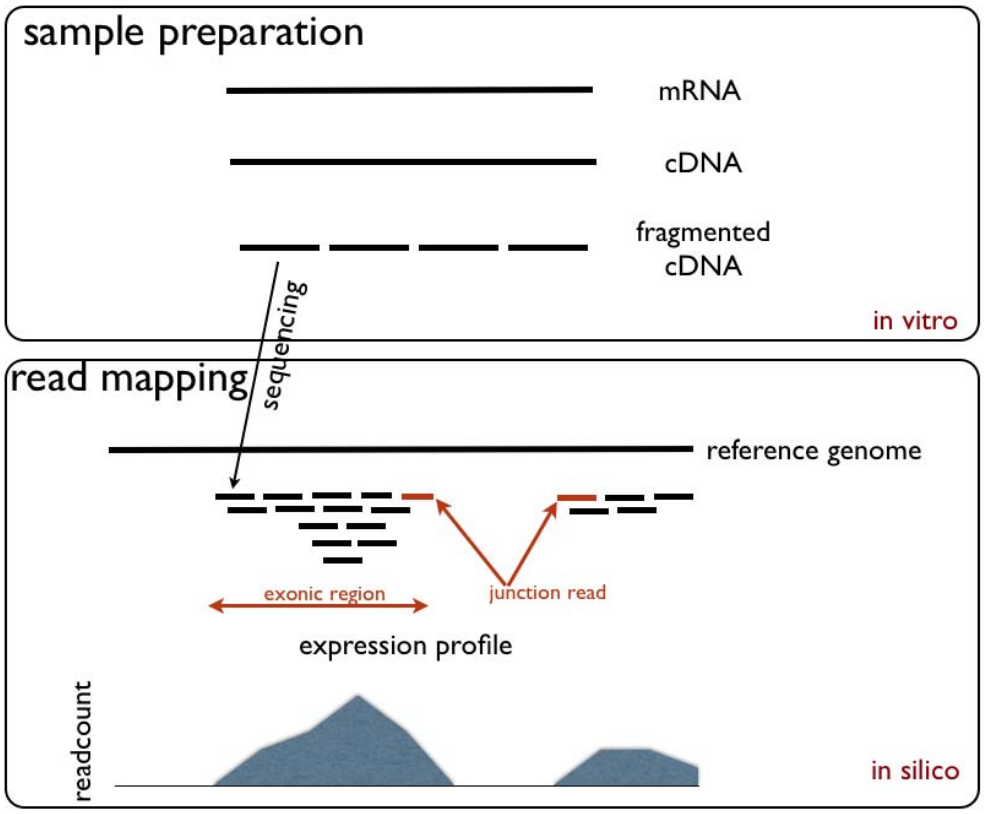

FIGURE 1.3: A typical RNA-Seq experiment. Long native RNAs are first converted into a library of cDNA fragments. Sequencing adaptors are subsequently added to each cDNA fragment and a short sequence is obtained from each cDNA using high-throughput sequencing technology. The resulting sequence reads are aligned with the reference genome or transcriptome. The expression profile is therefore calculated.

Progress has been brought on in areas including the definition of sense and antisense transcripts, alternative splicing events, fused transcripts and transcription initiation sites in physiologically normal and disease settings (Ozsolak and Milos 2010). There exists a wide range of RNA-Seq methods, such as mRNA, targeted RNA Exome Capture or total RNA, leading to a broad variety of possibilities in terms of studies. In fact, the transcriptomes of stem cells and cancer cells are of particular interest to researchers who try to understand the processes of cellular differentiation and carcinogenesis. The readable information that certain genes encode in a precise moment, and under a precise condition, can reveal the status of the cell. Therefore, a sick cell can

unveil pathologic mechanisms of the disorder itself and the determination of genes of importance (*signature* genes, see Chapter 3) sets control or alarm markers for detection and cure of the disease. RNA-Seq comprises gene expression across all cell types in the tumor, including immune cells: tumor data offer the possibility to investigate infiltrating immune cells, as well. Nonetheless, the count output produced by such a method is far from being self-explanatory: data interpretation needs a predetermined pipeline, suitable for the challenges to face in a certain type of analysis. The main usual steps comprise read mapping, counts computation, counts normalization and differential expression (DE) testing (Finotello and Di Camillo 2015). In the experimental section of this thesis (Chapters 3 and 4), some of these issues are going to be faced, with a particular attention given to DE testing, the real core of the entire analysis.

1.5 Immunohistochemistry

Immunohistochemistry (IHC) is an oncology screening technique that allows to analyse cells or tissue specimens collected from patients in early or advanced cancer stage of progression. This technique has the aim to identify the identity of tumoral cells to plan a clinical path of cure. Moreover, IHC is used as well to understand the distribution and localization of biomarkers and differentially expressed proteins, also in non-cancerous patients. Technically speaking, the procedure involves two alternative approaches, direct and indirect. The first one exploits the antibody action targeted to the molecule that is object of research: the antibody itself is binded to a coloured particle that highlights the target molecule when the complex between them is activated. The latter approach is similar, but instead uses a secondary antibody that is targeted to the main antibody and carries the coloured particle. Recently, the secondary antibody has been substituted by conjugate polymers (e.g. dextran). IHC is an alternative to immunofluorescency. Analysing images of the cancer immune contexture can be a powerful instrument both in cancer prevention and prognosis. This has been already stated in the previous subsection (Fridman et al. 2012) in general terms. Immune-based therapies have taken over the past ten years, but many patients fail to respond to these treatments because of the development of a generalized resistance to them. Besides, tumors showing low mutational loads present insufficient reaction rates. Therefore, personalized cancer landscape analysis results vital in terms of understanding the obstacles of immune therapy through identification of in situ or circulating biomarkers.

According to Tsujikawa et al. 2017, response to therapy correlates with degree of infiltrated cell density in tumor contexture and, especially, with percentages of CD8+ T cells expressing the already mentioned markers. To validate this, the team led by Tsujikawa exploited an IHC approach, using either biopsy or surgically resected specimens in formalin-fixed paraffin-embedded tissue sections.

In the experimental section, a collection of all these premises will be translated into a

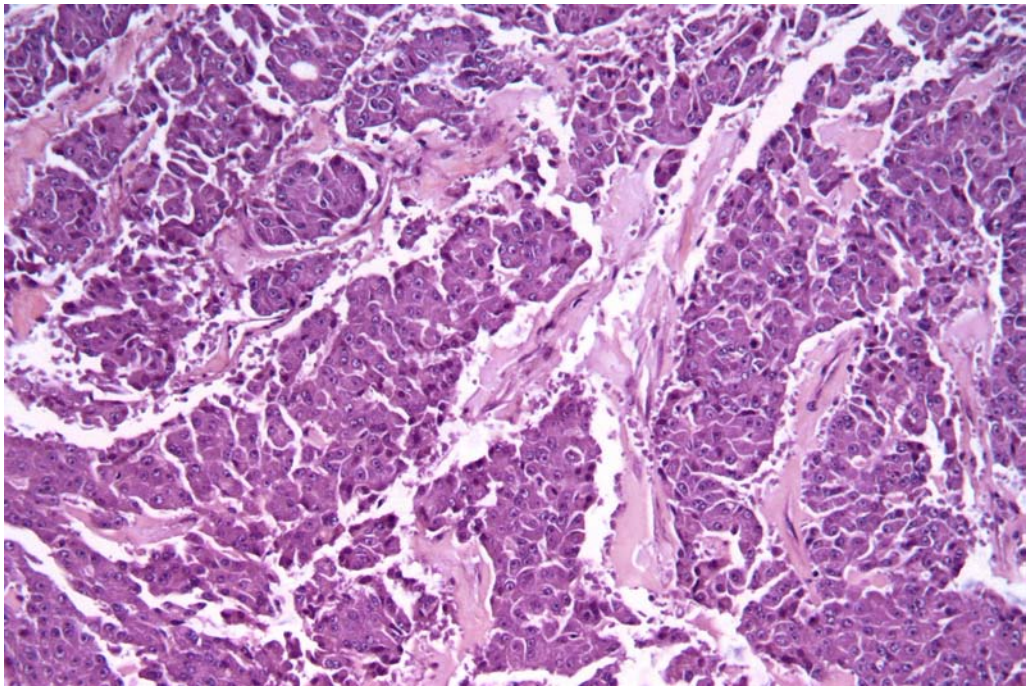

FIGURE 1.4: Example of digital scan representing bright field sequential IHC of one formalin-fixed paraffin embedded (FFPE) section of human cell carcinoma tissue.

pipeline to quantify the tumor contexture population of infiltrating immune cells. An important ally, common to the study mentioned a few lines above, will be Cellprofiler (Carpenter et al. 2006) software.

1.6 Types of tumor samples

Despite being not strictly related to the quantification problem, it is worth recalling the dualism between two main storing methodologies of cancerous tissue specimens, Fresh Frozen (FF) and Formalin-Fixed Paraffin-Embedded (FFPE).

As reviewed by Lüder Ripoli et al. 2016, FF tissues are the best option in terms of molecular analysis by gene expression measurements since their RNA is well preserved. However, these samples are not so abundant and their collection is mainly restricted to tissue banks and research groups. On the other side, FFPE tissues represent a unique source of archived biological material (Von Ahlfen et al. 2007). The advantages of strategies based on the FFPE approach are sundry such as easy handling, long-term cheap storage and suitability for immunohistochemical analyses. However, the RNA of FFPE specimens is of much lower quality than RNA obtained from FF samples (see Scicchitano et al. 2006, study on microarray). According to Sengüven et al. 2014, the chemical reactions induced by formalin are well known: the most important molecular change is the formation of cross-links between proteins or between proteins and nucleic acids, which make the eventual extracted RNA corrupted, or at least not identical to the original one.

The quantification algorithm we are going to introduce will be tested on data coming from both stored tissue examples for two reasons: confirming the robustness of the algorithm itself and, on the sidelines, assessing the feasibility of a FFPE storage instead of the FF one, for the beneficial reasons explained above.

The identification of genomic alterations based on the characterization of more than 500 GBMs has paved the way to understanding several novel mutated genes as well as complex rearrangements of signature receptors: these discoveries lead to new diagnostic and therapeutic target candidates to focus on (Brennan et al. 2013) in the act of analysing the feasibility of FFPE data exploitation, especially after the revelation of an automated protocol for the nucleic acid extraction by Hennig et al. 2010, even if it

was firstly assessed for breast cancer.

Many other studies previously faced the matter, with two-sided results: as in this report, also in Penland et al. 2007, data suggest that meaningful RNA expression analysis can be performed on FFPE samples, with the warnings that many samples are too degraded for analysis and that there is loss of information using FFPE-derived compared to analysis of frozen samples; nonetheless, criteria to predict which blocks will provide informative hybridizations were identified, with a consequent ready-to-action application: the use of microarray expression profiling on FFPE samples to identify tissue of origin in carcinoma of unknown primary (CUP), which represents approximately 3% of all new cancer diagnoses.

1.6.1 Success in RNA Extraction from FFPE Tissue Specimens

Formalin-Fixed Paraffin-Embedded tissue samples have become a standard technique to store and subsequently examine tumour specimens as it is cost effective and perfect to preserve morphology. However, it is also well known that formalin has a strong chemical power and it can modify the composition of the nucleic acids, especially RNA: therefore, the quality and the quantity of RNA extracted from FFPE tissues could be lower than the expectations. One attempt was to replace formalin with other types of fixatives, including Bouin's solution, acetone or alcohol: the differences using these protocols seem not to be consistent. The main steps forward came from the analysis of the factors that could interfere with the amplification of genes from the samples (Gouveia et al. 2014): low amplification rate was hypothesized as related to the presence of contaminants that inhibit the PCR main stages; the addition of a simple washing step after sample rehydration improved the general quality of the data subsequently obtained. Moreover, the introduction of the washing step is likely to increase the pH of the solution, assessing the value between 6.5 and 9.0, ideal for the efficiency of the entire process (Scorsato and Telles 2011).

1.6.2 RNA-Seq of FFPE tumor samples

One of the main issues that need to be investigated regarding FFPE RNA extraction is the action of paraffin, allegedly responsible for fragmentation, cross-linking and chemical modification of FFPE tissue-derived nucleic acids: though, new advances in NGS technologies suggest optimistic perspectives on the investigation of genomes and transcriptomes with limited sample material, situation ideally applicable to usually fragmented nucleic acids extracted from FFPE specimens. An important reference towards this direction is given by the study of Hedegaard et al. 2014, in which the purpose was to determine the effects of formalin fixation, extraction method, storage duration, tissue type, and tumour status on DNA and RNA NGS results using formalin-fixed paraffin-embedded (FFPE) and frozen (FF) specimens. They were able to identify strong correlations between the expression profiles in RNA-Seq data and detect some variants in DNA extracted from exomes for the FFPE/FF couples. Nevertheless, the DNA Exome-Seq library preparation was successful for only ten samples and eight variants out of 61 total specimens (29.5% of success), due to non-efficient amplification of the product during PCR and gene modifications allegedly caused by the chemical fixation: since the main kind of deamination was offered by cytosine to uracil, the idea of using a different uracil tolerant DNA polymerase was also tried, with no significant advantages. In total contrast to the issues encountered during FFPE-derived DNA extraction, the library preparation from RNA-Seq data gave complete satisfaction, offering 100% success rate, including samples collected two decades previously. Moreover, FFPE DNA generated a lower percentage of mapped reads, more unaligned ends, more non-perfectly aligned reads, and more transition errors than FF specimens. For FFPE RNA, still a higher percentage of non-perfect matches and a lower percentage of total exon mapping were observed than with the FF specimens, but the percentage of non-mapped reads was comparable. While the expression data between matched FFPE and FF specimens was strongly correlated, 1494 genes were found to be differentially expressed between FF and FFPE specimens in all analysed

tissue types.

1.6.3 Case study: RNA-Seq from FFPE and FF glioblastoma tumor samples

Year	FF samples			FFPE samples		
	Pre-selection			Pre-selection		
	ng/µl	RIN	SAMPLE CODE	ng/µl	RIN	SAMPLE CODE
2009	7.09	1.1	UNSELECTED	163.47	2.5	UNSELECTED
2008	155.72	4.7		83.63	1.1	
2011	2.79	N/A		282.43	2.4	
2009	372.25	7.1	FF_AA6360	95.31	2.5	FFPE_AA6364
2009	489.17	8	FF_AA6361	321.86	2.4	FFPE_AA6365
2007	145.64	6.8	FF_AA6362	48.69	N/A	FFPE_AA6366
2006	549	7.3	FF_AA6363	1452.14	1.9	FFPE_AA6367
2008	66.3	2.4	UNSELECTED	115.3	1	UNSELECTED
2008	225.13	4.7		322.91	2.5	
2009	211.98	1.9		53.97	2.4	
2010	154.31	3.8		37.55	N/A	

All 11 paired samples were from patients with pathologically confirmed GBM. Gray shading indicates samples that were selected for analyses.

doi:10.1371/journal.pone.0170632.t001

FIGURE 1.5: List of FF and FFPE samples selected based on Esteve-Codina et al. 2017.

By calculating the paired-end distances for each experiment it is possible to evaluate if the length of the RNA molecules is somehow affected by the storage technique: FFPE samples revealed smaller distances between read pairs than FF counterparts. Another important parameter, useful, once defined, to neutralize in vitro RNA degradation effect and improve differential gene expression analysis is the TIN (transcript integrity number, statistics visible in Figure 1.5): the lowest RINs were found in FFPE samples, in correspondence to higher values of GC content, assessing a major RNA degradation for these ones, but, while the values for FFs were homogeneous, the variability was more pronounced in FFPEs. Continuously focusing on variability, another study from Gravendeel et al. (2012) ascertained that, selecting the most variable sets on FFPE expression profiles, the degree of correlation improved dramatically, yielding to an increased success rate (87%) in the assignation of the FFPE samples to cancer subtypes precedently built on FF RNA-Seq data.

FFPE samples offered a higher number of duplicates than the FF counterpart and, consequently, a lower quantity of uniquely mapped reads, especially in the most degraded samples; moreover, these ones presented also the fewer genes to consume

25% of the reads: therefore, basically, the less intact libraries are strongly related to a small number of dominant genes and the libraries themselves present less heterogeneity.

The correlation within each pair was substantially high (around 0.9), with only one exception presented by the most degraded FFPE sample (correlation around 0.3). So, it is easy to presume that FF and FFPE hide high similarities in gene expression. Although a high correlation between FFPE and FF expression profiles has been detected, it has also become clear that FFPE presented, more frequently than the FF data, G to A and C to T transitions, due to the chemical action of paraffin during the fixation process: in extreme cases, if the gene was partially degraded in FFPE data, there would be no chance to recover the single nucleotide polymorphism (SNP).

Chapter 2

Objectives

The division of Bioinformatics of the Medical University of Innsbruck (*MUI*) has recently been seriously active in this direction and this is the place where the reported analysis has been entirely conducted. Numerous studies have been reported on the topic, leading to the introduction of the so-called *immunophenoscore* (Charoentong et al. 2017), capable of identifying and summing up determinants of tumor immunogenicity into a scoring scheme that targets the cure path for a single patient. This thesis is therefore inserted in a wider movement to develop a better understanding on the consequences that the cell population implies in the cancerous tissue origin, growth and maturation. For a catch up, it is worth citing Dander et al. 2014, Angelova et al. 2015 and Hackl et al. 2016, that paved the way to this work.

The main concern of the experimental set-up, described in Chapter 3, is to integrate the potential of RNA-Seq and imaging based quantification methods of the infiltrating lymphocytes in the tumor microenvironment, being associated with prognosis (Fridman et al. 2012). Specifically, the focus on CD8+ T cell subsets is due to their killing role, important for the mechanisms of immunotherapy.

The declared objectives of this thesis are enumerated below:

1. build a deconvolution algorithm from RNA-seq healthy human data, based on SVR, which takes as input a mixture matrix, a signature matrix and a parameter ν, and computes the relative T cell subtypes fractions (CM, EM, N);

2. validate the deconvolution algorithm through healthy human data;
3. apply the deconvolution algorithm to tumor (glioblastoma) human data and discuss the results;
4. test the immunoistochemistry-based pipeline (property of the Division of Bioinformatics, MUI) on IHC CD8+ stained images from cancerous patients and discuss the results.

To quickly overview the experimental set-up that we are switching to in the next Chapter, the first section will focus on the quantification method as well as the core algorithm *CD8quant*. The deconvolutional approach selected for the RNA-Seq side of the study is going to be formulated based on human healthy donors data and subsequently furtherly tested on cancerous tissues in order to prove, on one hand, the robustness of the algorithm and, on the other hand, also its applicability to FFPE stored tissues, that, as it has been previously acknowledged, are more durable in terms of conservation, but potentially less reliable in terms of nature of the extracted data.

In conclusion, the imaging section will show a quantification pipeline, *runCP.py*, intellectual property of the Medical University of Innsbruck (Division of Bioinformatics), from IHC high resolution tissue tiles, applied to data consistent with the previous analysis in order to calculate the number of total CD8+ T cells. Unfortunately, due to legal issues of divulgation, it has been not possible to give a complete view on the entire topic, data and results, but what we presented is sufficient to be introduced to this novel approach, complementary to the deconvolutional one.

Chapter 3

Materials and Methods

3.1 Development of an algorithm for the quantification of CD8+ T cell sub-populations from tumor RNA-Seq data

3.1.1 Selection and preprocessing of RNA-seq data of CD8+ T cells subpopulations

Since the aim of the project is to develop a computational method for the quantification of the relative fractions of CD8+ T-cells subsets, the first step was to select suitable publicly available RNA-seq data sets from enriched/purified Naïve CD8+ T cells (N), effector memory CD8+ T cells (EM), and central memory CD8+ T cells (CM) to reconstruct their expression signatures. We selected two studies, from Pulko et al. 2016 (data available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80306) and Bonnal et al. [referred to De Simone et al. 2016] (data available at ftp://ftp.sra.ebi.ac.uk/vol1/fastq/ERR431/). Table 3.1 lists the considered samples and summarizes the data characteristics. From now on, the data will be referred to as Pulko and Bonnal data sets.

Pulko data, available in SRA format, were converted into FASTQ format, with the fastq-dump function of the SRA toolkit (https://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?view=software).

TABLE 3.1: Characteristics and information concerning the data of Pulko and Bonnal datasets: GEO ID (https://www.ncbi.nlm.nih.gov/geo/), cell type, markers used to isolate the cells, Pubmed ID.

GEO ID	Cell Type	Marker	ID Pubmed	study	SE/PE
ERR431612	N	CCR7+ CD45RA+ CD45RO-	26451251	Bonnal et al.	PE
ERR431605	N	CCR7+ CD45RA+ CD45RO-	26451251	Bonnal et al.	PE
ERR431623	N	CCR7+ CD45RA+ CD45RO-	26451251	Bonnal et al.	PE
ERR431627	N	CCR7+ CD45RA+ CD45RO-	26451251	Bonnal et al.	PE
ERR431592	N	CCR7+ CD45RA+ CD45RO-	26451251	Bonnal et al.	PE
ERR431607	CM	CCR7+ CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431578	CM	CCR7+ CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431593	CM	CCR7+ CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431602	CM	CCR7+ CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431621	CM	CCR7- CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431585	EM	CCR7- CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431573	EM	CCR7- CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431590	EM	CCR7- CD45RA- CD45RO+	26451251	Bonnal et al.	PE
ERR431567	EM	CCR7- CD45RA- CD45RO+	26451251	Bonnal et al.	PE
GSM2124048	EM	CD45RA- CCR7- CD95+ CD28-	27270402	Pulko et al.	SE
GSM2124049	EM	CD45RA- CCR7- CD95+ CD28-	27270402	Pulko et al.	SE
GSM2124050	EM	CD45RA- CCR7- CD95+ CD28-	27270402	Pulko et al.	SE
GSM2124051	EM	CD45RA- CCR7- CD95+ CD28-	27270402	Pulko et al.	SE
GSM2124052	EM	CD45RA- CCR7- CD95+ CD28-	27270402	Pulko et al.	SE
GSM2124053	N	CD45RA+ CCR7+ CD95- CD28int	27270402	Pulko et al.	SE
GSM2124054	N	CD45RA+ CCR7+ CD95- CD28int	27270402	Pulko et al.	SE
GSM2124055	N	CD45RA+ CCR7+ CD95- CD28int	27270402	Pulko et al.	SE
GSM2124056	N	CD45RA+ CCR7+ CD95- CD28int	27270402	Pulko et al.	SE
GSM2124060	CM	CD45RA- CCR7+ CD95+ CD28+	27270402	Pulko et al.	SE
GSM2124061	CM	CD45RA- CCR7+ CD95+ CD28+	27270402	Pulko et al.	SE
GSM2124062	CM	CD45RA- CCR7+ CD95+ CD28+	27270402	Pulko et al.	SE
GSM2124063	CM	CD45RA- CCR7+ CD95+ CD28+	27270402	Pulko et al.	SE
GSM2124064	CM	CD45RA- CCR7+ CD95+ CD28+	27270402	Pulko et al.	SE

3.1.2 Construction of the signature matrices for CD8+ T cell subpopulations

We used the Bonnal data to derive the signature matrices for the CD8+ T cells subpopulations of interest to be used in our deconvolution algorithm. Gene expression was quantified from the FASTQ files using TIminer (Tappeiner et al. 2017). In particular, the matrices of raw counts and transcripts per million (TPM) were constructed using the *executeKallistoDir* function of TIminer. After quality control performed on the expression data, two samples of the Bonnal data set, ERR431585 and ERR431621, were excluded from the subsequent analyses (see Results). To identify the signature genes that are specific for the single CD8+ T cell sub-populations, we analysed the RNA-seq counts with two R packages for differential expression analysis which assume a negative binomial (NB) distribution of the data: edgeR (Robinson, McCarthy, and Smyth 2010) and DESeq2 (Love, Huber, and Anders 2014). edgeR and DESeq2 provided lists of differentially expressed genes with high overlap (data not shown), but the edgeR analysis required a much lower computational time. Therefore, we selected edgeR for the final selection of the signature genes. To identify the signature genes that are specific for the CD8+ T cell sub-populations of interest and do not present high expression in other cell types, additional expression data were included for the following cell types: B cells, dendritic cells, classically (M1) and alternatively (M2) activated macrophages, monocytes, neutrophils, natural killer cells, regulatory T cells, CD4+ T cells, and tumor cells. edgeR was applied to compare the libraries from the cell types of interest with the others, using two different approaches. In the first approach (*one against all*), the cell type of interest was compared with all the other libraries considered as a single group. In the second approach (*one against each*), the cell type of interest was compared with each of the other cell types singularly. Before running edgeR, the counts were pre-processed to filter out low expression genes: we kept only the genes that presented a CPM-value (counts per million) higher than 2 in at least three libraries. Given that the data were generated from purified/enriched cell types,

edgeR was run imputing a constant value as dispersion of 0.01 (parameter required by the package implementation). For each gene assessed in the comparisons, edgeR computed the p-values and the log-fold-changes. In the *one against each* comparison, the minimum log-fold-change and the maximal p-value for each gene were considered. We selected as signature genes the genes that had a log-fold-change higher than *logfc* and a p-value lower than *pval* (values presented below). Furthermore, to improve the specificity of the selected signature genes, we considered an additional heuristic: the average transcript per million (TPM) expression across all libraries in the cell type of interest $X_{i \mid i \in c}$ must be K times higher than the maximum TPM expression selected among all other libraries $X_{j \mid j \notin c}$:

$$X_{i \mid i \in c} \geq K \cdot X_{j \mid j \notin c}$$

with c being the cell type of interest. We tested different logfc and pval cut-offs and different values for the K heuristic:

- pval = 0.0005, 0.001, 0.005, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25;
- logfc = 0, 0.5, 1;
- K = 0.80, 0.85, 0.90, 0.95, 1.00, 1.05, 1.10, 1.15, 1.20, 1.25, 1.30.

Finally, for each set of signature genes, the corresponding signature matrix was built taking the mean gene expression in TPM across all libraries belonging to each cell type N, EM, or CM. In total, employing the one against all and the one against each approach, 594 signature matrices (297 for each configuration, see the following Table A.1) were built from the Bonnal RNA-seq data.

3.1.3 Simulation of RNA-seq data from mixtures of CD8+ T cells subpopulations

Pulko and Bonnal FASTQ data were used to generate 200 RNA-seq data sets from simulated mixtures of EM, CM, and N CD8+ subpopulations. The 100 data sets generated from Pulko data consisted in 1 million single-end reads, whereas the 100 data sets generated from Bonnal data consisted in 1 million read pairs. The CM, EM and N cell proportions were simulated sampling, for each cell type, a uniform distribution in [0,1] and then normalizing the fractions so to sum up to 1 (i.e. 100% in each simulated mixture) (Table B.1).

To generate the read mixtures, $F_c \cdot R$ reads (or read pairs), were extracted from a single FASTQ file of the Pulko or Bonnal study generated from the cell type *c*, where F_c is the cell fraction for cell type *c* and R=1e6 is the number or total reads (or read pairs) in the mixture. Reads from CM, EM and N were extracted from the original FASTQ files with the Seqtk tool (Li 2012) and admixed in the final single-end or paired-end FASTQ files. From each mixture FASTQ files, gene expression was quantified using the *executeKallistoDir* function of TIminer to compute the matrices of raw counts and TPM, analogously to what we have done for the signature matrices construction.

3.1.4 Implementation of the deconvolution algorithm

We implemented *CD8quant*, a deconvolution algorithm based on nu Support Vector Regression (nu-SVR) similar to that of CIBERSORT (Newman et al. 2015) but intended for the quantification of the relative cell proportions of three different CD8+ T cell subpopulations. CD8quant takes as input a *signature matrix* describing the signatures of the three CD8+ T cell types of interest, i.e. CM, EM, and N, and a *mixture matrix* describing the expression of different cell mixtures and estimate through deconvolution the relative fractions of the CD8+ sub-types. Both matrices must represent gene expression in TPM. The deconvolution with nu-SVR can be run with different

values of the nu parameter, which controls both an upper bound on training errors and a lower bound on support vectors. Higher values of ν yield narrower ε-tubes and consequently, more support vectors. Before performing deconvolution with nu-SVR, CD8quant normalizes the signature matrix and the mixture matrix using the following formula

$$Matrix_{\text{norm}} = \frac{Matrix - mean(Matrix)}{SD(Matrix)}$$

and are then filtered to keep only the genes in common. After deconvolution, possible negative cell fractions estimated with nu-SVR are casted to zero and normalized so that the final cell fractions to sum up to 1.

3.1.5 Identification of the optimal signature matrix and assessment of deconvolution performance

CD8quant was applied to all the mixture matrices from Pulko data using each time a different signature matrix (Table A.1) and using ν=0.5. Both the configurations, one against each and one against all, were tested. The cell fractions estimated by CD8quant where then compared with the true fractions (Table B.1) using Pearson's correlation and Root-mean-square error (RMSE).

$$corr_{\text{Pearson}} = \frac{cov(x,y)}{SD(x) \cdot SD(y)}$$

$$RMSE = \frac{\sum_{i=1}^{n} (x_{\text{i}} - \hat{x}_{\text{i}})^2}{n}$$

The values of RMSE and correlation were calculated considering both one single cell type of interest at a time and all cell types together (*overall* statistics). Finally, the signature matrix with the lowest overall RMSE was selected as basis signature for the CD8quant algorithm (called *CD8sig* from now on). The CD8sig signature was the used

to deconvolute the Pulko and Bonnal simulated mixtures varying the nu-parameter value, selected as 0.25, 0.5 and 0.75.

3.1.6 Application of CD8quant to RNA-seq data from fresh frozen and formalin-fixed paraffin-embedded data from human bulk tumors

In order to verify the applicability of CD8quant to cell mixtures from from fresh frozen (FF) and formalin-fixed paraffin-embedded (FFPE) samples of human tumors, we considered paired FF and FFPE samples from four glioblastoma tumors profiled with RNA-seq available in the literature (Table 3.2). The data, that from now on will be referred to as the Codina dataset (Esteve-Codina et al. 2017), were downloaded at the link https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP089805.

Run ID	Sample Name	Age	Sex	Tissue	SE/PE	Tumor Type
SRR4241104	Case 1_FF_AA6360	55	female	FF brain	PE	Glioblastoma
SRR4241105	Case 1_FFPE_AA6364	55	female	FFPE brain	PE	Glioblastoma
SRR4241106	Case 2_FF_AA6361	51	female	FF brain	PE	Glioblastoma
SRR4241107	Case 2_FFPE_AA6365	51	female	FFPE brain	PE	Glioblastoma
SRR4241108	Case 3_FF_AA6362	79	male	FF brain	PE	Glioblastoma
SRR4241109	Case 3_FFPE_AA6366	79	male	FFPE brain	PE	Glioblastoma
SRR4241110	Case 4_FF_AA6363	53	female	FF brain	PE	Glioblastoma
SRR4241111	Case 4_FFPE_AA6367	53	female	FFPE brain	PE	Glioblastoma

TABLE 3.2: Characteristics of Esteve-Codina glioblastoma RNA-seq data: library ID in the Sequence Read Archive, sample name, age and sex of the patient, tissue type (FF=Fresh Frozen, FFPE=Formalin-Fixed Paraffin-Embedded), and tumor type.

The Esteve-Codina FASTQ files were analysed with the executeKallistoDir function of TIminer to extract gene counts and TPM, as explained above. The mixture matrix of the gene TPM was then analysed with CD8quant, using the CD8sig signature and

three different v values (0.25, 0.5, 0.75), to estimate the fractions of EM, CM, and N CD8+ T cells subtypes. The cell fractions from paired FF-FFPE (i.e. samples generated from the same tumor) (Table 3.2) were compared considering RMSE and Pearson's correlation.

3.2 Quantification of CD8+ T cells from tumor images

This experimental section comprises the cell quantification, with a focus on nuclei of CD8 IHC (immunohistochemistry) stained images. The provenience and nature of patients and their cancer images is intellectual property of the team involved in the experiments, therefore this type of information cannot be divulgated. The pipeline, meant to go through all the analysis steps, has been developed by Clemens Mayer, under the supervision of Zlatko Trajanoski, head of the Bioinformatic department of the Medical University of Innsbruck. This workflow utilises several public available tools for the different steps of pre-processing (bftools, Linkert et al. 2010), classification (Ilastik, Sommer et al. 2011) and segmentation (CellProfiler, Kamentsky et al. 2011) and combined the individual steps in a python script (*runCP.py*, available at the link `https://github.com/mui-icbi/IHCount`), which is easy to use and to adapt.

3.2.1 Preprocessing

High resolution images were preprocessed by Mayer himself, who provided different sets related to various patients. Each set was organized in a folder, containing the tiles (2000x2000 pixels) the images have been cropped into. The tiles were converted from $*$.tiff to $*$.png.

3.2.2 Classification

The Pixel Classificator module of Ilastik was used to establish classifiers from a subset of the previously generated image tiles. I manually selected a few images, which represent the areas of interest on the tissue slide. Using manual annotation, that is exploiting the brush-like tool on Ilastik, the classifiers were trained to distinguish CD8+ cells (blue and brown round-shaped cells) and all nuclei (nuclei, only the darker cells) on the selected IHC images, as well as tissue and background. The color protocol, for convenience in the following quantification step, was decided in the beginning of the

analysis and not changed throughout the whole pipeline: red (either positives or nuclei), blue (background) and green (tissue). As a result of running the classifier as a batch process on all tiles of an image, I obtained two sets of so called probability maps. One set shows the probabilities for positives and the second for nuclei on the slide. These two image sets were used for the following segmentation and quantification step.

3.2.3 Segmentation and Cell Quantification

The final phase of the analysis has been actuated in collaboration with Mayer: segmentation and counting were performed with CellProfiler. This program presented the possibility to reduce noise and to isolate the three different classes by multiple intensity based operations, as good as possible: this optimization process was performed by Mayer as preliminary step to the segmentation. The probability maps, once obtained from Ilastik, were loaded into CellProfiler, exploiting one of the commands present in the pipeline runCP.py in the python environment. This step could also be manually performed through the CellProfiler's GUI. The Bioinformatics team, acknowledged above, created a pipeline (IHCount.cppipe4) that uses several internal modules to identify and count positive stained cells, nuclei and the area of the tissue. As a first step, the probability maps were split by the different colour channels (red, blue, green) and converted into grayscale. In the end, the final output involved tables for each image that list the questioned parameters for each single tile. The combination of the results of each single tile offered a uniform macro-table with the information extracted from all the patients.

Chapter 4

Results and Discussion

4.1 Validation of the Deconvolution Algorithm

4.1.1 Filtering of Bonnal's data

All Bonnal RNA-seq libraries listed in Table 1 were used to generate the signature matrices, with the exception of ERR431585 and ERR431621. The ERR431585 library was discarded due to a very low expression of the CD8A and CD8B maker genes (Figure 4.1), whereas the ERR431621 library was discarded for its dissimilarity from the other EM libraries revealed by principal component analysis (PCA) (Figure 4.2).

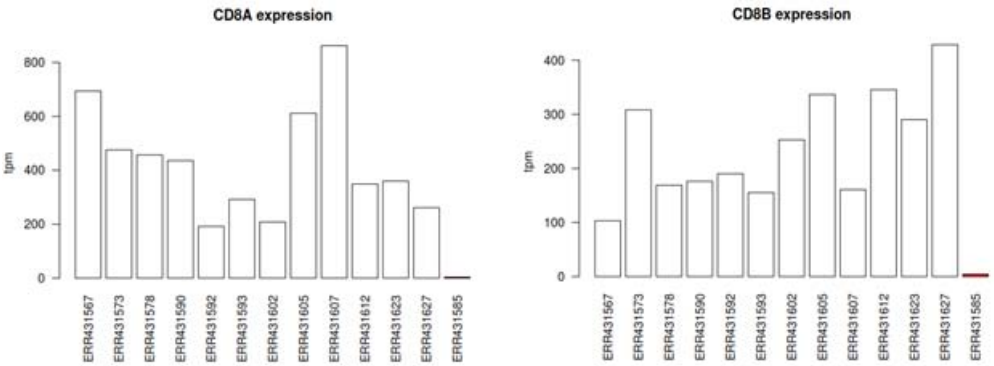

FIGURE 4.1: Barplots showing the low gene expression of CD8A and CD8B in the library ERR431585, therefore excluded from the conducted analysis, because of its irrelevance to the topic.

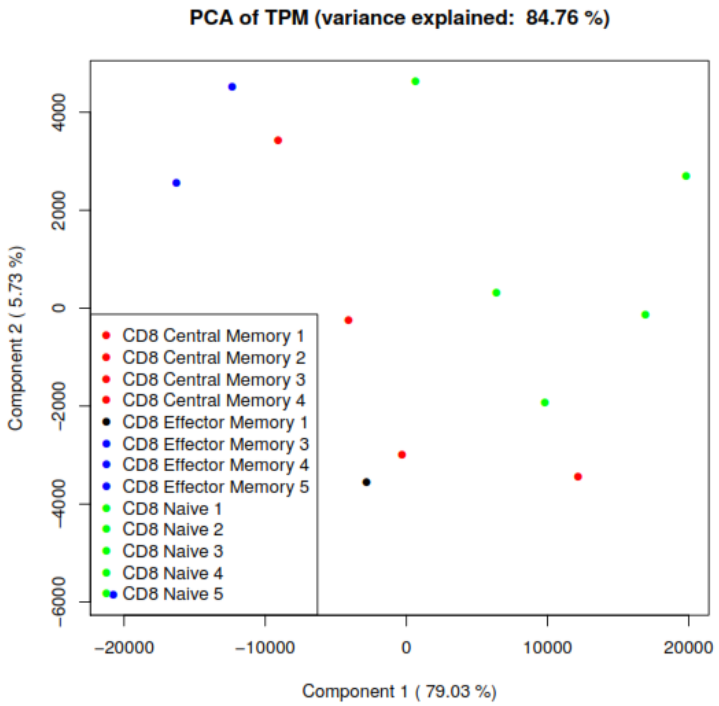

FIGURE 4.2: PCA of the Bonnal dataset (ERR431585 already excluded). ERR431621 is represented by the black dot and it is clear the non-belonging to the EM cluster.

4.1.2 Analysis of simulated data derived from the Pulko's study

100 simulated mixtures were generated from all the Pulko libraries (Table 3.1) with exception of the two libraries described in the previous section, and considering the relative cell fractions reported in Table B.1.

The deconvolution results obtained with CD8quant on these data are presented for both the configurations *one against each* and *one against all*. Table B.2 describes the deconvolution performance obtained by CD8quant (v=0.5) using the diffent the Pulko mixtures.

We selected the signature matrix with the lowest overall RMSE, i.e. *Sig_1vs1_103*, as basis signature for the CD8quant algorithm (*CD8sig*). The CD8sig signature was then used in the deconvolution of the Pulko simulated mixtures with CD8quant, using three values of the nu parameter: 0.25, 0.5, 0.75. The results are presented in Table 4.1 and Figure 4.3.

	RMSE all	RMSE CM	RMSE EM	RMSE N	corr all	corr CM	corr EM	corr N
nu = 0.25	0.18	0.19	0.2	0.14	0.3	0.36	0.45	0.63
nu = 0.5	0.16	0.18	0.15	0.13	0.47	0.49	0.53	0.69
nu = 0.75	0.15	0.17	0.15	0.14	0.5	0.49	0.53	0.69

TABLE 4.1: Deconvolution performances obtained by CD8quant with the CD8sig signature using different nu parameters. The columns show the performance parameters expressed as RMSE and Pearson's correlation.

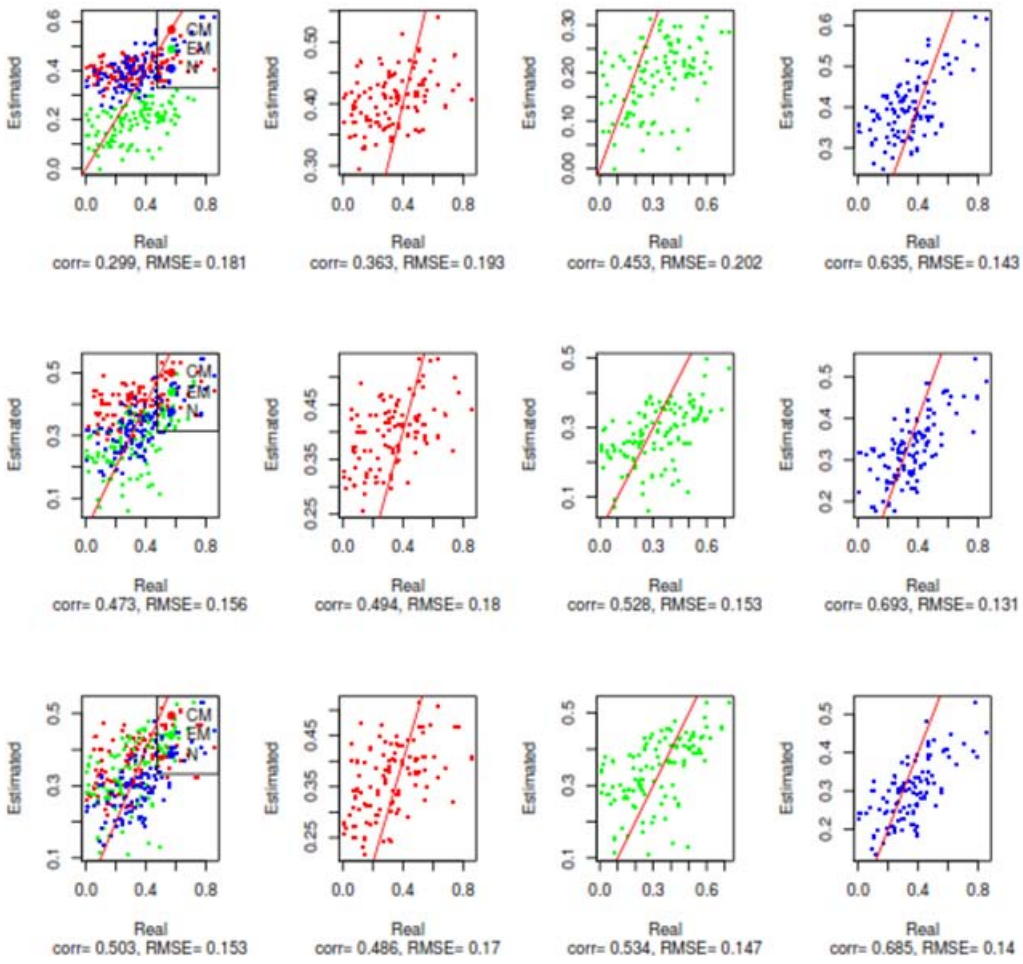

FIGURE 4.3: Deconvolution performances obtained by CD8quant with the CD8sig signature using different ν parameters. The X-axis presents the real concentrations for each simulation, while the Y-axis represents the estimated values after the application of the algorithm. The red line lays along the 45° diagonal.

4.1.3 Analysis of simulated data derived from the Bonnal's study

100 simulated mixtures were generated from the Bonnal libraries (Table 3.1) considering the relative cell fractions reported in Table B.1. To furtherly check the quality of the selected signature, we deconvoluted the Bonnal simulated mixtures using the CD8sig signature. On these data, a higher performance is obtained probably due to a higher similarity of the Bonnal signature and mixture profiles.

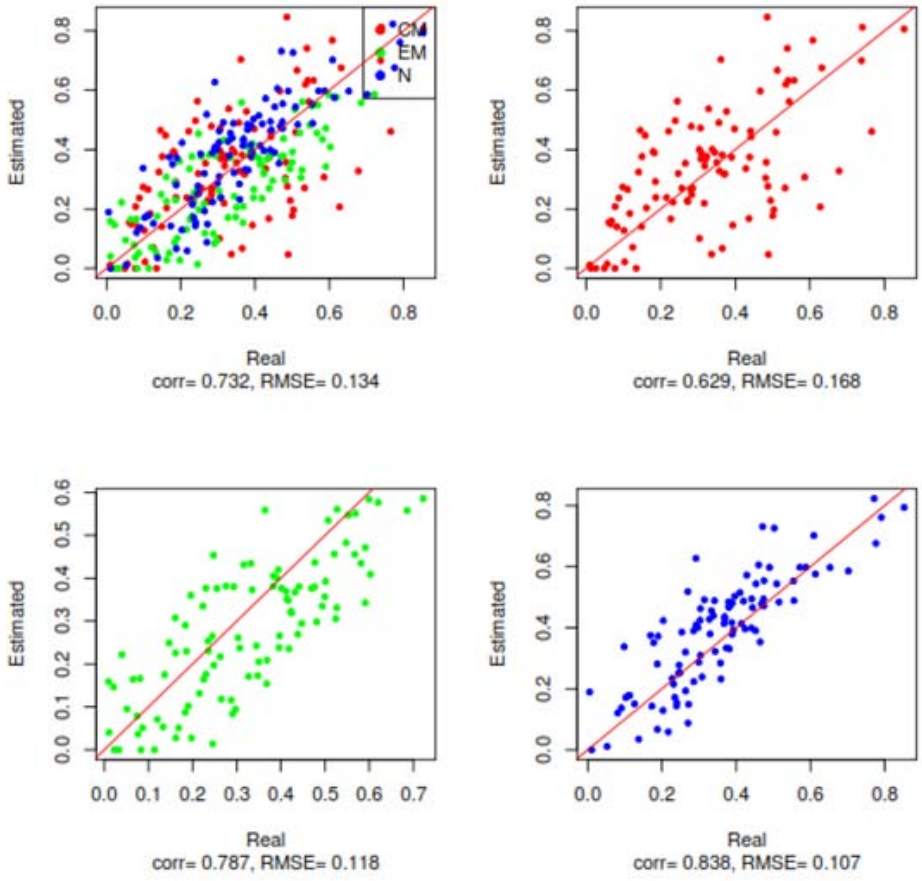

FIGURE 4.4: CD8sig and the preferred ν set-up applied to Bonnal mixture through CD8quant algorithm.

4.1.4 Deconvolution of the tumor RNA-seq data from FF and FFPE samples

To test the applicability of CD8quant to both FF and FFPE data, we considered the Esteve-Codina data set (see Chapter 3, Table 3.2) which consists in RNA-seq expression data form four pairs of FF-FFPE samples each taken from the same tumor. As it is evident in Figure 4.5 and Table 4.2, gene expression has high correlation for all pairs except for the Pair 2, due to the low quality of the FFPE sample for this pair (Esteve-Codina et al. 2017).

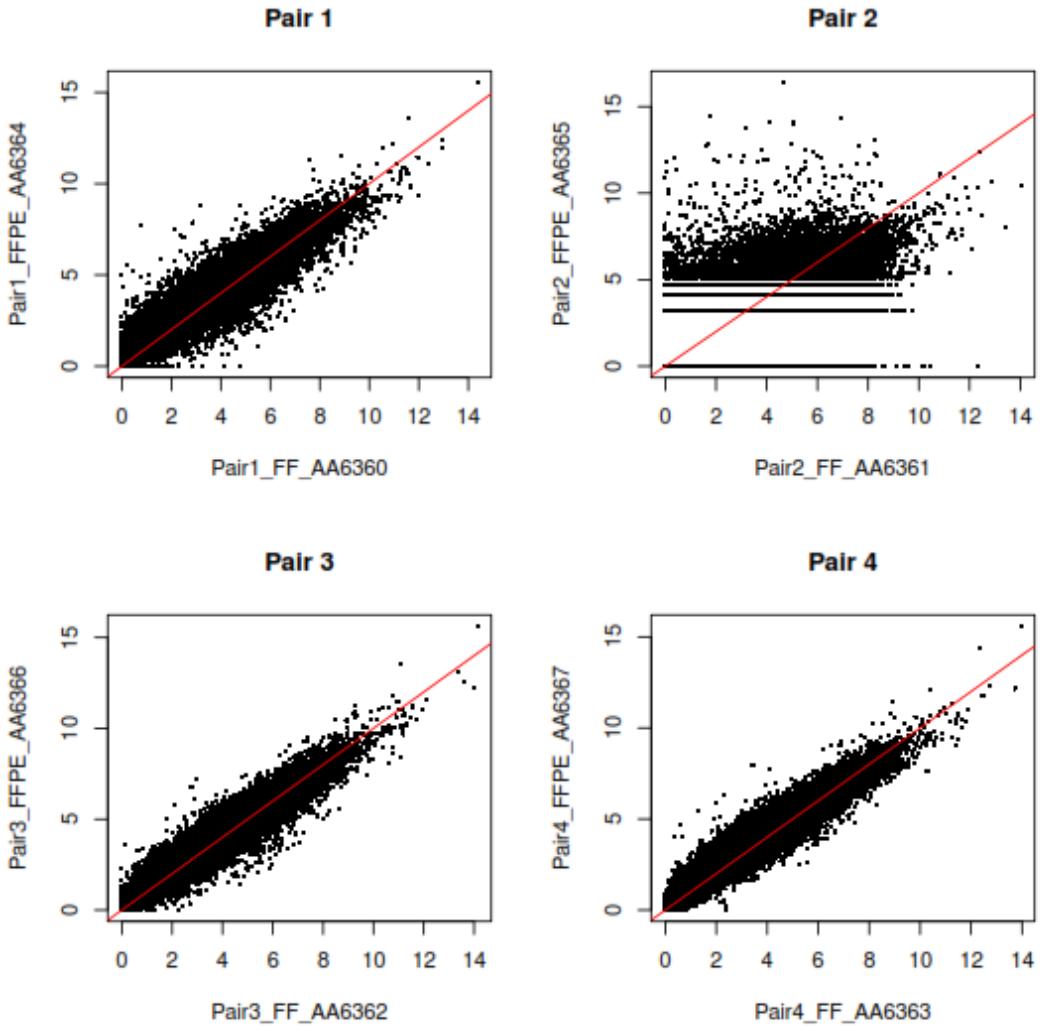

FIGURE 4.5: Correlation plots of log2(CPM+1) Codina data. Each plot is related to one single pair and shows the assessment of the data along the 45 degrees line (red).

	correlation	RMSE	ID_FF	ID_FFPE
Pair1	0.9356	0.8886	Pair1_FF_AA6360	Pair1_FFPE_AA6364
Pair2	0.6185	24.926	Pair2_FF_AA6361	Pair2_FFPE_AA6365
Pair3	0.9603	0.6797	Pair3_FF_AA6362	Pair3_FFPE_AA6366
Pair4	0.968	0.6315	Pair4_FF_AA6363	Pair4_FFPE_AA6367

TABLE 4.2: Correlation and RMSE within each pair, obtained comparing FF and FFPE samples' gene expression from the same tumor.

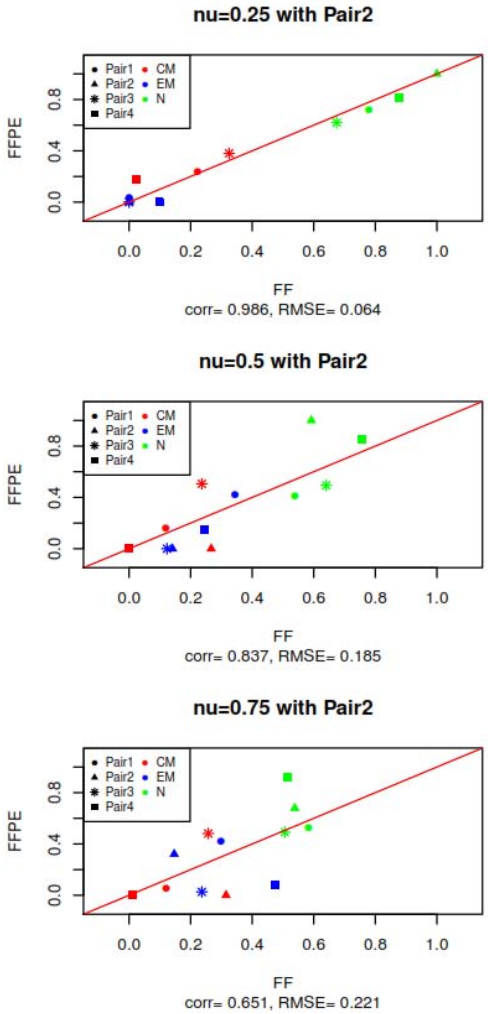

FIGURE 4.6: Correlation plots of the bulk tumor deconvolution output with Codina mixture matrix.

We analysed the Esteve-Codina dataset with CD8quant using the CD8sig and three

values for ν parameter, that is 0.25, 0.5, 0.75. The best results are obtained with ν=0.25. The high correlation between the FF and FFPE results confirms the robustness of CD8quant in the analysis of possibly degraded FFPE samples, as in the case of Pair2.

4.2 Quantification of CD8+ cells from IHC tumor images

The images for each patient were digitally cropped into squared subsets of 2000x2000 pixels of dimension. Figure 4.7 shows an example of full image (left) after the application of the cropping net, while in the same Figure on the right, a subset image is presented.

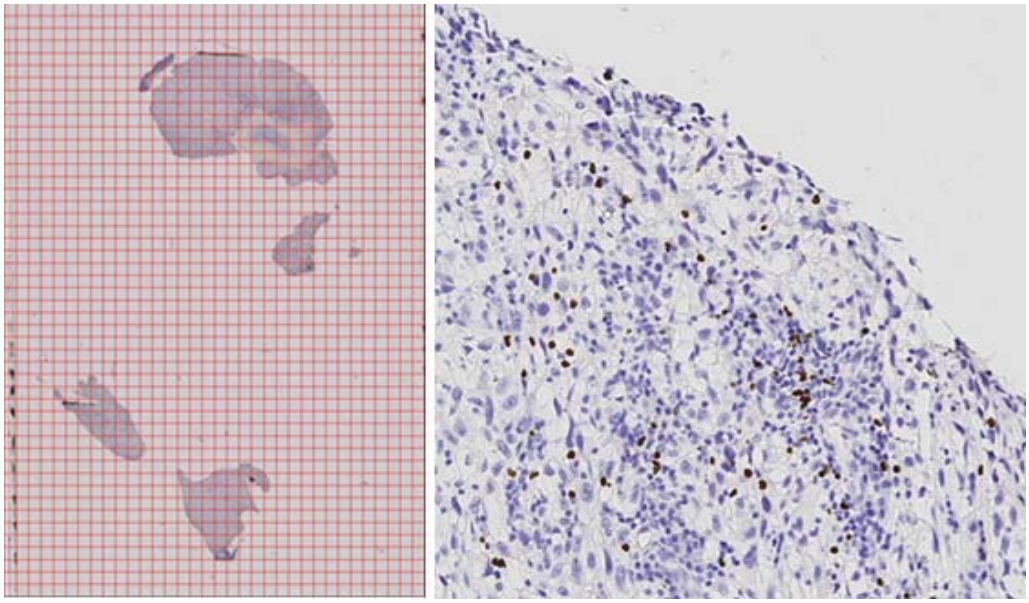

FIGURE 4.7: On the left, the full image, on the right an already cropped image extracted from the left one.

The following training of the pixel classifier was executed manually, exploiting a colour brush-like procedure provided by Ilastik itself. The process of training was 100% manual, sharing the strength and the weaknesses of the human eye: in Figure 4.8 it is possible to observe how a partially brushed tile looks like after a few steps into the training.

Figure 4.9, instead, presents a probability map, produced by the pixel classifier module, after the manual training. As mentioned above, the ability of human eye is not capable of achieving a perfect result, therefore we exploited a CellProfiler's feature to reduce the noise: especially in the background (blue), the green presence was reduced and the round-like shape of nuclei and positives (red) was enhanced (Figure

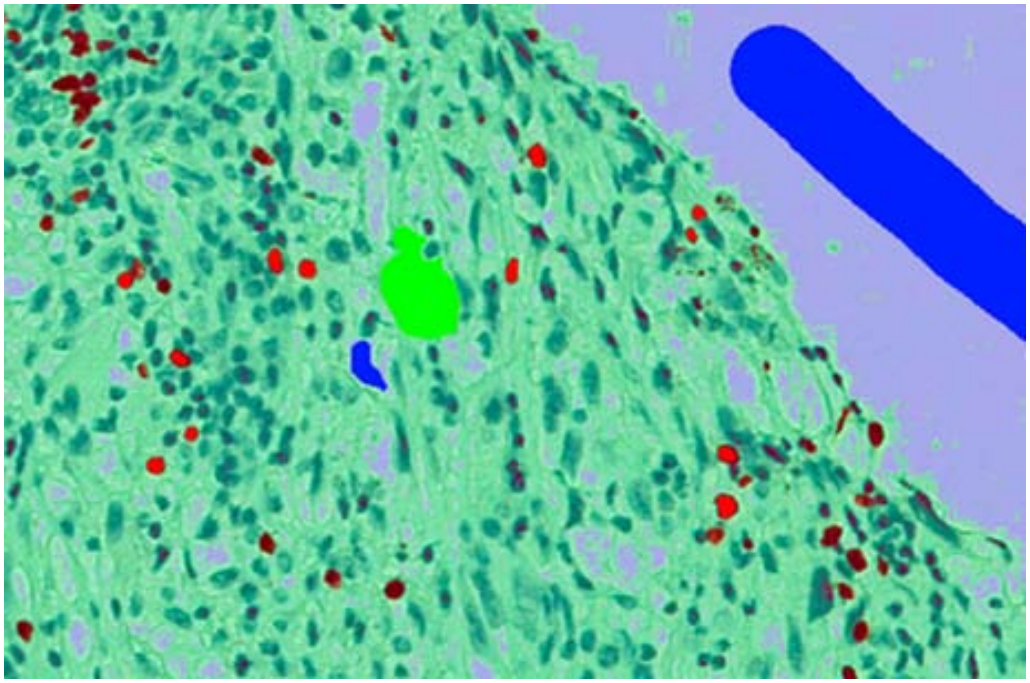

FIGURE 4.8: Example of training, positives session: the vivid colours (blue, red and green) have been traced by the paint brush manually, while the soft-faded coloured foil over the tile is due to the activation of the "live checking" feature.

4.9, right) to enable a better segmentation and object detection.

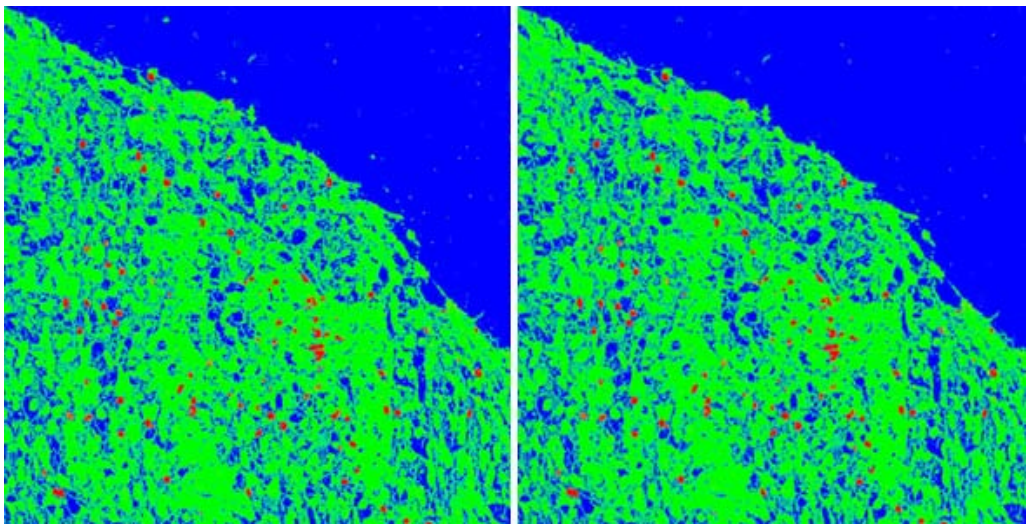

FIGURE 4.9: Classified tile before optimization *VS* Classified tile after optimization.

CellProfiler produced an imaging output that involves segmentation results for nuclei and positives, as well as images of the tissue area (Figure 4.10, background not

shown). The subsequent output was instead a table (Table 4.3) resuming the IDs of patients (obscured for privacy reasons) and the cell count for the images to each of them related: only the CD8 results are shown in this section for consistency to main topic.

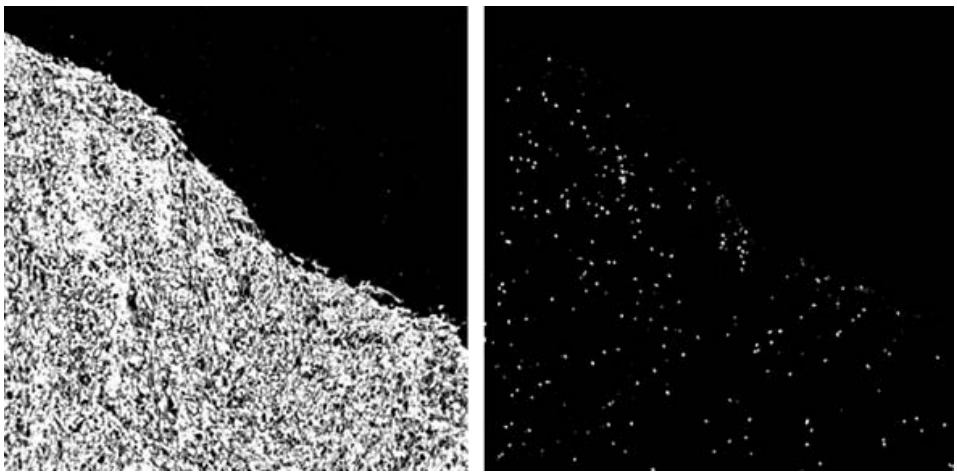

FIGURE 4.10: Segmentation output of CellProfiler regarding tissue and cells (nuclei in this case).

Image	Pixels of occupied Area (1px = 0.5 µm)	Sum of nuclei (including positives)	Sum of positives	Fraction Nuclei/Positives %
Patient 01	1.94E+08	191155	17823	9.323847
Patient 02	4.47E+08	1085883	200424	18.45724
Patient 03	1.9E+08	96719	6901	7.135103
Patient 04	51387976	62796	7874	12.53902
Patient 05	7.98E+08	1094620	75971	6.940399
Patient 06	5.01E+08	1331952	75937	5.701181
Patient 07	5.83E+08	1316218	133625	10.15219
Patient 08	1.9E+09	1499273	85280	5.68809
Patient 09	1.42E+09	1462025	81232	5.556129
Patient 10	1.65E+09	822383	126197	15.34528
Patient 11	8.68E+08	1445921	303031	20.95765

Patient 12	1.49E+09	1560646	88535	5.672971
Patient 13	1.23E+09	2264420	43349	1.914353
Patient 14	4.27E+08	878257	43698	4.975537
Patient 15	1.82E+09	1322095	69681	5.270499
Patient 16	60110177	88409	7217	8.163196
Patient 17	9.76E+08	1472746	114593	7.780907
Patient 18	1.32E+09	1425728	95948	6.729755
Patient 19	6.7E+08	919512	133490	14.51748
Patient 20	1.26E+08	142645	6842	4.796523
Patient 21	6.01E+08	1286295	55883	4.344493
Patient 22	1.23E+08	225973	5241	2.319304
Patient 23	1.52E+09	2382471	286175	12.01169
Patient 24	1.45E+09	1373325	54197	3.946407
Patient 25	9.46E+08	2219886	173755	7.827204
Patient 26	1.06E+09	2097571	79708	3.800014
Patient 27	1.14E+08	278083	4446	1.598803
Patient 28	8.94E+08	536626	69129	12.88216
Patient 29	1.24E+08	274285	6909	2.518913

TABLE 4.3: Results of the CD8+ cell quantification.

Chapter 5

Conclusion

The tumor immune contexture, defined as the type and density of tumor-infiltrating immune cells, has a strong impact on patients' prognosis and response to therapy. In particular, high infiltration of cytotoxic CD8+ T cells, which has the capability of recognize and kill tumor cells, has been associated with a good prognosis. Thus, the quantification of CD8+ T cells is of paramount importance for the stratification of cancer patients and for the development of effective combination therapies.

In this thesis, we developed two computational approaches for the quantification of CD8+ T cells and their subtypes. First, we developed a deconvolution algorithm, CD8quant, which quantifies the relative cell fractions of naïve, central memory, and effector memory CD8+ T cells from bulk tumor RNA sequencing (RNA-seq) data. Second, we optimized a bioinformatics pipeline for the for the quantification of total CD8+ T cells from immunohistochemistry (IHC) images of tissue slides derived from formalin-fixed paraffin-embedded (FFPE) samples.

As we already acknowledged, RNA-Seq data are difficult to handle and process due to their complexity. Therefore, in the deconvolutional section (Chapter 4, page 33), the necessity of a careful preprocessing was made clear: the quality of the results we proposed strikingly improved after the detection and filtering of libraries that seemed not suitable for the analysis we were trying to carry on. For example, in a study that focuses on CD8+ T cells, we could only accept samples with at least high levels of *CD8A*

and *CD8B* gene expression. Once reassured about the quality of the data, the attention switched to the extraction of the genes that could represent the different T cell subsets of interest (CM, EM, N), the so-called *signature genes*, searched by processing processing data of healthy humans. The real challenge was to detect an universal gene set, applicable to every patient, healthy or not. The entire process has been computationally demanding, but the aim of a robust signature gene matrix could only be pursued through the testing of a high number (almost 600, see A.1 with candidates and B.2 with results) of different candidates on a different, but comparable dataset. CD8quant, validated on 200 simulated RNA-seq data sets, obtained high deconvolution performance.

Besides this, a parallel aspect that is worth to outline is referred to the feasibility of the comparison between FF and FFPE stored data; the results obtained primarly in Esteve-Codina et al. 2017 and then confirmed by our analysis suggest that, under certain conditions, it is definitely possible to substitute the FF technique with the handier FFPE one. Surely, the initial check on the quality of the samples is a mandatory habit to avoid unprocessable and therefore misleading data. In fact, the one we solely referred to as *Pair 2* has been immediately detected as poorely qualitative. The decision not to exclude it from the analysis but to persevere in trying to process it was intentionally made to highlight the difficulties that can be encountered without a proper preprocessing.

What confirmed the robustness of the algorithm *CD8quant* and the *CD8sig* has been the application of the algorithm itself on the Codina dataset: testing a signature matrix, extracted from healthy donors, on data belonging to cancerous (glioblastoma) donors, was the real challenge. The results seem absolutely encouraging for two reasons: the feasibility of the comparison between FF and FFPE data is confirmed and the quality and strength of the pair algorithm-signature is established by a further proof.

Given the availability of large collections of FFPE tumor samples exploited for image pathology, and now also for RNA-seq, these methods represent valuable tools for the efficient extraction of immunological features from tumor data and may provide mechanistic insights for the optimization of anti-cancer treatments.

Concerning the imaging section, despite, as we already recalled, the possibilities of divulgation of precise information on the data are strongly reduced by legal restrictions, the project revealed itself as a convincing complementary approach in the quantification of cell subsets in the cancerous contexture. The pipeline is not easily accessible to a simple clinician, whose competences are definitely far from the informatic boundaries. However, for whom is qualified enough to handle it, the quantification tool will be available and free to adapt. In fact, besides the precision of the counting, one of the strengths of the pipeline is the possibility to intervene and modify according to necessities, as well as exploit only one sub-tool at a time, regardless of the original flow.

Future work will focus on the integration of the strengths of these two approaches: the image results (i.e. absolute cell counts on total CD8+ T cells) will be used to scale the deconvolution results (i.e. relative cell fractions, but available for more cell types) and obtain an absolute quantification of naïve, effector memory and central memory CD8+ T cells infiltrating human tumors. The hope is that the ambivalent aim of this thesis has been reached: summing up the state of the art in the bioinformatic tools related to cancer immunology and, simultaneously, approaching an integration attempt of two of the most promising methodologies, RNA-Sequencing and IHC imaging. At the moment, the particular kind of data used in the experimental section is not easy to retrieve online, firstly because human-bound studies are still less popular than mice experiments, secondarily because the consent to share sensitive data on delicate diseases, like cancer, is complicated to obtain. Ideally, it is desirable a global effort on the production of data, especially CD8+ which, as seen in Chapter 1, are strongly related to cancer patients prognosis. More data implies more effort in the topic, and

therefore achieving a shared consensus, that will be exceptionally relevant in the run towards personalized medicine, a powerful instrument to make cancer treatable.

Chapter 6

Author Contributions

Z.T. conceived the project. F.F. selected Bonnal and Codina datasets. A.C. selected Pulko dataset. A.C. obtained and preprocessed Bonnal, Codina and Pulko datasets. A.C. wrote the deconvolution algorithm. F.F. conceived the validation process of the deconvolution algorithm. A.C. applied the validation of the deconvolution algorithm. A.C and F.F. interpreted the results of the deconvolution algorithm. C.M. wrote the imaging IHC quantification pipeline. C.M. and A.C. applied and interpreted the results of the imaging IHC quantification pipeline. A.C. wrote the manuscript. B.D.C., Z.T. and F.F. revised the manuscript.

- Z.T. = Zlatko Trajanoski
- F.F. = Francesca Finotello
- A.C. = Antonio Collesei
- C.M. = Clemens Mayer
- B.D.C. = Barbara Di Camillo

Appendix A

Materials and Methods - Appendix

A.1 Signature gene matrix - Candidates' characterization

Signature ID	pval	K	logfc
Sig_1vs1_001	5,00E-04	0.8	0
Sig_1vs1_002	5,00E-04	0.85	0
Sig_1vs1_003	5,00E-04	0.9	0
Sig_1vs1_004	5,00E-04	0.95	0
Sig_1vs1_005	5,00E-04	1	0
Sig_1vs1_006	5,00E-04	1.05	0
Sig_1vs1_007	5,00E-04	1.1	0
Sig_1vs1_008	5,00E-04	1.15	0
Sig_1vs1_009	5,00E-04	1.2	0
Sig_1vs1_010	5,00E-04	1.25	0
Sig_1vs1_011	5,00E-04	1.3	0
Sig_1vs1_012	5,00E-04	0.8	0.5
Sig_1vs1_013	5,00E-04	0.85	0.5
Sig_1vs1_014	5,00E-04	0.9	0.5
Sig_1vs1_015	5,00E-04	0.95	0.5
Sig_1vs1_016	5,00E-04	1	0.5
Sig_1vs1_017	5,00E-04	1.05	0.5

Sig_1vs1_018	5,00E-04	1.1	0.5
Sig_1vs1_019	5,00E-04	1.15	0.5
Sig_1vs1_020	5,00E-04	1.2	0.5
Sig_1vs1_021	5,00E-04	1.25	0.5
Sig_1vs1_022	5,00E-04	1.3	0.5
Sig_1vs1_023	5,00E-04	0.8	1
Sig_1vs1_024	5,00E-04	0.85	1
Sig_1vs1_025	5,00E-04	0.9	1
Sig_1vs1_026	5,00E-04	0.95	1
Sig_1vs1_027	5,00E-04	1	1
Sig_1vs1_028	5,00E-04	1.05	1
Sig_1vs1_029	5,00E-04	1.1	1
Sig_1vs1_030	5,00E-04	1.15	1
Sig_1vs1_031	5,00E-04	1.2	1
Sig_1vs1_032	5,00E-04	1.25	1
Sig_1vs1_033	5,00E-04	1.3	1
Sig_1vs1_034	0.001	0.8	0
Sig_1vs1_035	0.001	0.85	0
Sig_1vs1_036	0.001	0.9	0
Sig_1vs1_037	0.001	0.95	0
Sig_1vs1_038	0.001	1	0
Sig_1vs1_039	0.001	1.05	0
Sig_1vs1_040	0.001	1.1	0
Sig_1vs1_041	0.001	1.15	0
Sig_1vs1_042	0.001	1.2	0
Sig_1vs1_043	0.001	1.25	0
Sig_1vs1_044	0.001	1.3	0
Sig_1vs1_045	0.001	0.8	0.5
Sig_1vs1_046	0.001	0.85	0.5

Sig_1vs1_047	0.001	0.9	0.5
Sig_1vs1_048	0.001	0.95	0.5
Sig_1vs1_049	0.001	1	0.5
Sig_1vs1_050	0.001	1.05	0.5
Sig_1vs1_051	0.001	1.1	0.5
Sig_1vs1_052	0.001	1.15	0.5
Sig_1vs1_053	0.001	1.2	0.5
Sig_1vs1_054	0.001	1.25	0.5
Sig_1vs1_055	0.001	1.3	0.5
Sig_1vs1_056	0.001	0.8	1
Sig_1vs1_057	0.001	0.85	1
Sig_1vs1_058	0.001	0.9	1
Sig_1vs1_059	0.001	0.95	1
Sig_1vs1_060	0.001	1	1
Sig_1vs1_061	0.001	1.05	1
Sig_1vs1_062	0.001	1.1	1
Sig_1vs1_063	0.001	1.15	1
Sig_1vs1_064	0.001	1.2	1
Sig_1vs1_065	0.001	1.25	1
Sig_1vs1_066	0.001	1.3	1
Sig_1vs1_067	0.005	0.8	0
Sig_1vs1_068	0.005	0.85	0
Sig_1vs1_069	0.005	0.9	0
Sig_1vs1_070	0.005	0.95	0
Sig_1vs1_071	0.005	1	0
Sig_1vs1_072	0.005	1.05	0
Sig_1vs1_073	0.005	1.1	0
Sig_1vs1_074	0.005	1.15	0
Sig_1vs1_075	0.005	1.2	0

Sig_1vs1_076	0.005	1.25	0
Sig_1vs1_077	0.005	1.3	0
Sig_1vs1_078	0.005	0.8	0.5
Sig_1vs1_079	0.005	0.85	0.5
Sig_1vs1_080	0.005	0.9	0.5
Sig_1vs1_081	0.005	0.95	0.5
Sig_1vs1_082	0.005	1	0.5
Sig_1vs1_083	0.005	1.05	0.5
Sig_1vs1_084	0.005	1.1	0.5
Sig_1vs1_085	0.005	1.15	0.5
Sig_1vs1_086	0.005	1.2	0.5
Sig_1vs1_087	0.005	1.25	0.5
Sig_1vs1_088	0.005	1.3	0.5
Sig_1vs1_089	0.005	0.8	1
Sig_1vs1_090	0.005	0.85	1
Sig_1vs1_091	0.005	0.9	1
Sig_1vs1_092	0.005	0.95	1
Sig_1vs1_093	0.005	1	1
Sig_1vs1_094	0.005	1.05	1
Sig_1vs1_095	0.005	1.1	1
Sig_1vs1_096	0.005	1.15	1
Sig_1vs1_097	0.005	1.2	1
Sig_1vs1_098	0.005	1.25	1
Sig_1vs1_099	0.005	1.3	1
Sig_1vs1_100	0.01	0.8	0
Sig_1vs1_101	0.01	0.85	0
Sig_1vs1_102	0.01	0.9	0
Sig_1vs1_103	0.01	0.95	0
Sig_1vs1_104	0.01	1	0

Sig_1vs1_105	0.01	1.05	0
Sig_1vs1_106	0.01	1.1	0
Sig_1vs1_107	0.01	1.15	0
Sig_1vs1_108	0.01	1.2	0
Sig_1vs1_109	0.01	1.25	0
Sig_1vs1_110	0.01	1.3	0
Sig_1vs1_111	0.01	0.8	0.5
Sig_1vs1_112	0.01	0.85	0.5
Sig_1vs1_113	0.01	0.9	0.5
Sig_1vs1_114	0.01	0.95	0.5
Sig_1vs1_115	0.01	1	0.5
Sig_1vs1_116	0.01	1.05	0.5
Sig_1vs1_117	0.01	1.1	0.5
Sig_1vs1_118	0.01	1.15	0.5
Sig_1vs1_119	0.01	1.2	0.5
Sig_1vs1_120	0.01	1.25	0.5
Sig_1vs1_121	0.01	1.3	0.5
Sig_1vs1_122	0.01	0.8	1
Sig_1vs1_123	0.01	0.85	1
Sig_1vs1_124	0.01	0.9	1
Sig_1vs1_125	0.01	0.95	1
Sig_1vs1_126	0.01	1	1
Sig_1vs1_127	0.01	1.05	1
Sig_1vs1_128	0.01	1.1	1
Sig_1vs1_129	0.01	1.15	1
Sig_1vs1_130	0.01	1.2	1
Sig_1vs1_131	0.01	1.25	1
Sig_1vs1_132	0.01	1.3	1
Sig_1vs1_133	0.05	0.8	0

Sig_1vs1_134	0.05	0.85	0
Sig_1vs1_135	0.05	0.9	0
Sig_1vs1_136	0.05	0.95	0
Sig_1vs1_137	0.05	1	0
Sig_1vs1_138	0.05	1.05	0
Sig_1vs1_139	0.05	1.1	0
Sig_1vs1_140	0.05	1.15	0
Sig_1vs1_141	0.05	1.2	0
Sig_1vs1_142	0.05	1.25	0
Sig_1vs1_143	0.05	1.3	0
Sig_1vs1_144	0.05	0.8	0.5
Sig_1vs1_145	0.05	0.85	0.5
Sig_1vs1_146	0.05	0.9	0.5
Sig_1vs1_147	0.05	0.95	0.5
Sig_1vs1_148	0.05	1	0.5
Sig_1vs1_149	0.05	1.05	0.5
Sig_1vs1_150	0.05	1.1	0.5
Sig_1vs1_151	0.05	1.15	0.5
Sig_1vs1_152	0.05	1.2	0.5
Sig_1vs1_153	0.05	1.25	0.5
Sig_1vs1_154	0.05	1.3	0.5
Sig_1vs1_155	0.05	0.8	1
Sig_1vs1_156	0.05	0.85	1
Sig_1vs1_157	0.05	0.9	1
Sig_1vs1_158	0.05	0.95	1
Sig_1vs1_159	0.05	1	1
Sig_1vs1_160	0.05	1.05	1
Sig_1vs1_161	0.05	1.1	1
Sig_1vs1_162	0.05	1.15	1

Sig_1vs1_163	0.05	1.2	1
Sig_1vs1_164	0.05	1.25	1
Sig_1vs1_165	0.05	1.3	1
Sig_1vs1_166	0.1	0.8	0
Sig_1vs1_167	0.1	0.85	0
Sig_1vs1_168	0.1	0.9	0
Sig_1vs1_169	0.1	0.95	0
Sig_1vs1_170	0.1	1	0
Sig_1vs1_171	0.1	1.05	0
Sig_1vs1_172	0.1	1.1	0
Sig_1vs1_173	0.1	1.15	0
Sig_1vs1_174	0.1	1.2	0
Sig_1vs1_175	0.1	1.25	0
Sig_1vs1_176	0.1	1.3	0
Sig_1vs1_177	0.1	0.8	0.5
Sig_1vs1_178	0.1	0.85	0.5
Sig_1vs1_179	0.1	0.9	0.5
Sig_1vs1_180	0.1	0.95	0.5
Sig_1vs1_181	0.1	1	0.5
Sig_1vs1_182	0.1	1.05	0.5
Sig_1vs1_183	0.1	1.1	0.5
Sig_1vs1_184	0.1	1.15	0.5
Sig_1vs1_185	0.1	1.2	0.5
Sig_1vs1_186	0.1	1.25	0.5
Sig_1vs1_187	0.1	1.3	0.5
Sig_1vs1_188	0.1	0.8	1
Sig_1vs1_189	0.1	0.85	1
Sig_1vs1_190	0.1	0.9	1
Sig_1vs1_191	0.1	0.95	1

Sig_1vs1_192	0.1	1	1
Sig_1vs1_193	0.1	1.05	1
Sig_1vs1_194	0.1	1.1	1
Sig_1vs1_195	0.1	1.15	1
Sig_1vs1_196	0.1	1.2	1
Sig_1vs1_197	0.1	1.25	1
Sig_1vs1_198	0.1	1.3	1
Sig_1vs1_199	0.15	0.8	0
Sig_1vs1_200	0.15	0.85	0
Sig_1vs1_201	0.15	0.9	0
Sig_1vs1_202	0.15	0.95	0
Sig_1vs1_203	0.15	1	0
Sig_1vs1_204	0.15	1.05	0
Sig_1vs1_205	0.15	1.1	0
Sig_1vs1_206	0.15	1.15	0
Sig_1vs1_207	0.15	1.2	0
Sig_1vs1_208	0.15	1.25	0
Sig_1vs1_209	0.15	1.3	0
Sig_1vs1_210	0.15	0.8	0.5
Sig_1vs1_211	0.15	0.85	0.5
Sig_1vs1_212	0.15	0.9	0.5
Sig_1vs1_213	0.15	0.95	0.5
Sig_1vs1_214	0.15	1	0.5
Sig_1vs1_215	0.15	1.05	0.5
Sig_1vs1_216	0.15	1.1	0.5
Sig_1vs1_217	0.15	1.15	0.5
Sig_1vs1_218	0.15	1.2	0.5
Sig_1vs1_219	0.15	1.25	0.5
Sig_1vs1_220	0.15	1.3	0.5

Sig_1vs1_221	0.15	0.8	1
Sig_1vs1_222	0.15	0.85	1
Sig_1vs1_223	0.15	0.9	1
Sig_1vs1_224	0.15	0.95	1
Sig_1vs1_225	0.15	1	1
Sig_1vs1_226	0.15	1.05	1
Sig_1vs1_227	0.15	1.1	1
Sig_1vs1_228	0.15	1.15	1
Sig_1vs1_229	0.15	1.2	1
Sig_1vs1_230	0.15	1.25	1
Sig_1vs1_231	0.15	1.3	1
Sig_1vs1_232	0.2	0.8	0
Sig_1vs1_233	0.2	0.85	0
Sig_1vs1_234	0.2	0.9	0
Sig_1vs1_235	0.2	0.95	0
Sig_1vs1_236	0.2	1	0
Sig_1vs1_237	0.2	1.05	0
Sig_1vs1_238	0.2	1.1	0
Sig_1vs1_239	0.2	1.15	0
Sig_1vs1_240	0.2	1.2	0
Sig_1vs1_241	0.2	1.25	0
Sig_1vs1_242	0.2	1.3	0
Sig_1vs1_243	0.2	0.8	0.5
Sig_1vs1_244	0.2	0.85	0.5
Sig_1vs1_245	0.2	0.9	0.5
Sig_1vs1_246	0.2	0.95	0.5
Sig_1vs1_247	0.2	1	0.5
Sig_1vs1_248	0.2	1.05	0.5
Sig_1vs1_249	0.2	1.1	0.5

Sig_1vs1_250	0.2	1.15	0.5
Sig_1vs1_251	0.2	1.2	0.5
Sig_1vs1_252	0.2	1.25	0.5
Sig_1vs1_253	0.2	1.3	0.5
Sig_1vs1_254	0.2	0.8	1
Sig_1vs1_255	0.2	0.85	1
Sig_1vs1_256	0.2	0.9	1
Sig_1vs1_257	0.2	0.95	1
Sig_1vs1_258	0.2	1	1
Sig_1vs1_259	0.2	1.05	1
Sig_1vs1_260	0.2	1.1	1
Sig_1vs1_261	0.2	1.15	1
Sig_1vs1_262	0.2	1.2	1
Sig_1vs1_263	0.2	1.25	1
Sig_1vs1_264	0.2	1.3	1
Sig_1vs1_265	0.25	0.8	0
Sig_1vs1_266	0.25	0.85	0
Sig_1vs1_267	0.25	0.9	0
Sig_1vs1_268	0.25	0.95	0
Sig_1vs1_269	0.25	1	0
Sig_1vs1_270	0.25	1.05	0
Sig_1vs1_271	0.25	1.1	0
Sig_1vs1_272	0.25	1.15	0
Sig_1vs1_273	0.25	1.2	0
Sig_1vs1_274	0.25	1.25	0
Sig_1vs1_275	0.25	1.3	0
Sig_1vs1_276	0.25	0.8	0.5
Sig_1vs1_277	0.25	0.85	0.5
Sig_1vs1_278	0.25	0.9	0.5

Sig_1vs1_279	0.25	0.95	0.5
Sig_1vs1_280	0.25	1	0.5
Sig_1vs1_281	0.25	1.05	0.5
Sig_1vs1_282	0.25	1.1	0.5
Sig_1vs1_283	0.25	1.15	0.5
Sig_1vs1_284	0.25	1.2	0.5
Sig_1vs1_285	0.25	1.25	0.5
Sig_1vs1_286	0.25	1.3	0.5
Sig_1vs1_287	0.25	0.8	1
Sig_1vs1_288	0.25	0.85	1
Sig_1vs1_289	0.25	0.9	1
Sig_1vs1_290	0.25	0.95	1
Sig_1vs1_291	0.25	1	1
Sig_1vs1_292	0.25	1.05	1
Sig_1vs1_293	0.25	1.1	1
Sig_1vs1_294	0.25	1.15	1
Sig_1vs1_295	0.25	1.2	1
Sig_1vs1_296	0.25	1.25	1
Sig_1vs1_297	0.25	1.3	1

TABLE A.1: Parameter set-up of the signature genes candidates.

Appendix B

Results - Appendix

B.1 Simulation table

	CM	EM	N
CD8mix_001	0.54	0.41	0.05
CD8mix_002	0.12	0.69	0.19
CD8mix_003	0.28	0.24	0.47
CD8mix_004	0.53	0.36	0.1
CD8mix_005	0.28	0.31	0.41
CD8mix_006	0.22	0.55	0.23
CD8mix_007	0.36	0.44	0.2
CD8mix_008	0.19	0.45	0.37
CD8mix_009	0.09	0.44	0.47
CD8mix_010	0.43	0.23	0.34
CD8mix_011	0.74	0.02	0.24
CD8mix_012	0.31	0.35	0.34
CD8mix_013	0.08	0.4	0.52
CD8mix_014	0.48	0.14	0.38
CD8mix_015	0.06	0.55	0.4
CD8mix_016	0.63	0.01	0.36
CD8mix_017	0.59	0.16	0.25

CD8mix_018	0.01	0.6	0.39
CD8mix_019	0.27	0.12	0.61
CD8mix_020	0.63	0.08	0.29
CD8mix_021	0.29	0.42	0.29
CD8mix_022	0.25	0.05	0.7
CD8mix_023	0.19	0.02	0.79
CD8mix_024	0.3	0.33	0.37
CD8mix_025	0.39	0.15	0.46
CD8mix_026	0.15	0.38	0.47
CD8mix_027	0.07	0.16	0.77
CD8mix_028	0.47	0.23	0.31
CD8mix_029	0.11	0.5	0.39
CD8mix_030	0.06	0.09	0.85
CD8mix_031	0.18	0.25	0.57
CD8mix_032	0.51	0.11	0.37
CD8mix_033	0.54	0.29	0.17
CD8mix_034	0.44	0.48	0.08
CD8mix_035	0.28	0.42	0.3
CD8mix_036	0.37	0.25	0.38
CD8mix_037	0.23	0.33	0.44
CD8mix_038	0.32	0.4	0.29
CD8mix_039	0.24	0.29	0.46
CD8mix_040	0.33	0.37	0.3
CD8mix_041	0.31	0.58	0.11
CD8mix_042	0.18	0.26	0.56
CD8mix_043	0.32	0.07	0.61
CD8mix_044	0.54	0.08	0.38
CD8mix_045	0.5	0.29	0.2
CD8mix_046	0.34	0.42	0.24

CD8mix_047	0.48	0.26	0.26
CD8mix_048	0.44	0.23	0.33
CD8mix_049	0.49	0.34	0.18
CD8mix_050	0.31	0.18	0.51
CD8mix_051	0.03	0.52	0.45
CD8mix_052	0.54	0.18	0.27
CD8mix_053	0.1	0.42	0.47
CD8mix_054	0.74	0.17	0.09
CD8mix_055	0.49	0.3	0.22
CD8mix_056	0.3	0.32	0.38
CD8mix_057	0.01	0.5	0.49
CD8mix_058	0.16	0.6	0.24
CD8mix_059	0.38	0.4	0.23
CD8mix_060	0.61	0.03	0.36
CD8mix_061	0.56	0.43	0.01
CD8mix_062	0.07	0.34	0.59
CD8mix_063	0.1	0.35	0.55
CD8mix_064	0.24	0.53	0.23
CD8mix_065	0.34	0.26	0.39
CD8mix_066	0.36	0.22	0.42
CD8mix_067	0.51	0.3	0.19
CD8mix_068	0.5	0.2	0.3
CD8mix_069	0.36	0.37	0.27
CD8mix_070	0.36	0.39	0.25
CD8mix_071	0.44	0.19	0.37
CD8mix_072	0.32	0.57	0.11
CD8mix_073	0.15	0.59	0.26
CD8mix_074	0.12	0.49	0.39
CD8mix_075	0.14	0.08	0.78

CD8mix_076	0.49	0.04	0.47
CD8mix_077	0.16	0.52	0.31
CD8mix_078	0.35	0.24	0.41
CD8mix_079	0.08	0.62	0.3
CD8mix_080	0.37	0.38	0.25
CD8mix_081	0.05	0.51	0.44
CD8mix_082	0.68	0.22	0.1
CD8mix_083	0.22	0.59	0.19
CD8mix_084	0.1	0.57	0.33
CD8mix_085	0.44	0.13	0.43
CD8mix_086	0.77	0.07	0.17
CD8mix_087	0.08	0.48	0.44
CD8mix_088	0.27	0.28	0.45
CD8mix_089	0.49	0.24	0.27
CD8mix_090	0.4	0.48	0.13
CD8mix_091	0.13	0.53	0.34
CD8mix_092	0.39	0.19	0.42
CD8mix_093	0.34	0.16	0.5
CD8mix_094	0.27	0.72	0.01
CD8mix_095	0.37	0.25	0.39
CD8mix_096	0.85	0.01	0.14
CD8mix_097	0.32	0.39	0.29
CD8mix_098	0.23	0.47	0.3
CD8mix_099	0.26	0.41	0.33
CD8mix_100	0.15	0.2	0.65

TABLE B.1: Simulation table to build all the mixture matrices.

B.2 Overall deconvolutional results - CD8sig selection

Signature ID	gene num	RMSE all	RMSE CM	RMSE EM	RMSE N	corr all	corr CM	corr EM	corr N
Sig_1vs1_001	254	0.3627	0.4644	0.1792	0.3833	0.0339	0.1861	0.4048	0.3662
Sig_1vs1_002	233	0.4006	0.5345	0.2194	0.3842	0.0247	0.0962	0.402	0.3103
Sig_1vs1_003	213	0.3668	0.4813	0.191	0.3681	0.0757	0.1866	0.4746	0.5419
Sig_1vs1_004	194	0.3056	0.3902	0.1784	0.3102	0.1377	0.2788	0.4254	0.6984
Sig_1vs1_005	175	0.2666	0.3263	0.1935	0.2634	0.1964	0.2543	0.3818	0.6214
Sig_1vs1_006	149	0.2315	0.2688	0.1852	0.2328	0.2263	0.2362	0.3953	0.5428
Sig_1vs1_007	126	0.3263	0.4236	0.1982	0.3175	0.1022	0.1653	0.3896	0.571
Sig_1vs1_008	101	0.3583	0.4647	0.2219	0.3462	-0.0176	-0.2185	0.1621	0.5458
Sig_1vs1_009	90	0.4304	0.591	0.3017	0.3397	0.0452	0.1952	0.2588	0.5792
Sig_1vs1_010	79	0.4377	0.6034	0.3139	0.3347	0.0536	0.2902	0.3944	0.5582
Sig_1vs1_011	71	0.4203	0.5784	0.2888	0.3348	0.06	0.2554	0.463	0.5254
Sig_1vs1_012	193	0.4922	0.6784	0.3448	0.3842	-0.0089	0.1598	0.389	0.4474
Sig_1vs1_013	179	0.5011	0.6904	0.3591	0.3843	-0.0187	0.1405	0.3113	0.4137
Sig_1vs1_014	166	0.504	0.6941	0.3634	0.3851	-0.0216	0.1829	0.2396	0.3789
Sig_1vs1_015	152	0.5079	0.699	0.3654	0.3897	-0.0279	-0.0606	-0.0575	0.299
Sig_1vs1_016	143	0.2603	0.3432	0.2023	0.2112	0.2188	0.1234	0.4639	0.5002
Sig_1vs1_017	130	0.2675	0.3569	0.2174	0.2002	0.2113	0.1478	0.4531	0.4943
Sig_1vs1_018	113	0.3298	0.4481	0.265	0.235	0.1278	0.1153	0.3639	0.523
Sig_1vs1_019	93	0.4548	0.6282	0.3237	0.3479	0.041	0.2754	0.4091	0.5689
Sig_1vs1_020	85	0.449	0.6198	0.3325	0.3318	0.0559	0.3671	0.4215	0.5833
Sig_1vs1_021	76	0.4406	0.6076	0.3206	0.3324	0.0573	0.3328	0.4289	0.5411
Sig_1vs1_022	69	0.4149	0.5709	0.2838	0.3318	0.0667	0.2551	0.478	0.5253
Sig_1vs1_023	82	0.2885	0.2827	0.2532	0.3251	0.2032	0.1472	0.2742	0.2498
Sig_1vs1_024	78	0.2942	0.2831	0.2561	0.3375	0.1774	0.1534	0.2793	0.1367
Sig_1vs1_025	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_026	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529

Sig_1vs1_027	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_028	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_029	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_030	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_031	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_032	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_033	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_034	276	0.3171	0.3724	0.1721	0.3652	0.0999	0.2722	0.4456	0.5623
Sig_1vs1_035	254	0.3506	0.447	0.1852	0.367	0.0776	0.1668	0.4364	0.5622
Sig_1vs1_036	225	0.3647	0.4726	0.186	0.3756	0.0673	0.1839	0.4577	0.5312
Sig_1vs1_037	205	0.3089	0.3853	0.1729	0.3286	0.1257	0.2671	0.4349	0.676
Sig_1vs1_038	183	0.2817	0.3419	0.1933	0.2896	0.1672	0.2633	0.3661	0.5877
Sig_1vs1_039	153	0.2559	0.3095	0.1869	0.2564	0.1982	0.2472	0.3913	0.5655
Sig_1vs1_040	129	0.3468	0.4558	0.208	0.3314	0.0917	0.1833	0.3923	0.5549
Sig_1vs1_041	102	0.3578	0.4641	0.2212	0.346	-0.0157	-0.2119	0.1665	0.5491
Sig_1vs1_042	91	0.4296	0.5899	0.3012	0.3391	0.0462	0.1967	0.2608	0.578
Sig_1vs1_043	80	0.4374	0.603	0.3134	0.3347	0.0538	0.2844	0.3955	0.5568
Sig_1vs1_044	72	0.4207	0.5789	0.289	0.335	0.0601	0.2551	0.4637	0.5244
Sig_1vs1_045	196	0.4917	0.6777	0.3442	0.384	-0.008	0.1692	0.3996	0.4107
Sig_1vs1_046	181	0.5009	0.6902	0.3589	0.3841	-0.0184	0.1421	0.3204	0.416
Sig_1vs1_047	167	0.5042	0.6944	0.3633	0.3854	-0.0219	0.1728	0.2382	0.3798
Sig_1vs1_048	153	0.5079	0.699	0.3654	0.3897	-0.0279	-0.0577	-0.0575	0.299
Sig_1vs1_049	144	0.2603	0.3429	0.2016	0.2122	0.2222	0.1368	0.4712	0.4971
Sig_1vs1_050	131	0.2673	0.3566	0.2168	0.2006	0.2094	0.1416	0.4538	0.4907
Sig_1vs1_051	114	0.3297	0.448	0.2646	0.2354	0.1275	0.1179	0.3635	0.5221
Sig_1vs1_052	94	0.4548	0.6282	0.3232	0.3483	0.0416	0.28	0.4135	0.5725
Sig_1vs1_053	86	0.4493	0.6202	0.3327	0.3321	0.0555	0.3628	0.4098	0.5824
Sig_1vs1_054	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_055	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246

Sig_1vs1_056	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_057	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_058	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_059	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_060	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_061	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_062	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_063	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_064	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_065	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_066	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_067	344	0.3825	0.2089	0.4934	0.3898	0.0448	0.4273	0.5337	NA
Sig_1vs1_068	314	0.4245	0.2577	0.5676	0.3898	0.0129	0.4102	0.3787	NA
Sig_1vs1_069	268	0.4513	0.3019	0.6066	0.3898	-0.0208	0.1305	0.3385	NA
Sig_1vs1_070	239	0.4609	0.3104	0.6236	0.3898	-0.026	0.1939	0.249	NA
Sig_1vs1_071	212	0.2877	0.3484	0.1931	0.2995	0.1581	0.2682	0.375	0.5263
Sig_1vs1_072	177	0.2591	0.3112	0.1862	0.2643	0.1974	0.2975	0.3914	0.5286
Sig_1vs1_073	141	0.3616	0.4775	0.2154	0.3435	0.0812	0.1782	0.396	0.508
Sig_1vs1_074	109	0.3908	0.519	0.2245	0.3719	-0.0176	-0.1651	0.2596	0.443
Sig_1vs1_075	94	0.4447	0.6115	0.3032	0.3569	0.0228	0.1254	0.3179	0.5491
Sig_1vs1_076	81	0.4389	0.6046	0.3124	0.3387	0.0466	0.2784	0.379	0.5459
Sig_1vs1_077	72	0.4207	0.5789	0.289	0.335	0.0601	0.2551	0.4637	0.5244
Sig_1vs1_078	212	0.496	0.6841	0.3563	0.3782	-0.0078	0.2477	0.3757	0.5168
Sig_1vs1_079	194	0.4994	0.6883	0.3577	0.383	-0.0164	0.1693	0.3269	0.4395
Sig_1vs1_080	175	0.5037	0.6937	0.3634	0.3846	-0.0212	0.1807	0.2288	0.3867
Sig_1vs1_081	159	0.5079	0.699	0.3654	0.3896	-0.0279	-0.0189	-0.0575	0.299
Sig_1vs1_082	149	0.2591	0.3403	0.2021	0.2116	0.2216	0.1338	0.4613	0.4971
Sig_1vs1_083	135	0.2689	0.3586	0.2167	0.2035	0.2051	0.1412	0.4401	0.4877
Sig_1vs1_084	115	0.3303	0.449	0.2651	0.2354	0.125	0.111	0.3574	0.5237

Sig_1vs1_085	95	0.4543	0.6276	0.3221	0.3485	0.0426	0.2863	0.43	0.5691
Sig_1vs1_086	86	0.4493	0.6202	0.3327	0.3321	0.0555	0.3628	0.4098	0.5824
Sig_1vs1_087	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_088	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_089	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_090	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_091	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_092	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_093	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_094	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_095	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_096	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_097	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_098	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_099	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_100	380	0.1759	0.211	0.1479	0.1626	0.3737	0.3791	0.5826	0.5121
Sig_1vs1_101	349	0.1816	0.2162	0.147	0.1749	0.334	0.2983	0.5394	0.5307
Sig_1vs1_102	297	0.1898	0.2297	0.1548	0.1771	0.3296	0.4379	0.4974	0.6763
Sig_1vs1_103	257	**0.1556**	0.1799	0.1525	0.1306	**0.4734**	0.4937	0.5283	0.6932
Sig_1vs1_104	224	0.2254	0.2793	0.1719	0.2119	0.2058	0.3056	0.4833	0.4359
Sig_1vs1_105	186	0.192	0.2306	0.1641	0.1748	0.2627	0.2556	0.4951	0.4317
Sig_1vs1_106	149	0.2496	0.322	0.1904	0.2168	0.1402	0.2239	0.4853	0.4109
Sig_1vs1_107	115	0.4218	0.5741	0.2595	0.3698	-0.0079	-0.1715	0.292	0.515
Sig_1vs1_108	96	0.4756	0.6557	0.3488	0.3564	0.0204	0.4123	0.3133	0.5702
Sig_1vs1_109	82	0.4405	0.6067	0.3139	0.3399	0.0438	0.2781	0.3683	0.5419
Sig_1vs1_110	73	0.4208	0.5792	0.2887	0.3355	0.0591	0.2422	0.4558	0.5258
Sig_1vs1_111	213	0.496	0.6841	0.3567	0.3779	-0.0079	0.2457	0.3669	0.5179
Sig_1vs1_112	195	0.4992	0.6879	0.3573	0.383	-0.0159	0.1771	0.3343	0.438
Sig_1vs1_113	176	0.5036	0.6936	0.3634	0.3845	-0.021	0.1917	0.2366	0.3813

Sig_1vs1_114	160	0.5079	0.699	0.3654	0.3896	-0.0279	-0.0148	-0.0575	0.299
Sig_1vs1_115	149	0.2591	0.3403	0.2021	0.2116	0.2216	0.1338	0.4613	0.4971
Sig_1vs1_116	135	0.2689	0.3586	0.2167	0.2035	0.2051	0.1412	0.4401	0.4877
Sig_1vs1_117	115	0.3303	0.449	0.2651	0.2354	0.125	0.111	0.3574	0.5237
Sig_1vs1_118	95	0.4543	0.6276	0.3221	0.3485	0.0426	0.2863	0.43	0.5691
Sig_1vs1_119	86	0.4493	0.6202	0.3327	0.3321	0.0555	0.3628	0.4098	0.5824
Sig_1vs1_120	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_121	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_122	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_123	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_124	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_125	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_126	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_127	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_128	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_129	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_130	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_131	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_132	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_133	53	0.2769	0.2918	0.2739	0.2647	0.1564	0.1389	0.3126	0.1056
Sig_1vs1_134	480	0.312	0.2036	0.3316	0.375	0.0189	0.1558	0.3916	0.3523
Sig_1vs1_135	402	0.3123	0.2017	0.3597	0.3501	0.0657	0.1719	0.4222	0.5702
Sig_1vs1_136	335	0.2972	0.1872	0.3615	0.315	0.1011	0.3169	0.4902	0.4937
Sig_1vs1_137	272	0.2843	0.3515	0.1622	0.3045	0.1107	0.2687	0.4964	0.2665
Sig_1vs1_138	212	0.2571	0.3253	0.1742	0.2493	0.1304	0.2635	0.4705	0.4063
Sig_1vs1_139	163	0.3127	0.4115	0.1964	0.2923	0.0783	0.2485	0.484	0.3519
Sig_1vs1_140	123	0.3144	0.1635	0.3673	0.3673	0.1078	0.4979	0.5116	0.5337
Sig_1vs1_141	99	0.4812	0.6634	0.3494	0.364	0.0106	0.3543	0.3058	0.5325
Sig_1vs1_142	83	0.4411	0.6077	0.3143	0.3401	0.0429	0.2709	0.3685	0.5433

Sig_1vs1_143	74	0.4216	0.5802	0.2891	0.3362	0.0577	0.2377	0.4535	0.5226
Sig_1vs1_144	224	0.497	0.6851	0.356	0.3806	-0.009	0.2634	0.3098	0.5017
Sig_1vs1_145	205	0.4996	0.6884	0.3576	0.3833	-0.0134	0.2984	0.3816	0.491
Sig_1vs1_146	186	0.5049	0.6953	0.3631	0.3866	-0.023	0.169	0.2273	0.4194
Sig_1vs1_147	169	0.508	0.6991	0.3653	0.3898	-0.0281	-0.1862	-0.0371	NA
Sig_1vs1_148	153	0.258	0.3381	0.2044	0.2087	0.2205	0.1332	0.4217	0.5787
Sig_1vs1_149	138	0.2629	0.3496	0.2147	0.1974	0.195	0.0974	0.4185	0.5145
Sig_1vs1_150	116	0.3262	0.4436	0.2589	0.2353	0.1148	0.047	0.3586	0.518
Sig_1vs1_151	96	0.4566	0.6303	0.3237	0.3515	0.0346	0.246	0.3695	0.5446
Sig_1vs1_152	87	0.4512	0.6235	0.3299	0.3364	0.0592	0.3722	0.4366	0.6014
Sig_1vs1_153	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_154	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_155	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_156	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_157	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_158	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_159	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_160	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_161	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_162	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_163	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_164	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_165	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_166	53	0.2769	0.2918	0.2739	0.2647	0.1564	0.1389	0.3126	0.1056
Sig_1vs1_167	549	0.1933	0.1904	0.2159	0.171	0.3131	0.3362	0.4467	0.4663
Sig_1vs1_168	451	0.1936	0.2089	0.2183	0.1456	0.3255	0.3418	0.4389	0.5399
Sig_1vs1_169	367	0.1843	0.214	0.1916	0.1391	0.3307	0.3754	0.4258	0.5892
Sig_1vs1_170	289	0.1698	0.179	0.1648	0.1652	0.3833	0.3645	0.4975	0.3859
Sig_1vs1_171	220	0.164	0.181	0.1627	0.1465	0.4142	0.3133	0.4536	0.5125

Sig_1vs1_172	169	0.3156	0.4149	0.1962	0.2971	0.0752	0.229	0.4804	0.328
Sig_1vs1_173	125	0.3148	0.1626	0.3658	0.3701	0.1068	0.5053	0.5339	0.4936
Sig_1vs1_174	100	0.4813	0.6635	0.3496	0.364	0.0107	0.357	0.3052	0.5321
Sig_1vs1_175	83	0.4411	0.6077	0.3143	0.3401	0.0429	0.2709	0.3685	0.5433
Sig_1vs1_176	74	0.4216	0.5802	0.2891	0.3362	0.0577	0.2377	0.4535	0.5226
Sig_1vs1_177	226	0.4971	0.6852	0.3563	0.3804	-0.0091	0.2627	0.3114	0.5031
Sig_1vs1_178	207	0.4997	0.6886	0.3579	0.3831	-0.0135	0.2907	0.3689	0.4927
Sig_1vs1_179	188	0.5049	0.6952	0.3632	0.3866	-0.0229	0.1714	0.2294	0.4205
Sig_1vs1_180	171	0.508	0.6991	0.3653	0.3898	-0.0281	-0.1796	-0.0212	0.2536
Sig_1vs1_181	155	0.2585	0.3391	0.2036	0.2098	0.225	0.1454	0.4301	0.5809
Sig_1vs1_182	139	0.2619	0.3483	0.2138	0.1967	0.2006	0.1119	0.4248	0.5267
Sig_1vs1_183	117	0.3265	0.4443	0.2582	0.2361	0.1151	0.0465	0.3614	0.5193
Sig_1vs1_184	97	0.4565	0.6302	0.3236	0.3512	0.0358	0.2535	0.3779	0.5512
Sig_1vs1_185	88	0.4501	0.6218	0.3295	0.3353	0.0618	0.3791	0.4336	0.6009
Sig_1vs1_186	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_187	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_188	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_189	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_190	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_191	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_192	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_193	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_194	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_195	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_196	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_197	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_198	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_199	53	0.2769	0.2918	0.2739	0.2647	0.1564	0.1389	0.3126	0.1056
Sig_1vs1_200	609	0.2281	0.2361	0.2683	0.1684	0.2232	0.3015	0.4374	0.4031

Sig_1vs1_201	490	0.2537	0.3003	0.2786	0.1589	0.169	0.1507	0.3762	0.4142
Sig_1vs1_202	395	0.2358	0.29	0.2448	0.1508	0.1997	0.2577	0.3943	0.5243
Sig_1vs1_203	306	0.1866	0.1929	0.1944	0.1718	0.2868	0.2346	0.378	0.3839
Sig_1vs1_204	226	0.1651	0.1825	0.1632	0.1476	0.4086	0.3107	0.4486	0.517
Sig_1vs1_205	173	0.3147	0.4148	0.1999	0.2915	0.0769	0.2208	0.4811	0.337
Sig_1vs1_206	127	0.3168	0.1639	0.3696	0.3709	0.1066	0.4984	0.5241	0.4874
Sig_1vs1_207	100	0.4813	0.6635	0.3496	0.364	0.0107	0.357	0.3052	0.5321
Sig_1vs1_208	83	0.4411	0.6077	0.3143	0.3401	0.0429	0.2709	0.3685	0.5433
Sig_1vs1_209	74	0.4216	0.5802	0.2891	0.3362	0.0577	0.2377	0.4535	0.5226
Sig_1vs1_210	226	0.4971	0.6852	0.3563	0.3804	-0.0091	0.2627	0.3114	0.5031
Sig_1vs1_211	207	0.4997	0.6886	0.3579	0.3831	-0.0135	0.2907	0.3689	0.4927
Sig_1vs1_212	188	0.5049	0.6952	0.3632	0.3866	-0.0229	0.1714	0.2294	0.4205
Sig_1vs1_213	171	0.508	0.6991	0.3653	0.3898	-0.0281	-0.1796	-0.0212	0.2536
Sig_1vs1_214	155	0.2585	0.3391	0.2036	0.2098	0.225	0.1454	0.4301	0.5809
Sig_1vs1_215	139	0.2619	0.3483	0.2138	0.1967	0.2006	0.1119	0.4248	0.5267
Sig_1vs1_216	117	0.3265	0.4443	0.2582	0.2361	0.1151	0.0465	0.3614	0.5193
Sig_1vs1_217	97	0.4565	0.6302	0.3236	0.3512	0.0358	0.2535	0.3779	0.5512
Sig_1vs1_218	88	0.4501	0.6218	0.3295	0.3353	0.0618	0.3791	0.4336	0.6009
Sig_1vs1_219	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_220	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_221	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_222	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_223	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_224	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_225	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_226	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_227	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_228	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_229	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187

Sig_1vs1_230	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_231	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_232	53	0.2769	0.2918	0.2739	0.2647	0.1564	0.1389	0.3126	0.1056
Sig_1vs1_233	647	0.2514	0.2337	0.3137	0.1912	0.1869	0.3439	0.4395	0.4267
Sig_1vs1_234	514	0.2689	0.3107	0.3079	0.1602	0.145	0.1587	0.4092	0.3857
Sig_1vs1_235	408	0.2522	0.302	0.281	0.1438	0.1575	0.2021	0.3715	0.5281
Sig_1vs1_236	311	0.1952	0.1909	0.2056	0.1886	0.2471	0.2401	0.3672	0.363
Sig_1vs1_237	228	0.1647	0.1835	0.1618	0.1467	0.409	0.2983	0.4636	0.5094
Sig_1vs1_238	174	0.3158	0.4141	0.1939	0.3002	0.0718	0.2043	0.4746	0.3215
Sig_1vs1_239	127	0.3168	0.1639	0.3696	0.3709	0.1066	0.4984	0.5241	0.4874
Sig_1vs1_240	100	0.4813	0.6635	0.3496	0.364	0.0107	0.357	0.3052	0.5321
Sig_1vs1_241	83	0.4411	0.6077	0.3143	0.3401	0.0429	0.2709	0.3685	0.5433
Sig_1vs1_242	74	0.4216	0.5802	0.2891	0.3362	0.0577	0.2377	0.4535	0.5226
Sig_1vs1_243	226	0.4971	0.6852	0.3563	0.3804	-0.0091	0.2627	0.3114	0.5031
Sig_1vs1_244	207	0.4997	0.6886	0.3579	0.3831	-0.0135	0.2907	0.3689	0.4927
Sig_1vs1_245	188	0.5049	0.6952	0.3632	0.3866	-0.0229	0.1714	0.2294	0.4205
Sig_1vs1_246	171	0.508	0.6991	0.3653	0.3898	-0.0281	-0.1796	-0.0212	0.2536
Sig_1vs1_247	155	0.2585	0.3391	0.2036	0.2098	0.225	0.1454	0.4301	0.5809
Sig_1vs1_248	139	0.2619	0.3483	0.2138	0.1967	0.2006	0.1119	0.4248	0.5267
Sig_1vs1_249	117	0.3265	0.4443	0.2582	0.2361	0.1151	0.0465	0.3614	0.5193
Sig_1vs1_250	97	0.4565	0.6302	0.3236	0.3512	0.0358	0.2535	0.3779	0.5512
Sig_1vs1_251	88	0.4501	0.6218	0.3295	0.3353	0.0618	0.3791	0.4336	0.6009
Sig_1vs1_252	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_253	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_254	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481
Sig_1vs1_255	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_256	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_257	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_258	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221

Sig_1vs1_259	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_260	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_261	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_262	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_263	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_264	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vs1_265	53	0.2769	0.2918	0.2739	0.2647	0.1564	0.1389	0.3126	0.1056
Sig_1vs1_266	53	0.2769	0.2918	0.2739	0.2647	0.1564	0.1389	0.3126	0.1056
Sig_1vs1_267	543	0.2749	0.3097	0.316	0.1758	0.123	0.1377	0.3929	0.2633
Sig_1vs1_268	427	0.2582	0.3052	0.2863	0.1579	0.1342	0.1327	0.3908	0.3697
Sig_1vs1_269	323	0.1966	0.2004	0.2015	0.1875	0.2333	0.1666	0.3811	0.2633
Sig_1vs1_270	232	0.1723	0.1914	0.1608	0.1629	0.3458	0.1742	0.4695	0.3669
Sig_1vs1_271	175	0.317	0.416	0.1964	0.2998	0.0696	0.1974	0.4654	0.3196
Sig_1vs1_272	128	0.3164	0.166	0.3676	0.371	0.1038	0.4835	0.5176	0.4794
Sig_1vs1_273	100	0.4813	0.6635	0.3496	0.364	0.0107	0.357	0.3052	0.5321
Sig_1vs1_274	83	0.4411	0.6077	0.3143	0.3401	0.0429	0.2709	0.3685	0.5433
Sig_1vs1_275	74	0.4216	0.5802	0.2891	0.3362	0.0577	0.2377	0.4535	0.5226
Sig_1vs1_276	226	0.4971	0.6852	0.3563	0.3804	-0.0091	0.2627	0.3114	0.5031
Sig_1vs1_277	207	0.4997	0.6886	0.3579	0.3831	-0.0135	0.2907	0.3689	0.4927
Sig_1vs1_278	188	0.5049	0.6952	0.3632	0.3866	-0.0229	0.1714	0.2294	0.4205
Sig_1vs1_279	171	0.508	0.6991	0.3653	0.3898	-0.0281	-0.1796	-0.0212	0.2536
Sig_1vs1_280	155	0.2585	0.3391	0.2036	0.2098	0.225	0.1454	0.4301	0.5809
Sig_1vs1_281	139	0.2619	0.3483	0.2138	0.1967	0.2006	0.1119	0.4248	0.5267
Sig_1vs1_282	117	0.3265	0.4443	0.2582	0.2361	0.1151	0.0465	0.3614	0.5193
Sig_1vs1_283	97	0.4565	0.6302	0.3236	0.3512	0.0358	0.2535	0.3779	0.5512
Sig_1vs1_284	88	0.4501	0.6218	0.3295	0.3353	0.0618	0.3791	0.4336	0.6009
Sig_1vs1_285	77	0.4402	0.6069	0.32	0.3325	0.058	0.3368	0.4249	0.543
Sig_1vs1_286	70	0.4145	0.5702	0.2827	0.3321	0.066	0.2467	0.4772	0.5246
Sig_1vs1_287	83	0.2888	0.2831	0.2533	0.3256	0.2012	0.1423	0.2741	0.2481

Sig_1vs1_288	79	0.2943	0.2836	0.2567	0.337	0.1744	0.1473	0.279	0.1328
Sig_1vs1_289	77	0.2934	0.2821	0.2562	0.3361	0.1771	0.1536	0.2795	0.1347
Sig_1vs1_290	74	0.3278	0.3234	0.2535	0.3919	0.1211	0.0179	0.2947	0.0529
Sig_1vs1_291	73	0.327	0.3155	0.2521	0.397	0.172	0.0923	0.2844	0.2221
Sig_1vs1_292	69	0.3338	0.3225	0.2531	0.4077	0.1504	0.0616	0.2862	0.1545
Sig_1vs1_293	67	0.3335	0.3226	0.2536	0.4065	0.148	0.0595	0.2803	0.1491
Sig_1vs1_294	62	0.2847	0.3036	0.2744	0.275	0.1384	0.1153	0.2982	0.0861
Sig_1vs1_295	58	0.2765	0.2927	0.2735	0.2625	0.1604	0.149	0.3193	0.1187
Sig_1vs1_296	56	0.2764	0.2926	0.2739	0.2617	0.1573	0.1398	0.3116	0.1232
Sig_1vs1_297	52	0.2768	0.2915	0.2739	0.2643	0.1515	0.1387	0.315	0.1005
Sig_1vsall_001	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_002	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_003	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_004	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_005	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_006	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_007	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_008	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_009	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_010	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_011	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_012	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_013	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_014	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_015	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_016	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_017	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_018	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_019	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317

Sig_1vsall_020	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_021	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_022	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_023	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_024	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_025	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_026	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_027	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_028	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_029	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_030	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_031	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_032	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_033	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_034	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_035	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_036	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_037	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_038	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_039	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_040	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_041	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_042	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_043	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_044	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_045	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_046	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_047	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_048	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317

Sig_1vsall_049	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_050	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_051	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_052	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_053	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_054	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_055	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_056	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_057	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_058	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_059	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_060	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_061	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_062	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_063	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_064	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_065	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_066	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_067	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_068	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_069	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_070	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_071	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_072	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_073	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_074	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_075	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_076	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_077	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181

Sig_1vsall_078	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_079	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_080	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_081	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_082	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_083	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_084	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_085	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_086	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_087	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_088	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_089	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_090	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_091	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_092	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_093	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_094	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_095	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_096	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_097	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_098	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_099	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_100	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_101	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_102	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_103	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_104	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_105	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_106	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181

Sig_1vsall_107	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_108	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_109	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_110	399	0.2057	0.217	0.2106	0.1883	0.2276	0.1153	0.3612	0.3181
Sig_1vsall_111	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_112	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_113	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_114	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_115	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_116	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_117	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_118	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_119	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_120	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_121	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_122	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_123	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_124	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_125	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_126	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_127	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_128	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_129	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_130	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_131	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_132	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_133	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_134	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_135	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232

Sig_1vsall_136	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_137	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_138	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_139	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_140	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_141	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_142	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_143	401	0.2056	0.2174	0.21	0.1882	0.2286	0.1094	0.3641	0.3232
Sig_1vsall_144	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_145	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_146	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_147	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_148	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_149	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_150	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_151	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_152	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_153	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_154	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_155	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_156	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_157	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_158	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_159	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_160	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_161	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_162	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_163	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_164	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596

Sig_1vsall_165	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_166	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_167	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_168	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_169	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_170	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_171	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_172	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_173	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_174	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_175	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_176	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_177	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_178	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_179	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_180	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_181	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_182	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_183	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_184	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_185	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_186	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_187	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_188	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_189	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_190	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_191	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_192	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_193	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596

Sig_1vsall_194	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_195	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_196	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_197	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_198	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_199	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_200	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_201	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_202	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_203	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_204	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_205	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_206	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_207	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_208	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_209	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_210	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_211	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_212	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_213	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_214	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_215	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_216	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_217	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_218	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_219	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_220	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_221	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_222	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596

Sig_1vsall_223	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_224	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_225	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_226	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_227	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_228	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_229	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_230	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_231	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_232	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_233	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_234	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_235	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_236	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_237	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_238	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_239	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_240	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_241	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_242	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_243	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_244	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_245	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_246	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_247	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_248	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_249	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_250	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_251	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317

Sig_1vsall_252	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_253	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_254	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_255	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_256	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_257	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_258	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_259	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_260	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_261	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_262	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_263	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_264	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_265	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_266	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_267	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_268	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_269	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_270	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_271	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_272	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_273	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_274	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_275	402	0.2056	0.2171	0.2105	0.1881	0.2293	0.1111	0.3635	0.325
Sig_1vsall_276	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_277	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_278	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_279	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_280	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317

Sig_1vsall_281	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_282	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_283	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_284	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_285	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_286	396	0.2052	0.2164	0.2109	0.1872	0.229	0.1161	0.3584	0.317
Sig_1vsall_287	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_288	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_289	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_290	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_291	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_292	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_293	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_294	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_295	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_296	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596
Sig_1vsall_297	305	0.205	0.2249	0.2195	0.1652	0.2339	0.1114	0.3277	0.3596

TABLE B.2: Performances obtained by CD8quant in the deconvolution of the Pulko simulated mixtures using different signature matrix derived from the Bonnal data set. The first column offers each signature candidate which has been fed the algorithm, while the other columns present the signature genes for each candidate and the RMSE and Pearson's correlation output between the real fractions and the estimated ones (both all together and single cell type).

Bibliography

[1] O. K. Afanasiev et al. “Vascular E-selectin expression correlates with CD8 lymphocyte infiltration and improved outcome in Merkel cell carcinoma”. In: *Journal of Investigative Dermatology* 133.8 (2013), pp. 2065–2073.

[2] M. Angelova et al. “Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy”. In: *Genome biology* 16.1 (2015), p. 64.

[3] S. M. Ansell et al. “PD-1 blockade with nivolumab in relapsed or refractory Hodgkin’s lymphoma”. In: *New England Journal of Medicine* 372.4 (2015), pp. 311–319.

[4] C. W. Brennan et al. “The somatic genomic landscape of glioblastoma”. In: *Cell* 155.2 (2013), pp. 462–477.

[5] A. E. Carpenter et al. “CellProfiler: image analysis software for identifying and quantifying cell phenotypes”. In: *Genome biology* 7.10 (2006), R100.

[6] P. Charoentong et al. “Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade”. In: *Cell reports* 18.1 (2017), pp. 248–262.

[7] C. Conrad et al. “Plasmacytoid dendritic cells promote immunosuppression in ovarian cancer via ICOS costimulation of Foxp3+ T-regulatory cells”. In: *Cancer research* 72.20 (2012), pp. 5240–5249.

[8] A. Dander et al. “Personalized Oncology Suite: integrating next-generation sequencing data and whole-slide bioimages”. In: *BMC bioinformatics* 15.1 (2014), p. 306.

[9] S. S. Dave et al. “Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells”. In: *New England Journal of Medicine* 351.21 (2004), pp. 2159–2169.

[10] M. De Simone et al. “Transcriptional landscape of human tissue lymphocytes unveils uniqueness of tumor-infiltrating T regulatory cells”. In: *Immunity* 45.5 (2016), pp. 1135–1147.

[11] M.-C. Dieu-Nosjean et al. “Tertiary lymphoid structures in cancer and beyond”. In: *Trends in immunology* 35.11 (2014), pp. 571–580.

[12] A. Esteve-Codina et al. “A Comparison of RNA-Seq Results from Paired Formalin-Fixed Paraffin-Embedded and Fresh-Frozen Glioblastoma Tissue Samples”. In: *PloS one* 12.1 (2017), e0170632.

[13] D. L. Farber, N. A. Yudanin, and N. P. Restifo. “Human memory T cells: generation, compartmentalization and homeostasis”. In: *Nature reviews Immunology* 14.1 (2014), pp. 24–35.

[14] M.-T. Fernández-Figueras et al. “Expression profiles associated with aggressive behavior in Merkel cell carcinoma”. In: *Modern Pathology* 20.1 (2007), pp. 90–101.

[15] F. Finotello and B. Di Camillo. “Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis”. In: *Briefings in functional genomics* 14.2 (2015), pp. 130–142.

[16] W. H. Fridman et al. “The immune contexture in human tumours: impact on clinical outcome”. In: *Nature Reviews Cancer* 12.4 (2012), pp. 298–306.

[17] J. Galon et al. “Cancer classification using the Immunoscore: a worldwide task force”. In: *Journal of translational medicine* 10.1 (2012), p. 205.

[18] L. Gattinoni et al. “A human memory T cell subset with stem cell-like properties”. In: *Nature medicine* 17.10 (2011), pp. 1290–1297.

[19] V. Golubovskaya and L. Wu. “Different subsets of T cells, memory, effector functions, and CAR-T immunotherapy”. In: *Cancers* 8.3 (2016), p. 36.

[20] G. R. Gouveia et al. “Comparison of two methods of RNA extraction from formalin-fixed paraffin-embedded tissue specimens”. In: *BioMed research international* 2014 (2014).

[21] H. Hackl et al. “Computational genomics tools for dissecting tumour-immune cell interactions”. In: *Nature Reviews Genetics* (2016).

[22] S. Hadrup, M. Donia, and P. thor Straten. “Effector CD4 and CD8 T cells and their role in the tumor microenvironment”. In: *Cancer Microenvironment* 6.2 (2013), pp. 123–133.

[23] J. Hedegaard et al. “Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue”. In: *PloS one* 9.5 (2014), e98187.

[24] G. Hennig et al. “Automated extraction of DNA and RNA from a single formalin-fixed paraffin-embedded tissue section for analysis of both single-nucleotide polymorphisms and mRNA expression”. In: *Clinical chemistry* 56.12 (2010), pp. 1845–1853.

[25] L. Kamentsky et al. “Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software”. In: *Bioinformatics* 27.8 (2011), pp. 1179–1180.

[26] H Li. *seqtk Toolkit for processing sequences in FASTA/Q formats*. 2012.

[27] M. Linkert et al. “Metadata matters: access to image data in the real world”. In: *The Journal of cell biology* 189.5 (2010), pp. 777–782.

[28] M. I. Love, W. Huber, and S. Anders. “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2”. In: *Genome biology* 15.12 (2014), p. 550.

[29] F. Lüder Ripoli et al. “A comparison of fresh frozen vs. formalin-fixed, paraffin-embedded specimens of canine mammary tumors via branched-DNA assay”. In: *International journal of molecular sciences* 17.5 (2016), p. 724.

[30] Y. D. Mahnke et al. “The who’s who of T-cell differentiation: Human memory T-cell subsets”. In: *European journal of immunology* 43.11 (2013), pp. 2797–2809.

[31] A. M. Newman et al. “Robust enumeration of cell subsets from tissue expression profiles”. In: *Nature methods* 12.5 (2015), pp. 453–457.

[32] J. J. Obar and L. Lefrançois. “Memory CD8+ T cell differentiation”. In: *Annals of the New York Academy of Sciences* 1183.1 (2010), pp. 251–266.

[33] B. A. Olenchock et al. “Disruption of diacylglycerol metabolism impairs the induction of T cell anergy”. In: *Nature immunology* 7.11 (2006), pp. 1174–1181.

[34] F. Ozsolak and P. M. Milos. “RNA sequencing: advances, challenges and opportunities”. In: *Nature reviews genetics* 12.2 (2010), nrg2934.

[35] D. M. Pardoll. “Inducing autoimmune disease to treat cancer”. In: *Proceedings of the National Academy of Sciences* 96.10 (1999), pp. 5340–5342.

[36] D. M. Pardoll. “The blockade of immune checkpoints in cancer immunotherapy”. In: *Nature Reviews Cancer* 12.4 (2012), pp. 252–264.

[37] J Parker et al. “A supervised risk predictor of breast cancer based on biological subtypes”. In: *Journal of Clinical Oncology* 26.15_suppl (2008), pp. 11008–11008.

[38] J. S. Parker et al. “Supervised risk predictor of breast cancer based on intrinsic subtypes”. In: *Journal of clinical oncology* 27.8 (2009), pp. 1160–1167.

[39] S. K. Penland et al. “RNA expression analysis of formalin-fixed paraffin-embedded tumors”. In: *Laboratory investigation* 87.4 (2007), pp. 383–391.

[40] P. A. Prieto et al. “CTLA-4 blockade with ipilimumab: long-term follow-up of 177 patients with metastatic melanoma”. In: *Clinical Cancer Research* (2012).

[41] V. Pulko et al. “Human memory T cells with a naive phenotype accumulate with aging and respond to persistent viruses”. In: *Nature immunology* 17.8 (2016), pp. 966–975.

[42] M. D. Robinson, D. J. McCarthy, and G. K. Smyth. “edgeR: a Bioconductor package for differential expression analysis of digital gene expression data”. In: *Bioinformatics* 26.1 (2010), pp. 139–140.

[43] C. Sautès-Fridman et al. “Tertiary lymphoid structures in cancers: prognostic value, regulation, and manipulation for therapeutic intervention”. In: *Frontiers in immunology* 7 (2016).

[44] M. S. Scicchitano et al. “Preliminary comparison of quantity, quality, and microarray performance of RNA extracted from formalin-fixed, paraffin-embedded, and unfixed frozen tissue samples”. In: *Journal of Histochemistry & Cytochemistry* 54.11 (2006), pp. 1229–1237.

[45] A. P. Scorsato and J. E. Q. Telles. “Fatores que interferem na qualidade do DNA extraído de amostras biológicas armazenadas em blocos de parafina”. In: *Jornal Brasileiro de Patologia e Medicina Laboratorial* 47.5 (2011), pp. 541–548.

[46] B. Sengüven et al. “Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues”. In: *International journal of medical sciences* 11.5 (2014), p. 494.

[47] C. Sommer et al. “Ilastik: Interactive learning and segmentation toolkit”. In: *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE. 2011, pp. 230–233.

[48] E. Tappeiner et al. “TIminer: NGS data mining pipeline for cancer immunology and immunotherapy”. In: *Bioinformatics* (2017).

[49] T. Tsujikawa et al. “Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis”. In: *Cell Reports* 19.1 (2017), pp. 203–217.

[50] S. Von Ahlfen et al. “Determinants of RNA quality from FFPE samples”. In: *PloS one* 2.12 (2007), e1261.

[51] Z. Wang, M. Gerstein, and M. Snyder. “RNA-Seq: a revolutionary tool for transcriptomics”. In: *Nature reviews genetics* 10.1 (2009), pp. 57–63.