A Systems Immunology Approach
to Graft-Versus-Host Disease

Pedro Miguel de Ascensão Santos e Sousa

Thesis submitted to UCL in fulfilment of the requirements
for the degree of Doctor of Philosophy

August 2016

Department of Haematology
UCL Cancer Institute
Royal Free Campus

Supervisors: Prof Ronjon Chakraverty
Prof Stephen Mackinnon
Dr Clare Bennett
Declaration of originality

I, Pedro Miguel de Ascensão Santos e Sousa confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

It is not known why only certain tissues are prone to graft-versus-host disease (GVHD) injury following allogeneic hematopoietic stem cell transplantation despite widespread antigen expression. Although it is known that T cell effector pathways can have distinct effects upon individual GVHD organs, there has been no unbiased or systems-wide approach to defining the mechanisms underlying tissue-specific pathology. This thesis reports the results from a systems immunology approach to address the hypothesis that GVHD target tissues exert dominant, idiosyncratic roles in regulating effector T cell functions.

To test this concept, gene expression profiles of effector CD8$^+$ T cells infiltrating lymphoid and GVHD target organs were compared in two clinically relevant murine GVHD models. Using Weighted Gene Network Correlation Analysis, a dichotomy between the transcriptomes of T cells in peripheral tissues and lymphoid organs was identified. These profiles diverged sharply between the different GVHD target organs, and between individual sub-compartments of single organs, independently of the TCR repertoire and antigen distribution.

In the skin, expression of a broad effector program was determined by the transition of T cells from the dermis to the epidermis, in a process regulated by Langerhans cells (LC). In the absence of LC, T cells were rendered incapable of up-regulating the full panoply of effector genes, showed impaired
differentiation into resident memory cells and failed to induce cutaneous GVHD. By performing localized LC depletion, it was demonstrated that LC regulated T cell effector programs in situ within the epidermis by providing signals to enhance local cytokine production, promote resistance to apoptosis and enhance local survival.

Collectively, these data demonstrate that GVHD is defined by tissue-autonomous regulation of effector T cells; in the skin, this is dictated by interaction with epidermal LC in situ. This work provides a rationale for precision therapies directed at blocking GVHD in individual tissues.
A systems immunology approach to GVHD

Table of Contents

List of Figures 10
List of Tables 17
List of abbreviations 19

I. INTRODUCTION 23

1. Haematopoietic stem-cell transplantation 24
   1.1. Brief historical review 32

2. Graft-versus-host disease 40
   2.1. Classification and clinical features 42
   2.2. Acute GVHD immunopathology 45
       a) Priming of the immune response 45
       b) Activation of donor T cells 48
       c) Alloreactive T cell proliferation and differentiation 51
       d) Migration of effector cells to GVHD target tissues 52
       e) Target tissue destruction by T_{eff} cells 54

3. Pre-clinical models of GVHD 57
   3.1. Murine models 60

4. Unanswered questions in GVHD 67

5. Project rationale 69

6. Hypothesis 72

7. Aims and objectives 72
II. MATERIALS & METHODS

1. Animal work
   1.1. Mice
   1.2. Haematopoietic stem cell transplantation
   1.3. In vitro T cell activation
   1.4. Animal welfare
   1.5. Treatments
       a) Antibiotic prophylaxis
       b) Lymph node lymphocyte egress blockage
       c) Langerhans cell depletion
       d) Notch signalling blockage

2. Tissue and organ harvest
   2.1. Isolation of murine immune cells
       a) Lymph nodes and spleens
       b) Bone marrow
       c) Blood
       d) Liver
       e) Small intestine
       f) Skin

3. Cell counting

4. Immunolabelling

5. Flow cytometry

6. Sample preparation for gene expression
7. **Microarray analysis**
   7.1. Data acquisition and processing
   7.2. Samples relationship visualisation
   7.3. Differential gene expression
   7.4. Gene set enrichment analysis (GSEA)
   7.5. Weighted gene co-expression network analysis (WGCNA)
   7.6. Gene ontology (GO) overrepresentation analysis

8. **RNA-seq analysis**
   8.1. Library preprocessing
   8.2. Calculation of read-level metrics
   8.3. Single-sample GSEA

9. **Statistical analysis**

III. **RESULTS**

1. **Transcription profiles of donor-derived T\textsubscript{eff} cells are different between SLO and GVHD target organs**

   1.1. Brief characterization of the minor histocompatibility antigen mismatched models
   a) B6\textrightarrow{}129 model
   b) MataHari T cell model

   1.2. Purification and microarray analysis of donor derived CD8\textsuperscript{+} T cells

   1.3. Exploratory data analysis
   a) B6\textrightarrow{}129 model
   b) MataHari T cell model
1.4. SLO - GVHD target organ dichotomy reflects disparities in differentiation status of donor T\textsubscript{eff} cells 121

1.5. Transcriptome diversity is neither due to early trafficking of effectors to the peripheral tissues nor to differential homing imprinting of donor-derived T cells 128

1.6. Discussion 134

2. \textbf{T\textsubscript{eff} cell transcription profiles show compartmental specificity in acute GVHD} 138

2.1. Identification of gene clusters based on correlation and co-expression across tissues using WGCNA 139

2.2. WGCNA defined modules are preserved in an independent murine model of GVHD 147

2.3. Organ specific modules are conserved across species 149

2.4. Module annotation reveals biologically meaningful clusters of genes 156

2.5. Identification of highly connected hub genes 158

2.6. Discussion 163

3. \textbf{Donor CD8\textsuperscript{+} T\textsubscript{eff} pathogenicity in the skin is regulated \textit{in situ} by epidermal Langerhans cells} 167

3.1. LC depletion reduces the differences between dermal and epidermal donor CD8\textsuperscript{+} T\textsubscript{eff} cells transcriptional profiles 173

3.2. In the absence of LC, epidermal donor T cells fail to fully up-regulate the effector cell programs 175

3.3. Donor CD8\textsuperscript{+} T cells expression of the epidermal specific transcriptional signature is regulated by host LC 179
3.4. LC dictate donor T cell accumulation, survival, effector function and T\textsubscript{RM} differentiation potential in the epidermis 181

3.5. Responsiveness to INF-\(\gamma\) by LC and Notch pathway signalling are central in the host LC – donor T cell crosstalk 188

3.6. Discussion 198

IV. DISCUSSION 203

1.1 Systems immunology – a comprehensive approach to acute GVHD study 204

1.2 A modular conception of the cellular programs underpinning GVHD development 208

1.3 A “two-hit” model of cutaneous acute GVHD development 211

1.4 Study limitations & Future work 215

V. CONCLUSION 221

VI. ACKNOWLEDGEMENTS 225

VII. REFERENCES 229

VIII. APPENDIX 263

List of Supplementary Figures 264

List of Supplementary Tables 268
List of Figures

Figure 1. Partial spectrum of conditioning regimens of various intensities and their impact on toxicity, and the dependence of transplant success upon GVT effects.

Figure 2. Timeline showing the main milestones in haematopoietic stem cell transplantation, 1957–2016.

Figure 3. Causes of death after HLA-matched sibling and unrelated donor HSCT.

Figure 4. Progression of events in aGvHD development.

Figure 5. Critical interactions for induction of GVHD.

Figure 6. Antigen presentation in GVHD in MHC-matched allogeneic stem-cell transplantation.

Figure 7. Genomic map of the mouse MHC complex.

Figure 8. B6→129 model.

Figure 9. MataHari T cell model.

Figure 10. Adapted MataHari T cell→Langerin-DTR model.

Figure 11. Characterisation of the B6→129 model.
**Figure 12.** Donor CD8$^+$ T cell infiltration of the secondary lymphoid organs and GVHD target organs at day 6 post-transplant.

**Figure 13.** Frequency of H60-specific from tissue infiltrating donor CD8$^+$ T cells.

**Figure 14.** Kinetics of MataHari T cell infiltration of the GVHD target organs.

**Figure 15.** Characterisation of the MataHari T cell model.

**Figure 16.** Donor CD8$^+$ (MataHari) T cell infiltration of the secondary lymphoid organs and GVHD target organs at day 7 post-transplant.

**Figure 17.** Strategy for high-purity FACS sorting of donor CD8$^+$ T cells.

**Figure 18.** Staining strategy to isolate cells with equivalent number of divisions and homing imprinting.

**Figure 19.** Sort quality check and evaluation of sample purity.

**Figure 20.** Principal components analysis of the B6→129 data set.

**Figure 21.** Hierarchical clustering of the B6→129 data set, validated through non-parametric bootstrap methodology.

**Figure 22.** Multidimensional scaling analysis of the alloBMT samples from the B6→129 data set.

**Figure 23.** Hierarchical clustering of the MataHari T cell data set, validated through non-parametric bootstrap methodology.

**Figure 24.** Multidimensional scaling analysis of the alloBMT samples from the MataHari T cell data set.
**Figure 25.** Transcriptional profiles are independent of the TCR repertoire and do not reflect antigen distribution.

**Figure 26.** Transcriptional differences between SLO and GVHD target organs are biologically meaningful.

**Figure 27.** Venn diagram representing the overlap of enriched gene sets from the KEGG PATHWAY collection in the GVHD target organs vs the SLO, found in the B6→129 and the MataHari T cell models.

**Figure 28.** Heat map of selected genes commonly used to characterise T cell differentiation, including: transcription factors, chemokine/cytokine receptors, signalling molecules, integrins, activation markers, coregulatory molecules, and effector function molecules.

**Figure 29.** Comparison of the differentiation status of donor CD8$^+$ T cells across the tissue. Single-sample GSEA was used to determine the enrichment for gene sets associated with Tc1/Tc17 polarisation, pathogenicity, and cytokine or effector function.

**Figure 30.** Experimental design to evaluate the effect of SLO lymphocyte egress blockage on the transcriptional profile of donor CD8$^+$ T$_{eff}$ cells in the lymph nodes.

**Figure 31.** Treatment with FTY720 blocks lymphocyte egress from the SLO.

**Figure 32.** Differences in gene expression between T$_{eff}$ cells are not due to early egress of optimally activated T$_{eff}$ cells.
Figure 33. Diversity of transcription profiles is not explained by differences in homing imprinting of donor T cells.

Figure 34. Selection of the power $\beta$ for gene correlation network generation.

Figure 35. Dendrogram obtained by hierarchical clustering of genes based on topological overlap, showing the assigned modules colour-coded.

Figure 36. Visualization of the eigengene network representing the relationships among the modules.

Figure 37. Heat map representation of the module-trait relationships.

Figure 38. Eigengene network generated using the edge-weighted spring embedded layout algorithm.

Figure 39. Module preservation study.

Figure 40. Validation of the murine modules on a data set from rhesus macaque.

Figure 41. Validation of the murine modules on a data set from human samples.

Figure 42. GSEA revealed statistically significant enrichment for module 24 in the samples from patients without GVHD.

Figure 43. GVHD target organ-associated modules accurately distinguish between samples isolated from the epidermis and samples isolated from the blood of transplanted patients with cutaneous acute GVHD.
**Figure 44.** Immunofluorescence staining of the skin of alloBMT recipients with cutaneous acute GVHD (MataHari T cell model).

**Figure 45.** Host-derived LC replacement kinetics after alloBMT.

**Figure 46.** Effect of systemic host LC depletion on cutaneous acute GVHD.

**Figure 47.** Effect of localised host LC depletion on donor CD8$^{+}$ T cell infiltration of the epidermis.

**Figure 48.** Correlation matrix comparing the transcription profiles of donor CD8$^{+}$ T_{eff} cells isolated from the lymph nodes, dermis and epidermis of LC depleted and LC replete alloBMT recipients, at D+7 post-transplant.

**Figure 49.** Heat map showing the enrichment for the 91 gene sets significantly over- or under-represented in the dermis vs the lymph nodes and in the epidermis vs the dermis, with or without LC.

**Figure 50.** Classification of the gene sets that were differentially enriched in the epidermis in the presence of LC according to KEGG PATHWAY mapping.

**Figure 51.** Heat map of the 443 genes that form module 28, comparing the transcriptional profiles of the lymph nodes, dermis and epidermis, in the presence or absence of LC.

**Figure 52.** Module 28 top 100 gene network, where the circles represent the genes and the lines represent the connectivity between the genes.
Figure 53. Experimental design to evaluate the effect of LC depletion on proliferation, effector function and survival of skin infiltrating donor CD8$^+$ T cells.

Figure 54. Assessment of the proliferation, survival and effector function of donor CD8$^+$ T cells by FACS staining.

Figure 55. Summary data referring to the study of the effect of LC depletion upon proliferation, survival and effector function of donor CD8$^+$ T$_{eff}$ cells.

Figure 56. Experimental design to evaluate the effect of LC depletion donor CD8$^+$ T cells capacity to differentiate into epidermal T$_{RM}$ cells.

Figure 57. Kinetics of donor derived CD8$^+$ epidermal T$_{RM}$ cells in the presence or absence of LC.

Figure 58. Experimental design to evaluate the requirement of LC responsiveness to IFN-γ for donor CD8$^+$ T cell accumulation in the epidermis.

Figure 59. Comparison of alloBMT driven changes in LC expression of MHC class I H-2Db, CD40, CD70 and CD86.

Figure 60. Study of the survival and effector function of donor CD8$^+$ T cells isolated from alloBMT recipients incapable of detecting IFN-γ signalling (Ifngr1$^{-/-}$) vs WT.

Figure 61. Summary data refering to the study of the effect of LC responsiveness to IFN-γ signalling upon accumulation, survival and effector function of donor CD8$^+$ T$_{eff}$ cells.
Figure 62. Experimental design to evaluate the effect of Notch signalling blockade on donor CD8$^+$ T cell effector function in the epidermis.

Figure 63. Study of the effect of Notch signalling blockade on donor CD8$^+$ T cell effector function in the spleen, dermis and epidermis.
List of Tables

Table 1. Comparison of the characteristics of the 3 different sources of stem cells.

Table 2. Main clinical indications for HSCT in adult patients.

Table 3. Distinguishing acute and chronic GVHD.

Table 4. Organ involvement in acute GVHD.

Table 5. Organ involvement in chronic GVHD.

Table 6. Advantages and disadvantages of various preclinical model systems.

Table 7. H-2 haplotypes and alloantigens of some of the most commonly used laboratory mouse strains.

Table 8. Commonly used mouse models of HSCT.

Table 9. Assessment of clinical GVHD severity in transplanted subjects.

Table 10. Antibodies used for cell surface immunolabelling.

Table 11. Antibodies used for intracellular immunolabelling.

Table 12. Top 10 pathways enriched in GVHD target organs vs SLO common to both data sets (MSigDB C2:CP:KEGG collection).
Table 13. WGCNA modules.

Table 14. Module specific driver genes summary.

Table 15. Top 10 type I and type II IFN responsive genes upregulated by LC after alloBMT.
List of abbreviations

aGVHD    acute GVHD
allo     allogeneic
auto     autologous
APC      antigen presenting cells
AraC     cytosine arabinoside
ATG      anti-thymocyte globulin
BM       bone marrow
BMT      bone marrow transplantation
BSA      body surface area
BU       busulphan
CB       cord blood
cGVHD    chronic GVHD
CMV      cytomegalovirus
ConA     concanavalin A
CSA      cyclosporine A
CTL      cytotoxic T cells
CY       cyclophosphamide
DAMP     damage-associated molecular patterns
DE       differential expression
DETC     dendritic epidermal γδ T cells
DLI      donor lymphocyte infusion
DOT Vδ1+ T cells
DT diphtheria toxin
EdU 5-ethyl-2'-deoxyuridine
FACS Fluorescence-activated cell sorting
FLU fludarabine
FTY720 fingolimod
GO Gene Ontology
GSEA gene set enrichment analysis
GVHD graft-versus-host disease
GVL graft-versus-leukaemia
GVT graft-versus-tumour
HLA human leukocyte antigen
Hp haploidentical donor
HSCT haematopoietic stem cell transplantation
IEL intraepithelial lymphocytes
ILC3 group 3 innate lymphoid cells
KEGG Kyoto Encyclopedia of Genes and Genomes
LC Langerhans cells
LFA lymphocyte function-associated antigen
LN lymph nodes
LP lamina propria
MDS multidimensional scaling
MHC major histocompatibility complex
miHA minor histocompatibility antigen
MLN mesenteric lymph nodes
MSigDB Molecular Signatures Database
<table>
<thead>
<tr>
<th><strong>Abbreviation</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>MUD</td>
<td>matched unrelated donor</td>
</tr>
<tr>
<td>NES</td>
<td>normalised enrichment score</td>
</tr>
<tr>
<td>PBSC</td>
<td>peripheral blood stem cells</td>
</tr>
<tr>
<td>PCA</td>
<td>principal components analysis</td>
</tr>
<tr>
<td>PLN</td>
<td>peripheral lymph nodes</td>
</tr>
<tr>
<td>RIC</td>
<td>reduced intensity conditioning</td>
</tr>
<tr>
<td>S1P</td>
<td>sphingosine 1-phosphate</td>
</tr>
<tr>
<td>SIB</td>
<td>sibling</td>
</tr>
<tr>
<td>SLO</td>
<td>secondary lymphoid organs</td>
</tr>
<tr>
<td>SPF</td>
<td>specific pathogen-free</td>
</tr>
<tr>
<td>ssGSEA</td>
<td>single-sample gene set enrichment analysis</td>
</tr>
<tr>
<td>syn</td>
<td>syngeneic</td>
</tr>
<tr>
<td>TBI</td>
<td>total body irradiation</td>
</tr>
<tr>
<td>TCR</td>
<td>T cell receptor</td>
</tr>
<tr>
<td>T\textsubscript{eff}</td>
<td>effector T cells</td>
</tr>
<tr>
<td>TOM</td>
<td>topological overlap matrix</td>
</tr>
<tr>
<td>T\textsubscript{RM}</td>
<td>tissue resident memory T cells</td>
</tr>
<tr>
<td>WGCNA</td>
<td>weighted gene co-expression network analysis</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
</tbody>
</table>
I. Introduction
1. Haematopoietic stem-cell transplantation

Hematopoietic cell transplantation (HSCT) includes a variety of procedures in which a patient receives an infusion of hematopoietic progenitor cells following a preparative regimen of chemo- and/or radiotherapy, and can be broadly classified into 3 subtypes based upon the hematopoietic cell donor:

- **Autologous HSCT** (autoHSCT) refers to the reinfusion of the patient's hematopoietic progenitor cells, collected from the patient prior to the administration of high-dose chemotherapy targeted at an underlying malignancy, and is performed to rescue the patient from long-lasting and profound pancytopenia;

- **Syngeneic HSCT** (synHSCT) refers to the use of hematopoietic progenitor cells obtained from an identical twin, offering the advantage of providing a graft that is free of tumor cells;

- **Allogeneic HSCT** (alloHSCT) refers to the use of hematopoietic progenitor cells collected from a related or unrelated donor, who may display variable degrees of immunologic disparity with the recipient, thus adding the potentially beneficial of an immunological reaction against the tumour.

Allogeneic haematopoietic stem cell transplantation (alloHSCT) has revolutionised the treatment of many haematological and non-haematological
malignancies, not only by permitting the usage of more intensive chemo- and radiotherapy regimens, but also, and most importantly, by introducing adoptive cellular immunotherapy into the clinical practice.\(^1\) Initially regarded as a means of replacing an irreversibly damaged haematopoietic compartment and overcoming the radio- and chemotherapy dose limitation imposed by bone marrow toxicity, it was soon recognised that this technique had a second powerful therapeutic effect, the graft-versus-tumour (GVT) effect, in which the host’s tumour cells are recognized as foreign and killed by the lymphocytes contained in the donor graft.\(^2\) However, this alloreactive immune response is not restricted to the tumour sites and, while useful in combating any residual disease, when it affects the host’s healthy tissues it can cause serious life-threatening damage, the graft-versus-host disease (GVHD).

The success of alloHSCT is thus based on three basic principles: (1) conditioning of the patient in order to “create space” for donor stem cells to access the host stem cell niches and for engraftment to occur; (2) long-term disease control, either through the eradication of the malignant cells or by reconstituting the damaged environment; and (3) modulation of the immune system, to balance the beneficial GVT effect while avoiding the development of GVHD.\(^3\)

The conditioning of the recipient is a crucial element in HSCT, thus, the selection of the optimal preparative regimen for any given patient is dependent upon disease-related factors (i.e. diagnosis and remission status) and patient-related factors (e.g. age, donor availability, comorbid conditions).\(^4\) Traditionally the conditioning of the patients was achieved by delivering
maximally tolerated doses of multiple chemotherapeutic agents with
non-overlapping toxicities, with or without radiation. However, even though this
high-dose myeloablative approach was very effective against most leukaemias
and lymphomas, and was able to penetrate the sanctuary sites, the associated
toxicity limited its use to young patients with a good performance status. The
increasing evidence that the eradication of malignant disorders could be
accomplished through the GVT effect led to the development of reduced-
intensity and nonmyeloablative regimens, making allogeneic HSCT a viable
therapeutic option for frail and older and patients, unable to tolerate high-dose
conditioning (Figure 1).

Initially, the main graft source was the bone marrow, harvested from the
posterior iliac crests under general anaesthesia (hence the designation of bone
marrow transplantation, BMT). Further characterisation of the haematopoietic
stem cell compartment prompted their identification in the peripheral blood,
and the development of new techniques to enrich for and harvest them.
Subsequently, mobilized peripheral blood stem cells (PBSC) have become one
of the main stem cell source for HSCT, and although no standardised
indications have been established to prefer PBSC over bone marrow, operating
room unavailability, lack of personnel to perform bone marrow aspiration,
contraindication to general anaesthesia and choice of the donor, are some of
the factors that may determine this trend. The discovery of cord blood as a
rich source of stem cells, in 1978, and its successful use in allogeneic
HSCT, provided a third source of stem cell
Figure 1. Partial spectrum of conditioning regimens of various intensities and their impact on toxicity, and the dependence of transplant success upon GVT effects. AraC, cytosine arabinoside; ATG, anti-thymocyte globulin; BU, busulfan; CY, cyclophosphamide; FLU, fludarabine; $^{131}$I, radioiodine; TBI, total body irradiation; *High-dose TBI (800-1320 cGy); §Low-dose TBI (200-400 cGy).

devoid of risks to the donors, with reduced likelihood of transmitting infections, particularly CMV, which can be fully tested and HLA typed, cryopreserved, and made available for immediate use (Table 2).⁶

Moreover, manipulation of the cellular components of the grafts via CD34⁺ and T cell depletion has expanded the field even further, extending the option of HSCT to patients without an HLA-matched related or unrelated donor.¹⁴ The success of this approach has fuelled the search for the optimal graft composition, providing rapid engraftment kinetics and enhanced immune reconstitution, with high antitumoural activity but low GVHD-inducing potential. The leading strategies already in clinical trials include infusion of regulatory T cells,¹⁵ graft depletion of αβ⁺ T and B cells,¹⁶,¹⁷ NK cell adoptive immunotherapy,¹⁸ and selective allo-depletion of T cells.¹⁹

HSCT has thus evolved, in just a few decades, from an experimental procedure for patients with bone marrow failure or incurable leukaemia to a standard of care therapy for an expanding range of patients with congenital or acquired disorders of the hematopoietic system or radio-, chemo- or immune-sensitive malignancies (Table 2).²⁰
Table 1. Comparison of the characteristics of the 3 different sources of stem cells.

<table>
<thead>
<tr>
<th></th>
<th>BM</th>
<th>PBSC</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical time frame from initiation of search to transplantation</td>
<td>3-6 months</td>
<td>3-6 months</td>
<td>2-4 weeks</td>
</tr>
<tr>
<td>Adverse effects for donor</td>
<td>wound infection, bleeding, general anesthesia, etc.</td>
<td>bleeding/thrombosis, infection, hypotension, electrolyte disturbance, etc.</td>
<td>None</td>
</tr>
<tr>
<td>Minimal cell dose for transplant</td>
<td>Total nucleated cell: 2x10^6/kg</td>
<td>Total CD34+ cell: 2x10^6/kg</td>
<td>Total nucleated cell: 2.5x10^7/kg</td>
</tr>
<tr>
<td>Possibility to give additional stem cell dose</td>
<td>Possible</td>
<td>Possible</td>
<td>Impossible</td>
</tr>
<tr>
<td>HLA matching requirement</td>
<td>More stringent (7/8 matched)</td>
<td>More stringent (7/8 matched)</td>
<td>Less stringent (4-6/6 matched)</td>
</tr>
<tr>
<td>Neutrophil engraftment</td>
<td>About 3 weeks</td>
<td>About 2 weeks</td>
<td>About 4 weeks</td>
</tr>
<tr>
<td>Immune reconstitution</td>
<td>Faster</td>
<td>Faster</td>
<td>Slower</td>
</tr>
<tr>
<td>GVHD risk</td>
<td>Medium</td>
<td>Highest</td>
<td>Lowest</td>
</tr>
<tr>
<td>Post-HSCT infection risk</td>
<td>Lower</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Latent virus transmission risk</td>
<td>Higher</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>CMV transmission risk</td>
<td>Higher</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>Relapse risk</td>
<td>Higher</td>
<td>Lower</td>
<td>Higher</td>
</tr>
</tbody>
</table>

BM, bone marrow; CB, cord blood; CMV, cytomegalovirus; GVHD, graft-versus-host disease; HLA: human leukocyte antigen; PBSC, peripheral blood stem cell.
<table>
<thead>
<tr>
<th>Disease</th>
<th>Disease status</th>
<th>SIB*</th>
<th>MUD*</th>
<th>CB/Hp*</th>
<th>Auto*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myeloid leukaemia</td>
<td>CR1 (low risk)</td>
<td>CO (II)</td>
<td>D (II)</td>
<td>GNR (II)</td>
<td>CO (I)</td>
</tr>
<tr>
<td></td>
<td>CR1 (intermediate)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>D (II)</td>
<td>S (I)</td>
</tr>
<tr>
<td></td>
<td>CR1 (high risk)</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>CO (I)</td>
</tr>
<tr>
<td></td>
<td>CR2</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>CO (I)</td>
</tr>
<tr>
<td></td>
<td>CR3, incipient relapse</td>
<td>S (III)</td>
<td>CO (III)</td>
<td>D (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>M3 Molecular persistence</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>M3 Molecular CR2</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>S (II)</td>
</tr>
<tr>
<td></td>
<td>Relapse or refractory</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>D (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Acute lymphoblastic leukaemia</td>
<td>Ph(+) CR1 (standard risk)</td>
<td>D (II)</td>
<td>GNR (II)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td></td>
<td>Ph(+) CR1 (high risk)</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>Ph(-), CR1</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>CO (I)</td>
</tr>
<tr>
<td></td>
<td>CR2, incipient relapse</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (II)</td>
</tr>
<tr>
<td></td>
<td>Relapse or refractory</td>
<td>CO (II)</td>
<td>D (II)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Chronic myeloid leukaemia</td>
<td>1st CP, failing TKI</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>Accelerated phase or&gt;1st CP</td>
<td>S (II)</td>
<td>S (II)</td>
<td>D (II)</td>
<td>S (III)</td>
</tr>
<tr>
<td></td>
<td>Blast crisis</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Myelofibrosis</td>
<td>Primary or secondary with an int. or high DIPSS score</td>
<td>S (II)</td>
<td>S (II)</td>
<td>S (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Myelodysplastic syndromes</td>
<td>RA, RCMD, RAEB I and II</td>
<td>S (II)</td>
<td>S (II)</td>
<td>S (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>sAML in CR1 or CR2</td>
<td>S (II)</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
</tr>
<tr>
<td></td>
<td>More advanced stages</td>
<td>S (II)</td>
<td>S (II)</td>
<td>S (II)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Chronic lymphocytic leukaemia</td>
<td>Poor risk disease</td>
<td>S (II)</td>
<td>S (II)</td>
<td>D (III)</td>
<td>GNR (I)</td>
</tr>
<tr>
<td>Diffuse large B cell lymphoma</td>
<td>CR1 (intermediate/high IPI at dx)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (I)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse, 2CR2</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>D (III)</td>
<td>S (I)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse after autoHSCT failure</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>Refractory</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>D (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td>Mantle cell lymphoma</td>
<td>CR1</td>
<td>D (III)</td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>S (I)</td>
</tr>
<tr>
<td></td>
<td>CR/PR&gt;1, no prior autoHSCT</td>
<td>CO (III)</td>
<td>CO (III)</td>
<td>D (III)</td>
<td>S (II)</td>
</tr>
<tr>
<td></td>
<td>CR/PR&gt;1, prior autoHSCT</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (III)</td>
<td>GNR (II)</td>
</tr>
<tr>
<td></td>
<td>Refractory</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>D (III)</td>
<td>GNR (II)</td>
</tr>
<tr>
<td>Follicular lymphoma</td>
<td>CR1</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>D (II)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse, 2CR2</td>
<td>CO (III)</td>
<td>CO (III)</td>
<td>GNR (III)</td>
<td>S (II)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse after autoHSCT failure</td>
<td>S (II)</td>
<td>S (II)</td>
<td>D (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>Refractory</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>WALDENSTROM MACROGLOBULINAEMIA</td>
<td>CR1</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>D (II)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse, 2CR2</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>D (II)</td>
</tr>
<tr>
<td></td>
<td>Poor risk disease</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>D (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>T cell lymphoma</td>
<td>CR1</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse, 2CR2</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td></td>
<td>Refractory</td>
<td>CO (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>GNR (II)</td>
</tr>
<tr>
<td>PRIMARY CUTANEOUS T CELL LYMPHOMA</td>
<td>EORTC/ISCL I-II(A) (early)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>EORTC/ISCL II-B-IV (advanced)</td>
<td>CO (III)</td>
<td>CO (III)</td>
<td>D (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Hodgkin’s lymphoma</td>
<td>CR1</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse, no prior autoHSCT</td>
<td>D (III)</td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>S (I)</td>
</tr>
<tr>
<td></td>
<td>Chemosensitive relapse, prior autoHSCT</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (III)</td>
<td>CO (III)</td>
</tr>
<tr>
<td></td>
<td>Refractory</td>
<td>D (II)</td>
<td>D (II)</td>
<td>D (III)</td>
<td>CO (III)</td>
</tr>
<tr>
<td>Multiple myeloma</td>
<td></td>
<td>CO (I)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>S (I)</td>
</tr>
</tbody>
</table>
A systems immunology approach to GVHD

Chapter I

Table 2. (continued)

<table>
<thead>
<tr>
<th>Disease</th>
<th>Disease status</th>
<th>SIB (I)</th>
<th>MUD (I)</th>
<th>CB/Hp (I)</th>
<th>Auto (I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amyloidosis</td>
<td></td>
<td>CO (III)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td>Acquired severe aplastic anaemia</td>
<td>Newly diagnosed</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Relapse/Refractory</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>NA</td>
</tr>
<tr>
<td>Constitutional severe</td>
<td>Fanconi anaemia</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>NA</td>
</tr>
<tr>
<td>aplastic anaemia</td>
<td>Dyskeratosis congenital</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>NA</td>
</tr>
<tr>
<td>Acquired aplastic</td>
<td>Newly diagnosed</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>GNR (III)</td>
<td>NA</td>
</tr>
<tr>
<td>anaemia/PNH</td>
<td>Relapse/Refractory</td>
<td>S (II)</td>
<td>S (II)</td>
<td>CO (II)</td>
<td>NA</td>
</tr>
<tr>
<td>Haemolytic PNH</td>
<td></td>
<td>GNR (II)</td>
<td>GNR (II)</td>
<td>GNR (II)</td>
<td>NA</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>Adjuvant high risk, HER2-negative</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td></td>
<td>Metastatic, chemosensitive</td>
<td>D (II)</td>
<td>D (II)</td>
<td>GNR (III)</td>
<td>D/CO (II)</td>
</tr>
<tr>
<td>Germ cell tumours</td>
<td>Second line, high risk</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
</tr>
<tr>
<td></td>
<td>Primary refractory, second &amp; further</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>S (II)</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>High risk/recurrent</td>
<td>D (II)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (I-II)</td>
</tr>
<tr>
<td>Medulloblastoma</td>
<td>Post-surgery, high risk</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>D/CO (III)</td>
</tr>
<tr>
<td>Small cell lung cancer</td>
<td>Limited</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>D (I-II)</td>
</tr>
<tr>
<td>Soft tissue sarcoma</td>
<td>Metastatic</td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (II)</td>
</tr>
<tr>
<td>Ewing’s sarcoma</td>
<td>Locally advanced/metastatic,</td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (III)</td>
</tr>
<tr>
<td>family of tumours</td>
<td>chemosensitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal cell carcinoma</td>
<td>Metastatic, cytokine-refractory</td>
<td>D (II)</td>
<td>D (II)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td>Advanced</td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>Metastatic</td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>GNR (III)</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td></td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
<tr>
<td>Systemic scleroderma</td>
<td></td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
<tr>
<td>Systemic lupus</td>
<td></td>
<td>D (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
<tr>
<td>erythematous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crohn’s disease</td>
<td></td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td></td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
<tr>
<td>Vasculitis</td>
<td></td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
<tr>
<td>Polymyositis-dermatomyositis</td>
<td></td>
<td>GNR (III)</td>
<td>GNR (III)</td>
<td>CO (II)</td>
<td></td>
</tr>
</tbody>
</table>

* S, standard of care (generally indicated in suitable patients); CO, clinical option (can be carried after careful assessment of risks and benefits); D, developmental (further trials are needed); GNR, generally not recommended. Levels of evidence are displayed in brackets.

Abbreviations: CP, chronic phase; CR1/2/3, first/second/third complete remission; DIPSS, Dynamic International Prognostic Score System; IPI, International prognostic index; M3, promyelocytic leukaemia; RA, refractory anaemia; RAEB, refractory anaemia with excess blasts; RCMD, refractory cytopenia with multilineage dysplasia; sAML, secondary AML; TKI, tyrosine kinase inhibitors.

1.1. Brief historical review

The rational basis for bone marrow transplantation (BMT) stems from pioneer experiments performed in the post-World War II period in an effort to overcome the lethal complications of high-dose radiation exposure.\(^{(1,21)}\) Jacobson et al. demonstrated in mice that survival from otherwise lethal radiation exposure could be achieved with the shielding of the spleen and/or femur by a lead foil,\(^{(22,23)}\) and proposed the existence of “a substance of noncellular nature within the spleen and bone marrow” which protected the subjects from irradiation toxicity (humoral hypothesis). Further work by Lorenz et al. showing that infusion of spleen or bone marrow cells had a similar protective effect in irradiated rodents,\(^{(24)}\) introduced an alternative hypothesis which viewed the spleen and bone marrow cellular constituents as being responsible for the clinical recovery observed (cellular hypothesis). A few years later, several research groups provided definite evidence supporting the latter, documenting in mice that after irradiation and infusion of bone marrow or spleen cells the haematopoietic recovery was dependent on the cells derived from the donor grafts.\(^{(25-29)}\)

Many of the factors known to be responsible for a successful bone marrow engraftment were established in early studies in murine BMT models:\(^{(30)}\) (1) Van Bekkum et al. showed that repopulation of the marrow spaces was effectively achieved if bone marrow cells were given intravenously;\(^{(31)}\) (2) conditioning with cyclophosphamide was sufficient for
allogeneic engraftment;\(^{(32)}\) (3) successfully engrafted allogeneic bone marrow cells could induce an aggressive reaction against the host causing a wasting syndrome (secondary disease), now recognised as being graft-versus-host disease (GVHD);\(^{(33)}\) (4) genetic factors controlled the severity of the immune reaction of donor cells against the host;\(^{(34)}\) (5) GVHD could be prevented or ameliorated through immunosuppression with methotrexate.\(^{(35)}\)

However, transferring the knowledge obtained from experimental animal models to humans was met with many challenges. The results from the initial attempts at translating the procedure into the clinic were disappointing: although in some of the patients treated with BMT aimed at reconstituting the BM following radiotherapy, successful engraftment and leukaemia remission were achieved, none survived for more than a few months,\(^{(36)}\) succumbing to early complications associated with bone marrow aplasia (e.g. bleeding, severe infections) or after developing severe fatal GVHD. These discouraging results tempered the clinicians’ enthusiasm, with some prominent haematologists stating that “… these failures have occurred mainly because the clinical applications were undertaken too soon, most before even the minimum basic knowledge required to bridge the gap between mouse and patient had been obtained”.\(^{(30)}\)

Further research was subsequently conducted in dogs as it was predicted that the wide genetic diversity of out-bred dogs would be a much more suitable animal model for preclinical studies.\(^{(37)}\) Although the existence of murine histocompatibility systems had already been documented in the mid-1950s,\(^{(38)}\) it was only with Epstein et al. in 1968 that it was demonstrated
that leukocyte antigen compatibility between donor and recipient was absolutely required for the success of allogeneic transplantation.\textsuperscript{(39)} Moreover, disparities involving minor histocompatibility antigens (miHA) were found to be equally capable of eliciting GVHD.\textsuperscript{(40)} On the other hand, mutual graft-versus-host tolerance was documented to be established after 3 to 6 months of methotrexate treatment, suggesting that suspension of immunosuppressive therapy would be possible thereafter.\textsuperscript{(41)}

The encouraging results in pre-clinical models, together with the discovery of the Human Leukocyte Antigen (HLA) system by J. Dausset\textsuperscript{(42)} and J.J. Van Rood\textsuperscript{(43)} which permitted selection of HLA identical donors, inspired further trials of allogeneic HSCT between matched human siblings. Additionally, improvements in supportive care (in particular blood component transfusions, antibiotic / antifungal / antiviral therapy, vascular access, and parenteral nutrition) lead to a reduction in pancytopenia associated complications and permitted the development of new high-dose conditioning regimens.\textsuperscript{(44)}

In the late 1960s, Gatti \textit{et al.}, Bach \textit{et al.} and deKoning \textit{et al.} reported the first 3 successful cases of allogeneic marrow transplantation in patients with severe combined immunological deficiency using HLA-matched sibling donors,\textsuperscript{(45-47)} with a long term survival of over 25 years.\textsuperscript{(48)} In the following years, the Seattle Marrow Transplant Team described their results with leukaemia and aplastic anaemia patients transplanted after failure of conventional therapy, showing some long-term disease-free survival, in spite of still having a high transplant-related mortality.\textsuperscript{(49,50)} Due to the low success rate and morbidity
associated with the procedure, at this stage, BMT was restricted to patients with immunoglobulin deficiencies, severe aplastic anaemia, or refractory advanced leukaemia/lymphoma who had HLA-matched sibling donors. This remained the leading practice for over a decade, until 1977 when Thomas et al. showed that patients in poorer conditions at transplantation had a significantly lower survival,\(^{(50)}\) suggesting that higher success rates would be achieved if transplantation occurred earlier in the course of the disease, when tumour burdens are reduced and patients are still in good medical condition.

Consistent with earlier findings in dogs, half of the patients receiving a transplant from an HLA-matched sibling donor and treated with methotrexate alone still developed GVHD. However, by associating cyclosporine A (CSA) or tacrolimus, GVHD management was significantly improved,\(^{(51-53)}\) to such an extent that these combinations of immunosuppressants are widely used to this day. Long-term survival of allogeneic transplanted patients introduced two new concepts into clinical practice: chronic GVHD and the graft-versus-leukaemia (GVL) effect. As it emerged that tumour was controlled both by the conditioning’s cytotoxicity and by the effects of the donor immune system against malignant host cells, a link between GVL and GVHD was recognised.\(^{(54-56)}\)

Gradually, allogeneic BMT began to be viewed as a valid treatment, not to be used as a desperate measure for patients with refractory end-stage disease, but rather as a planned procedure offered to patients in first complete remission of an acute leukaemia, in first blast phase of chronic myeloid leukaemia, or for as soon as a donor was identified in patients with aplastic
anaemia. Concurrently, autologous HSCT was demonstrated to be an effective treatment option for patients with chemo- or radiotherapy-sensitive malignancies, such as lymphoid malignancies\(^{57}\) and some childhood tumours (e.g. neuroblastoma),\(^{58}\) as extramedullary toxicity of the conditioning became the sole limiting factor of dose intensification.

This new chapter in HSCT was marked by important breakthroughs (Figure 2):

- introduction of new conditioning regimens – the combination of high dose busulphan (BU) and cyclophosphamide (CY)\(^{59}\) avoided irradiation, making HSCT possible in centres without access to total body irradiation (TBI);
- creation of a network of national donor registries\(^ {60}\) – inclusion of tissue-typed volunteers in the donor pool extended the applicability of HCT to patients who didn’t have HLA-matched sibling donors;
- recognition of the negative effect of blood transfusions on subsequent bone marrow grafts\(^{61}\) — transplant recipients can become sensitized to non-HLA antigens by receiving unirradiated transfusional support, increasing the risk of graft rejection;
- advances in immunogenetics of HLA\(^ {62}\) — replacement of the serological typing system by molecular techniques of allele characterisation revealed the full heterogeneity of the major histocompatibility complex (MHC), and made possible a more adequate donor-recipient matching, improving the results of unrelated grafts to the level of HLA-identical sibling transplantation, at least for young patients\(^{63}\).
A systems immunology approach to GVHD

Successful remission with donor lymphocyte infusion after cord-blood transplantation. Cure of sickle cell anaemia with BMT. Recognition of human GVT effect. Cure of lymphoma with autologous BMT. Initial report on use of BMT as cancer treatment. Methotrexate to prevent and treat GVHD. Successful transplantation from unrelated donor. Use of PBSC for alloHSCT.

E. Donnall Thomas - Nobel Prize for advances in cell transplantation in the treatment of human disease.

Allogeneic BMT for leukaemia. Allogeneic BMT for aplastic anaemia. Cure of sickle cell anaemia with BMT. T cell depletion of the graft for GVHD prevention. DNA-based HLA typing. Introduction of reduced-intensity transplants.

Donor Treg infusion to protect from GVHD. Whole-tumour-cell GVAX effective in post-transplant malignancies. Role of NK cell alloreactivity in transplant outcomes.


Figure 2. Timeline showing the main milestones in haematopoietic stem cell transplantation, 1957–2016.

BMT, bone marrow transplantation; CAR, chimeric antigen receptor; DOT, Vδ1+ T cells; GVAX, vaccine comprised of cancer cells genetically modified to secrete granulocyte-macrophage colony-stimulating factor; GVHD, graft-versus-host-disease; GVT, graft-versus-tumour; HLA, human leukocyte antigen; HSCT, haematopoietic stem cell transplantation; PBSC, peripheral blood stem cells.

Adapted with permission from Appelbaum FR. N Engl J Med 2007, 357:1472-1475; Copyright Massachusetts Medical Society.
- use of alternative graft sources – the identification of the glycoprotein CD34 as a marker of haematopoietic progenitor cells and the development of centrifuge technology for bulk blood leukocyte collection provided the means to isolate stem cells from the peripheral blood\textsuperscript{(64)} and from umbilical cord blood\textsuperscript{(13)}.

As the understanding of GVT biology grew, it became clear that the key to a successful alloHSCT resided in the immunologic elimination of malignant stem cells. This concept shifted the focus from the traditional myeloablative preparative regimens and led to the development of reduced-intensity conditioning (RIC) regimens extending HSCT to older and frailer patients.\textsuperscript{(65,66)} However, although safe and effective in patients with slow-growing cancers, the low transplant related mortality rate associated with RIC regimens may be offset by higher relapse rates in patients with advanced diseases.\textsuperscript{(67,68)}

In recent years, research has focused on developing new strategies for increasing disease free survival. However, achieving a separation between GVHD and GVL has proved challenging. It has been shown that graft T cell depletion is associated with a lower incidence of both acute and chronic GVHD, but at the cost of a delayed immune reconstitution with an increased risk of infectious complications, and higher disease relapse rates.\textsuperscript{(69,70)} Moreover, adoptive immunotherapy through subsequent donor lymphocyte infusion (DLI), although effective in boosting immune reconstitution and treating leukaemia relapses, still carry the risk of triggering GVHD.\textsuperscript{(71)} Other approaches include: manipulation of the graft to reduce its alloreactive T cell component, either by eliminating activated T cells,\textsuperscript{(72)} selecting for memory T cells,\textsuperscript{(73,74)} inducing
A systems immunology approach to GVHD

Chapter I

donor T cells tolerance for host-type alloantigens,\(^{(75)}\) or selecting for regulatory T cells;\(^{(76)}\) vaccinations against tumor-associated antigens;\(^{(77,78)}\) immunotoxin-based therapy;\(^{(79)}\) and targeted adoptive cellular immunotherapy.\(^{(80)}\)

Altogether, HSCT is at present the treatment which provides the best chance for cure for many pathologies, including non-haematological malignancies and autoimmune diseases.
2. Graft-versus-host disease

Despite significant progress in the past 20 years, GVHD is still a major barrier to successful HSCT, affecting 30-50% of transplanted patients, and remaining one of the leading causes of long-term morbidity and mortality (Figure 3). The risk of GVHD is multifactorial, increasing with certain conditioning regimens, graft type (peripheral blood stem cells > bone marrow > cord blood), type of HSCT (HLA-mismatched > HLA-matched unrelated > HLA-matched sibling), age of the donor, multiparous female donor, and age of the recipient.

This immunological complication of alloHSCT is caused by donor T cells recognizing the genetically disparate transplant recipient (host) as foreign, thus initiating an immune reaction that leads to tissue damage, which can be further amplified by disturbances in immunological reconstitution and failure to acquire immunological tolerance.
Figure 3. Causes of death after HLA-matched sibling (A) and unrelated donor (B) HSCT.

Adapted with permission from Pasquini MC, Zhu X. Current uses and outcomes of haematopoietic stem cell transplantation: CIBMTR Summary Slides, 2015; Copyright The Medical College of Wisconsin, Inc. and the National Marrow Donor Program.
2.1. Classification and clinical features

Two main forms of GVHD can be recognised, acute and chronic. Although the classical distinction of the two forms of the disease, based on the time of its occurrence (before versus after D+100 post-HSCT) and the main clinical features, are still valid today, the boundaries between acute and chronic GVHD have become less rigid as there is cumulative evidence supporting that GVHD should be viewed as a continuum process (Table 3).\(^{(83)}\) Nevertheless the clinical manifestation of the disease remain central in the current classification of GVHD syndromes. While acute GVHD (aGVHD) affects primarily the skin (erythroderma), liver (cholestasis) and gut (diarrhoea with or without haematochezia) (Table 4),\(^{(84)}\) chronic GVHD (cGVHD) can present itself with a myriad of manifestations across multiple organs, most frequently, the skin (dermatosclerosis), mouth (xerostomia), eyes (xerophthalmia), vagina (scarring and stonis), esophagus (strictures), lung (bronchiolitis obliterans), and musculoskeletal system (fasciitis and joint stiffness) (Table 5).\(^{(83)}\)

However, the pathophysiology underlying each of the syndromes is quite distinct. Whilst, aGVHD can be viewed as a distortion of the normal inflammatory cellular responses, involving alloreactive donor T cell-mediated cytotoxic damage to the recipient’s tissues via cell-surface and secreted factors,\(^{(85)}\) cGVHD is mainly characterised by end-organ fibrosis, and its features resemble those of well-recognized autoimmune diseases, suggesting similar pathophysiology.\(^{(86)}\)
Table 3. Distinguishing acute and chronic GVHD.

<table>
<thead>
<tr>
<th>Category</th>
<th>Time of symptoms</th>
<th>aGVHD features</th>
<th>cGVHD features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute GVHD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classic acute</td>
<td>≤ 100 days</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Persistent, recurrent or late onset acute</td>
<td>&gt; 100 days</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Chronic GVHD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classic chronic</td>
<td>No time limit</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Overlap syndrome</td>
<td>No time limit</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

aGVHD, acute GVHD; cGHVD, chronic GVHD


Table 4. Organ involvement in acute GVHD.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Skin</th>
<th>Liver (bilirubin)</th>
<th>Gut (stool output/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No GVHD rash</td>
<td>&lt; 2 mg/dl</td>
<td>&lt; 500 ml/day or persistent nausea.</td>
</tr>
<tr>
<td>1</td>
<td>Maculopapular rash</td>
<td>2–3 mg/dl</td>
<td>500–999 ml/day</td>
</tr>
<tr>
<td></td>
<td>&lt; 25% BSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Maculopapular rash</td>
<td>3.1–6 mg/dl</td>
<td>1000–1500 ml/day</td>
</tr>
<tr>
<td></td>
<td>25 – 50% BSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Maculopapular rash</td>
<td>6.1–15 mg/dl</td>
<td>Adult: &gt;1500 ml/day</td>
</tr>
<tr>
<td></td>
<td>&gt; 50% BSA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Generalized erythroderma plus bullous formation</td>
<td>&gt;15 mg/dl</td>
<td>Severe abdominal pain with or without ileus</td>
</tr>
</tbody>
</table>

BSA, body surface area

Adapted with permission from Jacobsohn DA. Acute graft-versus-host disease in children. Bone Marrow Transplant 2008, 41(2): 215-21; Copyright Macmillian Publishers Ltd.
Table 5. Organ involvement in chronic GVHD.

<table>
<thead>
<tr>
<th>Organ or site</th>
<th>Diagnostic features†</th>
<th>Distinctive features§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>• Poikiloderma</td>
<td>• Depigmentation</td>
</tr>
<tr>
<td></td>
<td>• Lichen planus-like features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Sclerotic features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Morphea-like features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Lichen sclerosus-like features</td>
<td></td>
</tr>
<tr>
<td>Nails</td>
<td>• Dystrophy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Longitudinal ridging, splitting, or brittle features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Onycholysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Pterygium unguis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Nail loss (usually symmetric; affects most nails)</td>
<td></td>
</tr>
<tr>
<td>Scalp and body hair</td>
<td>• New onset of scarring or nonscarring scalp alopecia (after recovery from chemoradiotherapy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Scaling, papulosquamous lesions</td>
<td></td>
</tr>
<tr>
<td>Mouth</td>
<td>• Lichen-type features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Hyperkeratotic plaques</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Restriction of mouth opening from sclerosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Xerostomia</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Mucocoele</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Mucosal atrophy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Pseudomembranes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ulcers</td>
<td></td>
</tr>
<tr>
<td>Eyes</td>
<td>• New onset dry, gritty, or painful eyes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Cicatricial conjunctivitis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Keratoconjunctivitis sicca</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Confluent areas of punctate keratopathy</td>
<td></td>
</tr>
<tr>
<td>Genitalia</td>
<td>• Lichen planus-like features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Vaginal scarring or stenosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Erosions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Fissures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ulcers</td>
<td></td>
</tr>
<tr>
<td>GI tract</td>
<td>• Esophageal web</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Strictures or stenosis in the upper to mid third of the esophagus</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>• Bronchiolitis obliterans diagnosed with lung biopsy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Bronchiolitis obliterans diagnosed with pulmonary function tests and radiology</td>
<td></td>
</tr>
<tr>
<td>Muscles, fascia, joints</td>
<td>• Fasciitis</td>
<td>• Myositis or polymyositis</td>
</tr>
<tr>
<td></td>
<td>• Joint stiffness or contractures secondary to sclerosis</td>
<td></td>
</tr>
</tbody>
</table>

† Sufficient to establish the diagnosis of chronic GVHD
§ Seen in chronic GVHD, but insufficient alone to establish a diagnosis of chronic GVHD

2.2. Acute GVHD immunopathology

Most of our current knowledge about the biology of GVHD is derived from observations made in animal models of the disease. Early studies established that the development of GVHD required: (1) an immunologically competent graft containing mature T cells; (2) inability of the recipient to reject the transplanted cells; and (3) expression of tissue antigens by the recipient absent in the transplant donor. The current conceptual model is that its pathogenesis can be divided into 5 sequential steps (Figure 4).

a) Priming of the immune response

The first phase of acute GVHD is triggered by tissue injury induced by the primary disease, prior chemo- and radiotherapy treatments or infectious complications, and further exacerbated by the HSCT conditioning regimen. The damaged tissues produce a storm of pro-inflammatory cytokines (e.g. TNF-α, IL-1, IL-6) which lead to the up-regulation of a variety of “danger signals”, such as adhesion and costimulatory molecules, MHC antigens, and chemokines, activating host antigen presenting cells (APC).

Damage of gastrointestinal mucosa and disruption of the intestinal microbiome are of particular importance in this process. Gastrointestinal toxicity is known to be one of the most important dose-limiting factors in high-dose conditioning for HSCT. Changes to the permeability of the gut mucosal barrier allows bacterial and/or endotoxin translocation into the systemic
Figure 4. Progression of events in aGvHD development: immune priming (A), activation (B), T-cell expansion (C), T-cell trafficking (D) and host tissue injury (E).
Endotoxins are potent stimulators of toll-like receptors, leading to the release of an array of cytokines that enhance tissue inflammation. Conversely, in murine HSCT models in which the integrity of intestinal barrier is protected, the severity of GVHD is significantly diminished.

Recent studies looking into the biology of intestinal epithelial cells and the intestinal microbiota have unveiled a mechanistic link between changes in microbiome diversity, intestinal epithelial cell preservation and GVHD immunopathology. For instance, it has been shown that a decrease in the amount of the microbial metabolite butyrate, a short-chain fatty acid histone deacetylase inhibitor, in the intestinal tissue of alloHSCT recipients resulted in decreased acetylation of histone H4 in intestinal epithelial cells, leading to apoptosis. Moreover, direct intragastric administration of exogenous butyrate restored acetylation of histone H4, protected epithelial barrier by altering the expression ratio of anti-apoptotic to pro-apoptotic molecules and increasing the expression of proteins relevant for junctional integrity, which resulted in decreased GVHD severity.

The role of TLR signalling has further been recognised in the skin; in studies where an activator of TLR7 was topically applied to the epidermis before inducing GVHD, extensive T cell infiltrates and GVHD histopathologic changes were only observed at the site of treatment. Other relevant molecules released following conditioning regimen-induced tissue damage, such as the damage-associated molecular patterns (DAMP), have been implicated in the induction of GVHD. For instance, binding of extracellular ATP to P2X7 receptor and activation of the inflammasome leads to the over-expression of
co-stimulatory molecules on antigen presenting cells (APC).\textsuperscript{(105)}

Altogether, these findings are consistent with the correlation observed between the risk of developing GVHD and the intensity of the HSCT conditioning regimens.\textsuperscript{(106-108)}

**b) Activation of donor T cells**

Once primed, tissue APC migrate to the secondary lymphoid organs (SLO) where they induce T cell activation. This process requires cognate interaction between donor T cells and alloantigen presenting APC. The recognition of host allopeptide presented on the MHC of APC by donor lymphocyte T cell receptor (TCR) is not by itself sufficient to induce T cell activation;\textsuperscript{(109)} costimulatory signals, such as CD27, OX40, ICOS, CD40, 4-1BB, are required for T cell proliferation, differentiation and survival (Figure 5).\textsuperscript{(110,111)}

It has been shown in murine models, that GVHD is primarily induced by donor T cell interactions with host-derived haematopoietic APC,\textsuperscript{(112)} even though, once initiated, both donor- and host-derived APC are sufficient to further promote T cell activation (Figure 6).\textsuperscript{(113)} Donor T cell activation can be achieved both by direct presentation on host APC (the peptide bound to allogeneic MHC molecules or allogeneic MHC molecules without peptide are recognised by donor T cells) and indirect presentation on donor APC (peptide generated by degradation of the allogeneic MHC molecules is presented on self-MHC).\textsuperscript{(114-116)} However, to this date, no study using stringent conditional ablation of individual APC subsets has been able to demonstrate the absolute requirement of any single recipient APC subset for GVHD induction.\textsuperscript{(95)}
**Figure 5.** Critical interactions for induction of GVHD.

Figure 6. Antigen presentation in GVHD in MHC-matched allogeneic stem-cell transplantation. (A) Allogeneic immune reaction is initiated by donor CD8+ T cells recognizing the peptide products of polymorphic genes (miHAs). In the class I antigen presentation pathway, host APC processes cytosolic proteins through the proteasome and the resulting peptides, including miHAs, are transported into the endoplasmic reticulum, where loading onto MHC class I molecules occurs. Peptide–MHC class I complexes exported to the cell surface are then made available for antigen recognition by donor CD8+ T cells. Initial CD4+ T-cell activation, on the other hand, can be directed against endocytosed antigens by host- or donor-derived APC and processed by the class II pathway. (B) When donor-derived APC have fully replaced the host-derived population, activation of donor CD8+ T cells would solely rely on cross-presentation of exogenously acquired antigens, through the uptake of apoptotic recipient or shed proteins in a process involving phagosome retrotranslocation. TAP, transporter associated with antigen processing.

The role of non-haematopoietic APC in the induction of GVHD is still controversial. Although it has been shown that donor CD4+ T cells can be activated by non-haematopoietic APC in the gastrointestinal tract, haematopoietic APC are critical for CD8+ T cell priming, particularly in the context of MHC-compatible but miHA-mismatched HSCT.

c) Alloreactive T cell proliferation and differentiation

Following priming, donor T cells proliferate and differentiate into effector cells, a process characterised by gain of cytolytic function and ability to secrete cytokines, such as IL-1, IL-3, IL-7, GM-CSF, TNF-α and IFN-γ. Differentiation into Th1/Tc1, Th2/Tc2 and Th17/Tc17 is dependent on the cytokine milieu. For instance, it has been shown that while in the presence of IL-12, CD4+ T cells differentiate into IFN-γ producing Th1 cells, IL-4 induces differentiation into Th2 cells producing IL-4, IL-5, and IL-13; as for Th17 cells, TGF-β and IL-6 is required for their differentiation, and IL-23 and IL-21 are critical for their expansion and survival. Although it has been demonstrated that the balance between Th1/Tc1, Th2/Tc2 and Th17/Tc17 subsets impacts the manifestation of GVHD, the exact contribution of each of these elements is still under active investigation.

The pro-inflammatory molecules IFN-γ, IL-2 and TNF-α produced by Th1/Tc1 cells have been implicated in the pathogenesis of GVHD, being linked to an increased mortality and to the classical syndrome characterised by weight loss, diarrhoea and skin changes. However, the role of some cytokines is still controversial, as IFN-γ was shown to both regulate immune suppression and support cellular cytotoxicity.
Conversely, Th2 polarization of T cells has been reported to reduce aGVHD.\textsuperscript{(123,124)} IL-2 and IL-4 treatment of donor mice was shown to induce generation of IL-4 and IL-10 producing Th2 cells which induced protection from tissue injury associated with GVHD when transferred into sublethally irradiated mice or nonirradiated F1 recipients.\textsuperscript{(125,126)} Furthermore, in miHA mismatch BMT models of GVHD, transplantation of Th2 polarized cells resulted in a significant reduction in GVHD associated mortality.\textsuperscript{(127)}

Th17/Tc17 cell subsets have been proposed to have a dual role after HSCT, in particular on the gastrointestinal tract. IL-17 pro-inflammatory action is well established in several autoimmune diseases,\textsuperscript{(128-130)} and its role mediating GVHD has also been demonstrated in experimental models.\textsuperscript{(131)} On the other hand, like group 3 innate lymphoid cells (ILC3),\textsuperscript{(132,133)} Tc17 cells produce IL-22, a cytokine that acts on the intestinal epithelial stem cell compartment to limit apoptosis and preserve mucosal barrier function.\textsuperscript{(134)} However, with the onset of acute GVHD the differentiation, homeostasis and function of host ILC3 are compromised, and decrease in intestinal ILC3 is directly associated with enhanced intestinal tissue damage owing to a loss of IL-22. Post-HSCT Th17/Tc17 differentiation may thus be central for mucosal immunity reconstitution and restoration of its barrier function, but when dysregulated the primary pro-inflammatory effects are responsible for the induction of severe gastrointestinal damage.

d) Migration of effector cells to GVHD target tissues

After being retained in the SLO for a period of 2 to 3 days during which proliferation and differentiation occurs,\textsuperscript{(135)} effector T (T\textsubscript{eff}) cells migrate into the
tissues in a process that involves changes in vascular permeability and requires specific selectin-ligand, chemokine-receptor, and integrin-ligand interactions.\textsuperscript{(136)}

T cell egress from the lymphoid organs is thought to be mediated by sphingosine 1-phosphate (S1P) signalling.\textsuperscript{(137)} Experimental work using FTY720, an antagonist drug that targets S1P receptors, demonstrated that the expression of S1P receptor 1 on T cells and the establishment of a S1P gradient, low in the LN parenchyma and high at the exit site, was critical for T cells to exit the SLO. Furthermore, in a murine model of GVHD, FTY720 treatment starting at the time of transplantation was shown to be an effective strategy to prevent target organ T cell infiltration.\textsuperscript{(138)}

During differentiation, T cells up-regulate a number of chemokine receptors, such as CCR2, CCR5, CXCR3, and CXCR6, and produce IFN-\(\gamma\) and TNF-\(\alpha\). These pro-inflammatory cytokines enter the circulation and induce host cells in the peripheral tissues to produce an array of chemokines (e.g. CXCL9, CXCL10, and CXCL11),\textsuperscript{(139)} creating an inflammatory environment that directs T\(_{\text{eff}}\) cells to the target organs.\textsuperscript{(136)}

Upon entering the peripheral tissues, donor T cells induce a second wave of chemokine ligand production, including CCL3, CCL4, and CCL5, which help maintain target organ infiltration when CXCL9, CXCL10, and CXCL11 production has subsided.\textsuperscript{(140,141)}

In addition to chemokines and their receptors, migration of inflammatory cells to target organs is regulated by the expression of selectins, integrins and their ligands.\textsuperscript{(142,143)} Homing of donor T cells to intestinal mucosa requires
interaction between the α4β7 integrin (also known as lymphocyte Peyer patch adhesion molecule – LPAM), expressed by T cells, and its ligand, the mucosal vascular addressin cell adhesion molecule 1 (MadCAM-1),\(^{(143,144)}\) while entry of effector T cells into the lung and liver involves interactions between αLβ2 (lymphocyte function-associated antigen 1 – LFA-1) / intercellular adhesion molecule 1 (ICAM-1), and αLβ2/ICAM-1, -2, and -3, respectively.\(^{(136)}\)

e) Target tissue destruction by T\(_{\text{eff}}\) cells

GVHD target organ damage is the result of a complex immune response in which cytotoxic T cells (CTL), NK cells and inflammatory cytokines induce antigen-dependent and -independent cell death (Figure 4-E).

Data from murine models of GVHD using T cell donors deficient for FasL, perforin or granzyme B, or by the in vivo administration of neutralizing anti-FasL antibodies, demonstrated that perforin and granzyme B dependent cytolysis and Fas-mediated apoptosis are two of the main mechanisms of cytotoxicity employed by CTL.\(^{(145,146)}\) Supporting these observations, it has been reported in human patients a positive correlation between the serum levels of soluble FasL and Fas and the clinical severity of GVHD.\(^{(147,148)}\) Notably, different cytotoxic pathways dominate in each organ, while CTL use preferentially the Fas/FasL pathways to induce liver damage, the perforin and granzyme pathways appear to be more important in the gastrointestinal tract and skin.\(^{(91,149)}\)

Although CTL and NK cells differ in the receptors involved in the recognition of target cells, the mechanism by which they induce apoptosis are fundamentally the same. However, NK cell cytotoxicity appears to be mediated
primarily through perforin/granzyme-dependent processes, even though they express FasL and are likely to use these molecules to kill certain target cells.\textsuperscript{(150)}

Other death ligand receptor signalling pathways that have been suggested to be implicated in GVHD target tissue damage are TWEAK/Fn14,\textsuperscript{(151)} TRAIL,\textsuperscript{(152)} LIGHT/LTβ,\textsuperscript{(153)} However, their distinct role in GVHD is still uncertain.\textsuperscript{(154,155)}

The secretion of inflammatory mediators, e.g. TNF-α, IL-1, IFN-γ, and nitric oxide (NO), can also induce GVHD target organ injury independent of any cognate interaction. The relevance of TNF-α in GVHD immunopathology has been made clear in trials where prophylaxis and treatment of steroid-resistant GVHD patients with a TNF-α blocker have shown efficacy, particularly against gastrointestinal disease.\textsuperscript{(156-158)}

IL-1 has long been considered the second major pro-inflammatory cytokine in the effector phase of acute GVHD,\textsuperscript{(159)} being predominantly secreted in the spleen and skin.\textsuperscript{(160)} Several studies in murine models supported this concept, as neutralization of IL-1 significantly reduced GVHD mortality\textsuperscript{(161)} and IL-1 blockade using IL-1Ra downregulated Th17 cell differentiation and reduced the severity of GVHD.\textsuperscript{(162)} However, its relevance in human GVHD pathophysiology is uncertain as IL-1Ra treatment in a controlled randomised clinical trial has proved to be insufficient to reduce GVHD or to improve survival.\textsuperscript{(163)}

IFN-γ has been reported to have ambivalent effects \textit{in vivo}, as it can both amplify and suppress acute GVHD.\textsuperscript{(122,164)} Regarding amplification, IFN-γ
increases the expression of several surface molecules such as chemokine receptors, MHC proteins, and adhesion molecules in APC and stromal cells, enhancing their response to pro-inflammatory stimuli.\(^{(165)}\) Moreover, NO mediated damage to the intestinal epithelium and the epidermis is primarily dependent on donor-derived IFN-\(\gamma\).\(^{(166)}\) However, by accelerating the apoptosis of activated donor T cells, IFN-\(\gamma\) can both dampen GVHD severity and induce profound deficits in immune function after HSCT.\(^{(167)}\)

Further to the traditional Th1/Th2 paradigm of T cell polarisation, increasing evidence supports the role of Th17/Tc17 cells in the pathophysiology of GVHD. Studies in which the cytokines that drive Th17 differentiation were neutralised, in particular IL-6 and IL-23, showed a significant reduction in the severity of acute GVHD.\(^{(168-171)}\) In accordance with these results, blockage of transcription factors critical to Th17 generation (e.g. ROR\(\gamma\)t, ROR\(\alpha\), IRF-4) resulted in diminished GHVD, and was associated with lower levels of IL-17A and TNF-\(\alpha\) being detected in the serum and GVHD target organs.\(^{(172)}\)

Many of the cytokines produces by Th17 cells (IL-21, IL-22, TNF-\(\alpha\), GM-CSF, IL-17F) have now been linked to the development of both acute and chronic GVHD.\(^{(173)}\) However, their individual pathogenic role is still difficult to interpret independently of coexpressed Th1 cytokines. Additionally, it is now recognised that Th17 cells display a broad context-dependent plasticity, associated with higher in vivo survival and self-renewal capacity, and are able of acquiring functional features of Th1 cells, a characteristic that may prove critical for GVHD pathology.\(^{(174)}\)
3. Pre-clinical models of GVHD

As described above, some of the major advances in HSCT and much of the current understanding of GVHD biology was derived from experimental work largely in preclinical animal models with rodents, dogs and, to lesser extent, non-human primates (Table 6). As an outbred large animal model, with a diversity of leukocyte antigens similar to the HLA, the dog has been invaluable in the development of conditioning regimens, prophylaxis and treatment of GVHD, and understanding of the pathology of GVHD. However, such models are greatly limited by cost and experimental group size, time/duration of experiments, lack of detailed or validated immune reagents, and inability to assess GVT effects due to the unavailability of compatible tumour lines.

Since the early experimental work performed in the mid-20th century, the inbred mouse model has had a central place in the progression of our knowledge in HSCT. Having a highly controlled genetic background, with a variety of genetic modified strains available, and being kept on a consistent diet, under tightly regulated specific pathogen-free housing conditions, they provide the ideal experimental system in which researchers can dissect individual variables of disease and exclude the effect of external factors that interfere disease development and progression. Even though the aforementioned characteristics of the murine models warrant great reproducibility of the results,
some disparities in radio-sensitivity and GVHD severity\textsuperscript{(180)} have been observed that may arise from differences in animals from different vendors, related to differences in feed, microbiomes, specific pathogen-free environments, or genetic drift.\textsuperscript{(181-183)}

Furthermore, there are significant inter-specie differences that need to be considered before broad conclusions can be drawn and translated into clinical practice:\textsuperscript{(91)} distinct conditioning regimen modalities induce different severities of tissue inflammation;\textsuperscript{(104,184)} the depth of genetic disparities between donor and recipient affect the kinetics and pathophysiology of GVHD and GVT;\textsuperscript{(185)} differences in ratio and homing characteristics of the immune cells infused affect the pattern of GVHD or GVT induction;\textsuperscript{(186)} the recipient’s microbiome and its changes due antibiotic prophylaxis and/or treatment can affect the response to the conditioning regimen, marrow engraftment and GVHD development;\textsuperscript{(100,187,188)} post-HSCT immune reconstitution and GVHD severity are influenced both by the donor an recipient’s age, related to increased immune senescence, impaired tissue repair mechanisms, increased alloreactivity of recipient APC and lymphocytes, and underlying chronic low-grade inflammation status.\textsuperscript{(189,190)}a
A systems immunology approach to GVHD

Table 6. Advantages and disadvantages of various preclinical model systems.

<table>
<thead>
<tr>
<th>Species</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| **Murine model**       | - MHC well defined (major mismatch vs minor mismatch HSCT models)  
                          - Variety of genetic knock out models available  
                          - Environmental/housing control  
                          - Relatively short life span  
                          - Predictable, reproducible response  
                          - Allows for specific experimental intervention testing  
                          - Transplantation models only recapitulate single organ-specific GVHD pathologies  
                          - SPF environment, cannot mimic the interactions between the host immune system and environmental pathogens  
                          - Microbiome diversities do not mimic the complexity of the human scenario  
                          - Tumours are derived from cell lines rather than spontaneous tumours |
| **Canine model**       | - Similar environmental exposures to humans  
                          - Massive diversity in dog leukocyte antigens similar to human leukocyte antigens  
                          - Mimics both unrelated and related haploidentical bone marrow transplant in the clinic  
                          - Spontaneous haematological tumour development |
| **Non-human primate model** | - MHC locus closely resembles that of humans and may mimic the different degrees of MHC mismatch observed in the transplantation setting  
                          - Share similar near-universal exposure to pathogens  
                          - Similar response to conditioning regimens, as well as toxicity and GVHD incidence; cellular immunity composition and markers are similar to humans |
|                        | - Haplotyping of non-human primates is difficult  
                          - Lack of spontaneous tumours  
                          - Lack of reagents available for immunological studies  
                          - Limited numbers of non-human primates available for studies, and prohibitive costs  
                          - Variability in the amount of irradiation conditioning required due to difference in body fat, age or health status makes standardization difficult  
                          - Requires specialized veterinary technician assistance, and housing requirements limit the numbers of animals per cage, increasing the housing costs and space needed  
                          - Ethical concerns regarding primates cognitive and emotional abilities, and very long lifespan |

SPF, specific pathogen-free

Adapted with permission from Stolfi JL, Pai CC, Murphy WJ. Preclinical modeling of haematopoietic stem cell transplantation - advantages and limitations. FEBS J 2016, 283(9): 1595-606; Copyright John Wiley and Sons.
3.1. Murine models

Current murine models of HSCT can be broadly grouped into those in which GVHD is directed to MHC (class I, class II, or both) or to single or multiple miHA only, and they usually involve the transplantation of bone marrow (as a source of haematopoietic stem cells), supplemented with donor lymphocytes, into lethally irradiated allogeneic recipients.\(^{(180)}\) Depending on the biological question being addressed, adjustments to the number, phenotype and time point of transfer of donor T cells can be made.

Several strain combinations of donor-recipient pairs have been studied and the dominant features of each model characterised.\(^{(191)}\) For instance, based on the differences in MHC class I and/or II molecules (Figure 7 and Table 7) or minor histocompatibility antigens it can be predicted whether the effectors of GVHD are either CD4\(^+\) or CD8\(^+\) T cells, or both (Table 8). GVHD tissue injury in the context of a full MHC mismatch, both for class I and II molecules (e.g. B6→BALB/c) is primarily dependent on CD4\(^+\) T cells, with CD8\(^+\) T cells providing only additive pathology.\(^{(154)}\) These models of CD4\(^+\)-dependent GVHD are characterised by the development an inflammatory “cytokine storm”, capable of inducing damage in the target tissues without requirement for cognate T cell interaction with MHC \textit{in situ}.\(^{(192)}\) By contrast, CD8\(^+\) T cell-mediated GVHD injury (e.g. B6→bm1) requires the TCR to engage MHC on target tissue in order to trigger the cytolytic machinery.\(^{(193)}\)
Figure 7. Genomic map of the mouse MHC complex

Table 7. H-2 haplotypes and alloantigens of some of the most commonly used laboratory mouse strains.

<table>
<thead>
<tr>
<th>Mouse Strains</th>
<th>MHC Haplotypes</th>
<th>MHC Class I</th>
<th>MHC Class II</th>
<th>MHC Class Iβ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>H-2K</td>
<td>H-2D</td>
<td>H-2L</td>
</tr>
<tr>
<td>129/−</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>A/J</td>
<td>a</td>
<td>k</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>BALB/B</td>
<td>d</td>
<td>b</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>C3H/He</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>null</td>
</tr>
<tr>
<td>C57BL/6</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>C57BL/10</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>C57L/J</td>
<td>bc</td>
<td>b</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>CBA/J</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>null</td>
</tr>
<tr>
<td>DBA/1</td>
<td>q</td>
<td>q</td>
<td>q</td>
<td>q</td>
</tr>
<tr>
<td>DBA/2</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>FVB/N</td>
<td>q</td>
<td>q</td>
<td>q</td>
<td>q</td>
</tr>
<tr>
<td>LP/J</td>
<td>bc</td>
<td>b</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>NOD</td>
<td>g7</td>
<td>d</td>
<td>b</td>
<td>null</td>
</tr>
<tr>
<td>P/J</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>RF/J</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>null</td>
</tr>
<tr>
<td>SWR/J</td>
<td>q2</td>
<td>q</td>
<td>q</td>
<td>q</td>
</tr>
</tbody>
</table>
Table 8. Commonly used mouse models of HSCT.

<table>
<thead>
<tr>
<th>Donor strain</th>
<th>Recipient strain</th>
<th>Genetics</th>
<th>T cell dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MHC-mismatched</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>BALB/C (H-2D&lt;sup&gt;6&lt;/sup&gt;)</td>
<td>Mismatched for MHC-I, MHC-II and miHA</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; and CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C3H/HeJ (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>C57BL/6 (H-2D&lt;sup&gt;6&lt;/sup&gt;)</td>
<td>Mismatched for MHC-I, MHC-II and miHA</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; and CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>B6C3F1 (H-2D&lt;sup&gt;kb&lt;/sup&gt;)</td>
<td>Mismatched for MHC-I, MHC-II and miHA</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; and CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>B6D2F1 (H-2D&lt;sup&gt;bkb&lt;/sup&gt;)</td>
<td>Mismatched for MHC-I, MHC-II and miHA</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; and CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C57Bl6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>B6AF1 (H-2D&lt;sup&gt;bka&lt;/sup&gt;)</td>
<td>Mismatched for MHC-I, MHC-II and miHA</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; and CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>B10.BR (H-2D&lt;sup&gt;6&lt;/sup&gt;)</td>
<td>Mismatched for MHC-I, MHC-II and miHA</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; and CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>B6.C-H2bm1 (bm1) (H2b background with mutation at MHC I)</td>
<td>MHC I mismatch</td>
<td>CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>B6.C-H2bm12 (bm12) (H2b background with mutation at MHC II)</td>
<td>MHC II mismatch</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>miHA mismatched</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B10.BR (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>CBA (H-2D&lt;sup&gt;6&lt;/sup&gt;)</td>
<td>miHA mismatches</td>
<td>CD8&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>B10.D2 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>DBA/2 (H-2D&lt;sup&gt;5&lt;/sup&gt;)</td>
<td>miHA mismatches</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; (minor CD8&lt;sup&gt;+&lt;/sup&gt; contribution)</td>
</tr>
<tr>
<td>B10.D2 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>BALB/C (H-2D&lt;sup&gt;6&lt;/sup&gt;)</td>
<td>miHA mismatches</td>
<td>CD4&lt;sup&gt;+&lt;/sup&gt; (minor CD8&lt;sup&gt;+&lt;/sup&gt; contribution)</td>
</tr>
<tr>
<td>C3H.SW (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>miHA mismatches</td>
<td>CD8&lt;sup&gt;+&lt;/sup&gt; (with CD4&lt;sup&gt;+&lt;/sup&gt; help)</td>
</tr>
<tr>
<td>C57BL/6 (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>129/Sv (H-2D&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>miHA mismatches</td>
<td>CD8&lt;sup&gt;+&lt;/sup&gt; (with CD4&lt;sup&gt;+&lt;/sup&gt; help)</td>
</tr>
</tbody>
</table>

Adapted with permission from Schroeder MA, DiPersio JF. Mouse models of graft-versus-host disease: Advances and limitations. Dis Model Mech 2011, 4(3): 318-33; Copyright Elsevier Inc.
Due to their capacity to induce the full spectrum of clinically relevant GVHD, allowing the *in vivo* and *ex vivo* measurement of immunologic pathways, the MHC disparate systems have been the most commonly used model employed to dissect the immunologic mechanisms of GVHD. However, as MHC-mismatched transplants in humans are not commonly performed, miHA-mismatched models of acute GVHD are considered to be better representative of human biology.\(^{(191)}\)

Similar to what is found in humans, miHA-mismatched models exhibit less morbidity than MHC-mismatched models, but still result in lethal acute GvHD. Interestingly, in mouse miHA-mismatched HSCT models, differences in MHC haplotypes (*Figure 7*) influence the type and specificity of the T-cell response underlying the GVHD phenotype.\(^{(191)}\) For example, in the B10.D2 → BALB/c or DBA/2 (H-2\(^{D}D\)) models, acute GVHD is primarily dependent on CD4\(^{+}\) T cells and less on CD8\(^{+}\) T cells,\(^{(194)}\) while in C3H.SW→B6 and B6→129/Sv (H-2\(^{D}D\)), GvHD depends primarily on CD8\(^{+}\) T cells.\(^{(193)}\)

Much of the described variability in these miHA-mismatched HSCT models is believed to be related to the fact that the disparities between donor and recipient are in multiple antigens.\(^{(195)}\) Although more than 200 miHA loci have been described in murine models using congenic mouse strains, immune responses are restricted to a few dominant miHA, including H4, H7, H13, H28, H60, and H-Y.\(^{(196-201)}\) Model-specific GVHD causing mechanisms and pattern of target organ involvement can be a result of the different tissue distribution of
dominant miHA\textsuperscript{(202)} or differences in MHC affinity for antigens found in certain target tissues.\textsuperscript{(191)}

To evaluate the relevance of individual miHA in triggering GVHD, several antigen-specific transgenic TCR models have been established.\textsuperscript{(191)} By restricting the specificity of the donor T cells to a single peptide epitope it is possible to overcome the effect of antigen-independent bystander activation of T cells and investigate several aspects of GVHD, including the role of cross presentation of antigens, antigen affinity, and antigen-induced T-cell activation and proliferation.\textsuperscript{(191)} Examples of TCR-transgenic mice used in GVHD research include the MataHari and Marilyn systems (H-Y),\textsuperscript{(118)} TEa system (H2-E\textsuperscript{a}),\textsuperscript{(203)} D011.10 and OT-I (ovalbumin).\textsuperscript{(204,205)}

Clinical data from MHC-matched HSCT show an increased risk of developing GVHD in male patients receiving grafts from female donors (F $\rightarrow$ M),\textsuperscript{(206)} in a process that appears to be dependent on H-Y-specific alloresponses.\textsuperscript{(207-209)} Due to its ubiquitous tissue expression both in human and mice, research focused on H-Y dependent alloreactions in GVHD have proven to be particularly useful, not only to study the pathologic mechanisms underlying gender mismatch HSCT, but also to dissect the biology of immune dysfunction in the context of miHA disparities in general.\textsuperscript{(118,210,211)}

To further explore the cellular and molecular mechanism involved in the pathogenesis of GVHD, several genetically modified strains have been created. The usage of knockout or transgenic mice as donors or recipients in HSCT has allowed the study of the effect of specific cell populations or molecules on the development of alloreactive responses. For example: the usage of
A systems immunology approach to GVHD

Chapter I

(β2m⁻⁻→B6) chimeras as recipients for HSCT in miHA-mismatch model of GVHD, created a system in which haematopoietic cells are unable to present antigen through MHC-I, revealed that antigen presentation by haematopoietic-derived APC is required but not obligatory for the induction of optimal CD8⁺ T cell mediated GVHD; transfer of Ccr5⁻⁻ T cells in the context of HSCT corroborated the role of CCR5-CCL5 interactions both for the influx of donor T cells into the small intestine and for the induction of acute GVHD; by using perforin-deficient (Prf1⁻⁻) and granzyme B-deficient (Gzmb⁻⁻) mice as T cells donors in MHC-mismatched model of GVHD it was demonstrated that perforin/granzyme-dependent mechanisms are required for MHC-I restricted acute GvHD but only have a minor role in MHC-II restricted acute GVHD.

However, usage of systems in which genes are constitutively silenced or activated, or where cell populations are altogether absent, requires caution in the interpretation of the results as there are some caveats that need to be considered, in particular regarding developmental abnormalities induced by the genetic alteration. The current controversy regarding the role of Langerhans cells (LC) in the development of cutaneous acute GVHD is an example of such a situation. Being one of the few populations of professional APC that are resistant to the conditioning regimens and persist for long periods after transplantation, LC were regarded as having an important role in GVHD. Merad et al. supported this hypothesis demonstrating that administration of donor T cells to bone marrow-chimeric mice with persistent host Langerhans cells resulted in marked cutaneous GVHD, but were unable to induce skin damage in mice whose host Langerhans cells had been replaced by donor-derived ones. Li et al. contested these results by studying the development
of cutaneous acute GVHD in Langerin-DTA recipients, a system which selectively lacks LC. In this model it was observed that the absence of recipient LC did not abrogate the development of GVHD. However, the authors failed to demonstrate that the constitutive LC depletion did not affect the skin's immune baseline, rendering other cell populations capable of inducing GVHD.

Alternative inducible models have now been created which allow for selective and time limited manipulations of the systems. Still in the context of the study of Langerhans cell biology, Dr Clare Bennett developed the Langerin-DTR model which allows for the selective depletion of LC upon treatment with diphtheria toxin (DT). Using Langerin-DTR mice as recipients in HSCT experiments, and treating them with DT only a few days before HSCT, a previous PhD student in Prof Chakraverty and Dr Clare Bennett’s lab, Dr Thomas Conlan, showed that at the time of transplantation the only change to the recipients skin immune network was the absence of LC in the epidermis. In this highly controlled system, it was confirmed that host LC are required for the development of GVHD.

Although animal models have many limitations that hinder the direct extrapolated of results into clinical application, preclinical modelling remains a valuable tool, allowing the simplification of complex systems and analysis of pathogenesis at the single-variable level in a controlled, reproducible manner.
4. Unanswered questions in GVHD

The results obtained in the animal models and in the clinical trials have been crucial for advancing HSCT practice over the last few years, leading to the development of new modalities of treatment and disease management. However, there are still some major gaps in our understanding of GVHD that require further research.

Although acute GVHD is a systemic disorder, the organs commonly targeted are consistently the gastrointestinal tract, the liver and the skin. It has been proposed that this pattern is due to the inherent pro-inflammatory milieu of these barrier organs, and that it may reflect differences in the distribution of antigens to which donor alloreactive T cells respond.\(^{(218)}\) However, it is well recognised that the conditioning regimen induces an inflammatory status that is not organ specific.\(^{(219)}\) Moreover, in experimental models using transgenic donor T cells that recognises a ubiquitously expressed antigen (e.g. MataHari), the pattern of organ involvement is maintained.\(^{(118)}\)

The current scheme of GVHD development is based on data from models using myeloablative regimens and early transfer of donor T cells.\(^{(220)}\) The advent of reduced intensity conditioning regimens and delayed DLI, introduced the concept of late-onset acute GVHD. Arising at a point when
conditioning-induced inflammation is subsiding, the pathophysiologic processes responsible for its onset have not yet been fully identified.\(^{(90)}\)

The biology of chronic GVHD remains poorly defined, and although specific aspects of its pathophysiology are recapitulated by some animal models, none has been able to reproduce the most frequent modes of induction (i.e. acute GVHD evolving into chronic GVHD; chronic GVHD developing as immunosuppression is weaned).\(^{(221)}\)

Moreover, the separation between the biologic pathways responsible for GVHD and GVL effect remain elusive; consequently, current approaches for GVHD prevention and treatment are broad-spectred and impact negatively on GVL responses and on immune reconstitution.\(^{(180)}\) As systemic corticosteroid therapy remains the standard primary therapy for GVHD, the development of steroid-refractory disease represents a life-threatening incident whose underlying mechanisms are still not understood.\(^{(90)}\)

Recent advances in proteomics, gene profiling and whole genome sequencing have given new insights into the genetic basis of GVHD and to potential risk factors associated with the development of GVHD.\(^{(191)}\) However, thus far, these studies have failed to identify biomarkers with diagnostic significance which may change the paradigm of GVHD treatment from prophylactic to pre-emptive, and provide new means of improving the ability to predict the severity and kinetics of the disease.
5. Project rationale

The current conceptual model of acute GVHD elegantly summarizes its biology, but in the attempt to deconvolute the complexity of the immune reactions involved in GVHD immunopathology to a linear, stepwise sequence of events compartmentalised into discrete stages, it fails to recognise the dynamic nature of the system, where cell populations play a diversity of roles depending on milieu, space and time. Indeed, most of the knowledge on the cellular and molecular mechanisms involved in GVHD pathogenesis come from a collection of results obtained in studies mainly focused on the spleen, lymph nodes (LN) and blood, at individual time points, from multiple model systems.

These SLO-centric approaches have thus failed to take into consideration the potential of interactions occurring in the peripheral tissues having a determinant role in regulating the effector phase of GVHD immunopathology. Data from experimental models of infection\(^{(222-225)}\) support the concept of a “two-hit” model of the immune response, in which the range the T cell effector function is not fixed at priming, but rather depend on additional antigen-dependent interactions with APC in the tissues. Such a model, if validated in the acute GVHD setting, could be useful for defining the determinants of target organ specificity.
In the last years, new molecular tools such as gene expression profiling and genomic sequencing have provided the means to better characterise the changes that immune cells undergo during differentiation, and uncover important clues regarding the mechanisms underlying the establishment of protective responses against infections and the pathophysiology of several immune mediated diseases. Microarray and RNA-seq based technologies have also been used in GVHD studies to investigate differences between experimental groups. Most commonly, these studies employ classical tools which are based on the analysis of differential gene expression, and subsequent classification according to gene ontology annotation. Although this methodology has succeeded in identifying genes that are up- or down-regulated in specific conditions (i.e. GVHD vs no GVHD; treated vs untreated; genetically modified vs wild-type), the pathophysiological significance of these findings is frequently difficult to determine, as the gene sets are often heterogeneous and biologically unrelated. Also, depending on the study design and on the heterogeneity of the populations included in each sample, the data may not be able to identify which cell type is involved. Moreover, this approach does not distinguish cause or effect nor how the gene expression changes occurs, and it fails to take into consideration gene-gene interactions which may be critical for normal cellular function or may be relevant for GVHD pathogenesis.\(^{226}\)

Advances in computational power have led to the development of more powerful bioinformatics tools which, instead of using preconceived thresholds of fold change and statistical significance to identify differentially expressed genes, analyse the correlation patterns among genes to generate networks of
co-expressed and co-regulated genes.\textsuperscript{(227,228)} Weighted gene correlation network analysis (WGCNA) is one of such applications which can be used for finding clusters (modules) of highly correlated genes and for summarizing them using the module eigengene (defined as the first principal component of a given module) or the intramodular hub genes (defined as the genes with the highest intramodular connectivity).\textsuperscript{(229)} By relating modules to one another and to external sample traits, new candidate biomarkers or therapeutic targets can be identified.\textsuperscript{(229)} These systems biology methods have been successfully applied in various biological contexts, identifying functionally related genes, enabling a more systematic and global interpretation of gene expression data.\textsuperscript{(230,231)}

In this regard, the work in this thesis represents a novel approach to the study of GVHD, designed to evaluate the changes in the cellular programs of donor CD8$^+$ T cells as acute GVHD develops in clinically relevant murine models of miHA-mismatch HSCT. By analysing the development of the alloreactive immune response at different sites and time points, within the same host, and employing both classical and systems biology methodologies to analyse the data obtained, this study constitutes an innovative multi-dimensional approach which provides valuable insights into the plasticity of CD8$^+$ T\textsubscript{eff} cell programs in the setting of acute GVHD.
6. **Hypothesis**

GVHD target organs are major participants in shaping tissue injury, providing idiosyncratic signals that regulate alloreactive CD8\(^+\) T cell effector function.

7. **Aims and objectives**

- Develop a “systems immunology” methodology to characterise the transcriptional response of donor CD8\(^+\) T\(_{eff}\) cells as they traffic to multiple sites during the evolution of GVHD.
- Describe the main differences in donor CD8\(^+\) T\(_{eff}\) cell gene expression between the SLO and the GVHD target organs.
- Define tissue-specific alloreactive T cell gene signatures and identify the associated cellular programs.
- Validate the gene signatures against independent data sets and across different species.
- Identify the cellular and molecular mechanisms underlying the establishment of tissue-specific gene signatures.
II. Materials & Methods
1. Animal work

1.1. Mice

C57BL/6 and 129/Sv mice were purchased from Charles River Laboratories (Margate, UK) and bred in house by UCL Biological Services (Royal Free Hospital, UK). C57BL/6 Langerin.DTREGFP (Langerin-DTR) mice were kindly provided by Bernard Malissen and Adrien Kissenpfennig (Université de la Méditerranée, France) and bred in house. C57BL/6 TCR transgenic anti-HY (MataHari) mice were kindly provided by Jian Chai (Imperial College London, UK) and bred in house. Ifngr1−/− mice were acquired from The Jackson Laboratory (USA) and bred in house.

Animals used as recipients for HSCT were 10 - 20 weeks old, and those used as donors were 8 - 16 weeks old.

All procedures were conducted in accordance with the United Kingdom Home Office Animals (Scientific Procedure) Act of 1986, and were approved by the Ethics and Welfare Committee of the Comparative Biology Unit, Royal Free and University College London Medical School, London.

1.2. Haematopoietic stem cell transplantation

HSCT was performed as described before with minor modifications.(210) Briefly, recipient mice were lethally irradiated (11 Gy TBI, split into two fractions
over a period of 48 hours, at D-2 and D0) and reconstituted 4 hours later with 5 x 10^6 bone marrow cells, 2 x 10^6 CD4^+ splenocytes and 1 x 10^6 CD8^+ splenocytes, administered by intravenous injection through the tail vein. Isolation of CD4^+ and CD8^+ T cells was performed by immunomagnetic selection of CD4^+ or CD8^+ splenocytes using Manual MACS® Cell Separation Technology (QuadroMACS Separator, LS columns, CD4 (L3T4) MicroBeads, CD8a (Ly-2) MicroBeads; Miltenyi, Germany), according to the manufacturer’s instructions.

In the B6→129 model: 129/Sv male mice (CD45.2^+/Thy-1.2^+) were used as recipients and C57BL/6 female mice (CD45.1^+/Thy-1.1^+) were used as bone marrow and splenocyte donors (alloBMT group); or C57BL/6 female mice (CD45.2^+/Thy-1.2^+) were used as recipients and C57BL/6 female mice (CD45.1^+/Thy-1.1^+) were used as bone marrow and splenocyte donors (synBMT group) (Figure 8).

In the MataHari T cell model: C57BL/6 male mice (CD45.2^+/Thy1.2^+) were used as recipients, C57BL/6 female mice (CD45.2^+/Thy1.2^+) were used
as bone marrow and CD4\(^+\) splenocyte donors, and MataHari female mice (CD45.2\(^+\)/Thy1.1\(^+\)) were used as CD8\(^+\) splenocyte donors (alloBMT); C57BL/6 female mice (CD45.2\(^+\)/Thy1.2\(^+\)) were used as recipients, C57BL/6 female mice (CD45.2\(^+\)/Thy1.2\(^+\)) were used as bone marrow and CD4\(^+\) splenocyte donors, and MataHari female mice (CD45.2\(^+\)/Thy1.1\(^+\)) were used as CD8\(^+\) splenocyte donors (synBMT) (Figure 9).

In the adapted MataHari T cell→Langerin-DTR model: Langerin-DTR male mice (CD45.2\(^+\)/Thy1.2\(^+\)) were used as recipients, C57BL/6 female mice (CD45.2\(^+\)/Thy1.2\(^+\)) were used as bone marrow and CD4\(^+\) splenocyte donors, and MataHari female mice (CD45.2\(^+\)/Thy1.1\(^+\)) were used as CD8\(^+\) splenocyte donors (alloBMT); Langerin-DTR female mice (CD45.2\(^+\)/Thy1.2\(^+\)) were used as recipients, C57BL/6 female mice (CD45.2\(^+\)/Thy1.2\(^+\)) were used as bone marrow and CD4\(^+\) splenocyte donors, and MataHari female mice (CD45.2\(^+\)/Thy1.1\(^+\)) were used as CD8\(^+\) splenocyte donors (synBMT) (Figure 10).
1.3. *In vitro* T cell activation

In the experiments in which *Ifngr*<sup>−/−</sup> mice were used as HSCT recipients, donor T cells were activated *in vitro* prior to transfer. For this purpose, CD4<sup>+</sup> and CD8<sup>+</sup> splenocytes were isolated as previously described and cultured overnight in complete T cell medium (RPMI, 1% FCS, 1% L-glutamine, 1% penicillin-streptomycin, 1% HEPES; Lonza, UK) supplemented with concanavalin A (ConA) (2 µg/ml, final concentration) and IL-7 (1 ng/ml, final concentration).

1.4. Animal welfare

All transplanted animals were kept in germ free conditions, housed in sterilised individually ventilated cages, and fed irradiated chow and acidified UV-treated water.
Animal well-being was assessed by weight monitoring and GVHD severity evaluation of the experimental subjects'. Weight monitoring was started at the day of HSCT (D0), baseline, and all subsequent measurements were reported as percentage changes in body weight compared to baseline. Severity of GVHD was evaluated using a 6 parameter scoring system that takes into account weight loss, posture, activity, fur texture, eye opening and the presence of diarrhoea (Table 9).

Table 9. Assessment of clinical GVHD severity in transplanted subjects.

<table>
<thead>
<tr>
<th>Score</th>
<th>Weight loss</th>
<th>Posture</th>
<th>Activity</th>
<th>Fur texture</th>
<th>Eye opening</th>
<th>Diarrhoea</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;10%</td>
<td>No hunch</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
<td>Absent</td>
</tr>
<tr>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eyes closed up to 25%</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>≥10%</td>
<td>Slight hunch, lost on movement</td>
<td>Mild to moderate reduction in spontaneous movement; resists handling</td>
<td>Angle 0-45°</td>
<td>Eyes closed 25-50%</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Significant hunch, reduced on movement</td>
<td>Severe reduction in spontaneous movement; does not resist handling</td>
<td>Angle 45-90°</td>
<td>Eyes closed 50-75%</td>
<td>Present</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eyes closed 75-100%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Severe hunch, persistent on movement</td>
<td>No spontaneous movement</td>
<td>Angle &gt;90°</td>
<td>Eyes completely closed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Euthanasia of the experimental subjects was performed whenever the weight loss exceeded 20% of the baseline body weight, or the global GVHD severity score (calculated as the sum of the individual characteristics score) was equal or greater than 4.

1.5. Treatments

a) Antibiotic prophylaxis

All HSCT recipients received antibiotic prophylactic treatment from D-2 until the end of the experiment: enrofloxacin (Baytril®; Bayer, Germany) was given in the drinking water at a final concentration of 200 mg/ml.

b) Lymph node lymphocyte egress blockage

Homing trafficking blockage of activated lymphocytes from the lymph nodes to the peripheral tissues was achieved through treatment of transplanted animals with the sphingosine-1-phosphate antagonist, fingolimod (FTY720; Sigma, UK), as previously described. Briefly, HSCT recipients were injected intraperitoneally with 1.0 mg/kg of FTY720 daily, from D+3 to D+7 (experiment terminus); control subjects received an equivalent volume of saline.

c) Langerhans cell depletion

Selective depletion of epidermal Langerhans cells was achieved through treatment of Langerin-DTR mice with diphtheria toxin (DT; Sigma, UK) three weeks prior to HSCT (D-21), as previously described. Briefly: for systemic Langerhans cell depletion, animals received a single intraperitoneal injection of
400 ng of DT; for localised Langerhans cell depletion, animals received intradermal injections of 25 ng of DT in dorsal and ventral sides of the left ear (total DT dose: 50 ng), under general anaesthesia with 2% isoflurane with an oxygen flow rate of 1 L/min; control subjects received an equivalent volume of saline.

d) Notch signalling blockage

Blockage of Notch signalling was achieved through treatment of transplanted animals with a potent gamma secretase inhibitor, LY411575 (Sigma, UK). Animals were treated with 5 mg/kg/day of LY411575 given intraperitoneally, from D+5 to D+7 (experiment terminus); control subjects received an equivalent volume of saline.

2. Tissue and organ harvest

Blood samples were collected into heparinised tubes by venepuncture of the lateral tail vein, for interim analysis, or by cardiac puncture under terminal anaesthesia, at the experiment terminus.

Animals were euthanized by exposure to carbon dioxide gas in a rising concentration. Following intracardiac perfusion with 20 ml of cold PBS to remove the blood from the vasculature, the organs of interest were harvested and stored in harvest medium (PBS, 2% FCS, 1% penicillin-streptomycin; Lonza, UK) on ice. Processing of the samples was started within 2 h from collection.
Harvested organs included: spleen, peripheral lymph nodes (cervical, axillary, brachial and inguinal), mesenteric lymph nodes, tibias and femurs, liver, small intestine (from 0.5 cm below the stomach to 1 cm above the cecum), skin (body and ears).

2.1. Isolation of murine immune cells

a) Lymph nodes and spleens

Preparation of cell suspensions from spleens and lymph nodes was performed according to the laboratory’s protocol, adapted from the procedure described in Current Protocols in Immunology (supplement 39). Briefly, the freshly removed organs were mashed and passed through a 40 µm cell strainer; red blood cells were removed by isotonic lysis with ammonium chloride (ACK Lysing Buffer; Lonza, UK). Cells were resuspended in FACS buffer (PBS, 2% FCS, 2 mM EDTA; Lonza UK) for counting and immunolabelling.

b) Bone marrow

Isolation of bone marrow cells was accomplished following an adaptation of the procedure described in Current Protocols in Immunology (supplement 67). Briefly, both epiphysis of the bones were cut and the bone marrow was flushed out with FACS buffer. The cell suspension was filtered through a 40 µm cell strainer and red blood cells were removed by isotonic lysis with ammonium chloride. Cells were resuspended in FACS buffer for counting and immunolabelling.
c) Blood

Erythrocytes were removed from whole blood samples by hypotonic lysis with distilled water.\textsuperscript{(235)} Cells were resuspended in FACS buffer for counting and immunolabelling.

d) Liver

An adapted version of the basic protocol described in \textit{Current Protocols in Immunology} (supplement 22)\textsuperscript{(236)} was used to isolate liver lymphocytes. Briefly, the excised liver was cut into small fragments (~0.5 cm in diameter) and incubated in digestion medium (RPMI, 2% FCS; Lonza, UK; 200 U/ml collagenase IV; LifeTechnologies, USA; 200 U/ml DNAse I; Sigma, UK) at 37°C with shaking at 150 rpm for 60 min. The liver fragments were further dissociated and passed sequentially through a 100 µm and a 40 µm cell strainer. Density gradient centrifugation with Ficoll-Paque\textsuperscript{TM} Plus (GE Healthcare, UK) was used for lymphocyte enrichment of the cell suspension. Cells were resuspended in FACS buffer for counting and immunolabelling.

e) Small intestine

Isolation of intraepithelial lymphocytes (IEL) and lamina propria (LP) cells from the small intestine was performed according to an adapted version of the protocol described in \textit{Current Protocols in Immunology} (supplement 99).\textsuperscript{(237)} Briefly, the entire small intestine was flushed and rinsed with 40 ml of ice cold harvest medium, and sectioned into ~0.5 cm pieces. The intestinal pieces were incubated with detaching medium (HBSS, 5% FCS, 1% penicillin-streptomycin, 5 mM EDTA; Lonza, UK) at 37°C with shaking at 150 rpm for 60 min. The supernatant, containing the IEL, was passed sequentially through
a 100 µm and a 40 µm cell strainer and enriched for lymphocytes by density gradient centrifugation with Ficoll-Paque™ Plus. The intestinal pieces were further incubated in digestion medium at 37°C with shaking at 150 rpm for 60 min, dissociated and passed sequentially through a 100 µm and a 40 µm cell strainer. The cell suspension, containing the LP cells was enriched for lymphocytes by density gradient centrifugation with Ficoll-Paque™ Plus. Cells were resuspended in FACS buffer for counting and immunolabelling.

f) Skin

Epidermal and dermal immune cells were isolated from the skin in accordance with the protocol described by Henri et al. (238) with minor modifications. Briefly, the body skin was cut into ~1 x 1 cm pieces, after having been shaved and the subcutaneous fat removed, and the ears were split in two parts (ventral and dorsal). The pieces of skin were incubated overnight in dispase medium (HBSS, 2% FCS; Lonza, UK; 2.5 mg/dl dispase II; Sigma, UK) at 4°C, and the epidermal and dermal sheets were separated and cut into ~0.5 cm fragments. The epidermal fragments were vortexed, mashed and passed sequentially through a 100 µm and a 40 µm cell strainer to disintegrate the remaining tissue and create a cell suspension. The dermal fragments were further incubated in digestion medium at 37°C with shaking at 150 rpm for 60 min, dissociated and passed through a 100 µm and a 40 µm cell strainer. Cells were resuspended in FACS buffer for counting and immunolabelling.
3. Cell counting

Total viable cell concentration of the cell suspensions was determined using a CASY® Model TT cell counter and analyser system (Schärfe-System, Germany). Absolute number of specific cell populations was determined by flow cytometry using CountBright™ Absolute Counting Beads (LifeTechnologies, USA) following the manufacturer’s protocol.

4. Immunolabelling

Cell immunolabelling was performed following the protocols described in Current Protocols in Immunology (supplement 44)\textsuperscript{(239)} with minor adaptations. Briefly, before staining, cells were plated out at up to $1 \times 10^6$ cells per well in a 96 well conical bottom plate and incubated with 2.4 G2 antibody at 4°C for at least 10 min to block Fc receptors. For cell surface immunolabeling, cells were incubated with the fluorochrome-conjugated antibodies (Table 10) diluted in 100 μl of FACS buffer at 4°C for 20 min in the dark. For detection of H60 specific CD8$^+$ T cells, samples were labelled with PE-conjugated pentamers for H60/H-2K$^b$ complexes (ProlImmune, UK). When intracellular staining was required, after having performed cell surface immunolabelling, samples were washed twice with FACS buffer, fixed in 100 μl of fixation solution (BD Cytofix/Cytoperm solution; BD Biosciences, UK) for 15 min at 4°C in the dark, washed twice with permeabilization solution (BD Perm/Wash™ buffer; BD
Biosciences, UK), and incubated with the fluorochrome-conjugated antibodies (Table 11) diluted in 100 μl of permeabilization solution at 4°C for 30 min in the dark. For detection of cytokine production, cells were treated with brefeldin A (Sigma, UK) for 2 h at 37°C, prior to immunolabelling. Samples were washed twice with FACS buffer and resuspended in 300 μl of FACS buffer for immediate analysis by flow cytometry. For dead cell exclusion, 2 μl of propidium iodide (LifeTechnologies, USA) was added to the unfixed samples prior to analysis.

Table 10. Antibodies used for cell surface immunolabelling.

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Clone</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4</td>
<td>GK1.5</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD8a</td>
<td>53-6.7</td>
<td>BioLegend, USA</td>
</tr>
<tr>
<td>CD11b</td>
<td>M1/70</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD25</td>
<td>PC61</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD40</td>
<td>HM40-3</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD44</td>
<td>IM7</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD45</td>
<td>30-F11</td>
<td>BioLegend, USA</td>
</tr>
<tr>
<td>CD45.1</td>
<td>A20</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD45.2</td>
<td>104</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD62L</td>
<td>MEL-14</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD69</td>
<td>H1.2F3</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD70</td>
<td>FR70</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD86 (B7-2)</td>
<td>GL1</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD90.1 (Thy-1.1)</td>
<td>HIS51</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>CD90.2 (Thy-1.2)</td>
<td>53-2.1</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD103</td>
<td>M290</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD326 (EpCAM)</td>
<td>G8.8</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>MHC class I (H-2Kd)</td>
<td>SF1-1.1</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>MHC class II (I-A/I-E)</td>
<td>M5/114.15.2</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>γδ TCR</td>
<td>eBioGL3 (GL-3, GL3)</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>Vβ 8.3 TCR</td>
<td>1B3.3</td>
<td>BD Biosciences, Germany</td>
</tr>
</tbody>
</table>
Table 11. Antibodies used for intracellular immunolabelling.

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Clone</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Caspase-3</td>
<td>C92-605</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>CD207 (Langerin)</td>
<td>eBioL31</td>
<td>eBioscience, USA</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>XMG1.2</td>
<td>BD Biosciences, Germany</td>
</tr>
<tr>
<td>TNF-α</td>
<td>MP6-XT22</td>
<td>BD Biosciences, Germany</td>
</tr>
</tbody>
</table>

Fluorescence minus one controls and isotype controls were used to accurately discriminate positive versus negative signals, and the gating strategy was established accordingly.

5. Flow cytometry

Multicolour flow cytometry data acquisition was done with BD LSRFortessa and BD LSR II cell analysers equipped with BD FACSDiva v6.2 software (BD Biosciences, Germany).

Fluorescence activated cell sorting was performed on a BD FACSARia equipped with BD FACSDiva v5.0.3 software (BD Biosciences, Germany). All samples were maintained at 4°C for the duration of the sort. Three biological replicates were obtained for every tissue, each containing a minimum of 5000 cells (2-5 animals samples were pooled when necessary).

Sort purity was accessed for all samples and only those with purity ≥ 95% were used for RNA extraction. Cells were sorted directly into Buffer RLT (Qiagen, USA) with 1% β-mercaptoethanol (Sigma, UK), disrupted
through vortexing at 3200 rpm for 1 min, and immediately stored at -80°C until further processing.

Flow cytometry data were analysed with FlowJo X v10 (Tree Star, Inc, USA).

6. Sample preparation for gene expression

RNA was extracted using the RNeasy Micro Kit (Qiagen, USA) following the manufacturer’s protocol. RNA yield, quality and integrity were evaluated using the RNA 6000 Pico LabChip kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

For microarray studies, the Ovation Pico WTA System V2 kit (NuGEN, USA) was used to prepare amplified cDNA from total RNA for fragmentation and labeling using the Encore Biotin Module kit (NuGEN, USA), according to kit instructions, and then hybridized onto GeneChip Mouse Gene 2.0 ST arrays (Affymetrix, USA).

For RNA-seq studies, RNA samples were amplified using the SMART-Seq® v4 Ultra® Low Input RNA Kit. Paired-end sequencing libraries were prepared from the amplified cDNA according to the Nextera® XT DNA library prep protocol, and sequenced using an Illumina NextSeq 500 (Illumina, USA) (38 base-paired reads).

Microarray and RNA-seq studies were performed in collaboration with UCL Genomics.
7. Microarray analysis

7.1. Data acquisition and processing

Hybridized arrays were scanned with a GeneChip 3000 7G scanner (Affymetrix, USA) and the image data processed to generate .cel files. Expression Console Software, version 1.4.1 (Affymetrix, USA) was used to generate quality control statistics for each sample and the internal reproducibility was assessed by calculating the Pearson correlation coefficient for the pairs of biological replicates for each tissue (Appendix - Supplementary Figure 1); samples with $r < 0.9$ were considered outliers and were excluded from the analysis (B6→129 model: mesenteric lymph nodes - replicate 1; MataHari T cell model: blood $\alpha 4\beta 7^+$ - replicate 1, bone marrow – replicate 2, gut lamina propria – replicate 1).

Raw sample expression signals were background subtracted, quantile normalized, and the probe level data were summarized using the Robust Multi-array Average algorithm,$^{(240-242)}$ implemented in the oligo BioConductor R package.$^{(243)}$ The ComBat algorithm$^{(244)}$ from the sva BioConductor R package$^{(245)}$ was employed to adjust for batch effects. Transcripts identified through multiple probes were collapsed based on maximum expression values using the CollapseDataset module of GenePattern software (Broad Institute)$^{(246)}$. To enhance statistical testing power, contaminant non-T cell genes were excluded from analysis by filtering out transcripts identified as
tissue-specific (for liver, small intestine and skin) in the PaGenBase database 
(Appendix - Supplementary Table 1)\textsuperscript{(247)}

7.2. Samples relationship visualisation

Multivariate statistical analysis methods implemented in the \textit{stats} R package,\textsuperscript{(248)} in particular principal components analysis and multidimensional scaling, were applied to perform dimensionality reduction of the datasets and visualization of the samples relationships.

Similarities between groups were evaluated by unsupervised hierarchical clustering analysis, using average-linkage criteria and Pearson correlation-based dissimilarity matrix; the validity and stability of the clusters was assess by random sampling with replacement of the genes using non-parametric bootstrap methodology implemented in the \textit{pvclust} CRAN R package.\textsuperscript{(249,250)}

Additional hierarchical clustering and heat maps were produced using the matrix visualization and analysis platform GENE-E (Broad Institute, USA).\textsuperscript{(251)}

7.3. Differential gene expression

The \textit{limma} BioConductor R package\textsuperscript{(252)} was used to perform analyses of gene differential expression, using an empirical Bayes moderated t-statistic,\textsuperscript{(253)} with a cut-off of 0.05, corrected for multiple hypothesis testing using Benjamini-Hochberg procedure and an absolute fold-change cut-off >2.0.
7.4. Gene set enrichment analysis (GSEA)

Gene Set Enrichment Analysis was performed as previously described \(^{(254)}\) using the GSEA software \(^{(255)}\) with the curated gene sets (C2) from the KEGG pathway database (Kyoto Encyclopedia of Genes and Genomes) \(^{(256,257)}\), the gene sets derived from the Biological Process Ontology database (C5) defined by the Gene Ontology Consortium \(^{(258,259)}\), collected in the Molecular Signatures Database (MSigDB v5.1), and the gene sets derived from the modules identified by WGCNA in the MataHari T cell data set.

7.5. Weighted gene co-expression network analysis (WGCNA)

Scale-free network topology analysis of microarray expression data was performed using the WGCNA R package, as previously described \(^{(229,260)}\). A signed hybrid weighted correlation network was constructed using a Pearson correlation matrix created from the pairwise comparison between all pairs of genes, and a soft thresholding power \(\beta = 8\). The topological overlap was calculated as a measure of network interconnectedness, and genes were grouped by average linkage hierarchical clustering on the basis of the topological overlap dissimilarity (1-topological overlap). Module eigengenes were calculated using a dynamic tree-cutting algorithm and merging threshold function at 0.25. The modules identified were correlated to the sample traits using a binary vector representation of the tissues of origin and study groups.
To validate the microarray analysis, preservation of the WGCNA modules identified in the MataHari model dataset was tested against the dataset of the B6→129 model, using the R function “modulePreservation” in the WGCNA R package, as previously described.\(^{(261)}\) Results were interpreted according to the following thresholds for \(Z_{\text{summary}}\): if \(Z_{\text{summary}} \geq 10\), module “strongly preserved”; if \(2 \leq Z_{\text{summary}} < 10\), module “weak to moderately preserved”; if \(Z_{\text{summary}} < 2\), module “not preserved”. Significance of the \(Z_{\text{summary}}\) scores was calculated by permutation analysis.

Intramodular hub genes, which are genes that have the highest number of connections within a module, were identified on the basis of having eigengene–based connectivity (kME) > 0.8 and gene significance (GS) > 0.2. Visualisation of the modules network of gene connections was accomplished with the Cytoscape v3.4 software.\(^{(262)}\)

### 7.6. Gene ontology (GO) overrepresentation analysis

The Web-based Gene Set Analysis Toolkit (WebGestalt),\(^{(263,264)}\) a suite of tools for functional enrichment analysis, was used to identify overrepresented GO annotation categories and translate gene lists into functional profiles. Enrichment of GO terms and associated p-values were calculated based on hypergeometric distribution statistics, adjusting the false discovery rate using the Benjamini-Hochberg procedure.
8. **RNA-seq analysis**

8.1. **Library preprocessing**

FASTQ Toolkit, version 1.0.0 (BaseSpace, Illumina), was used for adapter trimming of the reads in order to eliminate poor quality bases occurring at the ends of reads which could lead to erroneous read alignment. Read alignment was performed using TopHat Alignment, version 1.0.0 (BaseSpace, Illumina) by mapping the RNA-seq reads to the human genome. Transcripts were identified using Cufflinks Assembly & DE, version 2.0.0 (BaseSpace, Illumina), based on the *Homo sapiens* hg38 RefSeq gene annotations.

8.2. **Calculation of read-level metrics**

Gene expression levels (expressed in Reads Per Kilobase per Million, RPKM) were calculated using the Cufflinks Assembly & DE, version 2.0.0 (BaseSpace, Illumina), employing a geometric library normalization method and a fragment bias correction algorithm, which re-scales gene counts to correct for differences in both library sizes and gene length.

8.3. **Single-sample GSEA**

Single-sample GSEA was performed using the ssGSEAPrjorjection (v4) module of GenePattern software (Broad Institute), following the protocol
described by Barbie et al.\textsuperscript{(265)} using the modules identified by WGCNA in the MataHari T cell data set as gene sets. For this purpose the human orthologs for the mouse genes were using the MGI-Mouse Vertebrate Homology Database.

9. **Statistical analysis**

Apart from microarray and RNA-seq data, which were analysed with the aforementioned programs and methodologies, statistical analysis was performed using GraphPad Prism version 6.00 for Mac OsX (GraphPad Software, USA).\textsuperscript{(266)}

Survival curves were created using the Kaplan-Meier method.

The data are expressed as mean ± standard deviation. Significance was assessed using a student’s t-test, 1-way ANOVA or 2-way ANOVA with Sidak’s post hoc test. A p-value of less than 0.05 was taken to indicate a significant difference between groups; only statistically significant differences are marked in the figures. Sample sizes, n, are indicated in the figure legends.
III. Results
1. **Transcription profiles of donor-derived $T_{\text{eff}}$ cells are different between SLO and GVHD target organs**

In order to test the hypothesis that GVHD target tissues exert dominant, idiosyncratic roles in regulating the effector function of alloreactive CD8$^+$ T cells, two murine models of acute GVHD were studied: (1) B6→129 model;\(^{(267,268)}\) (2) MataHari T cell model.\(^{(118)}\) The models used were selected for their clinical relevance, as they are both miHA-mismatched models that reproduce the clinical features and histological abnormalities found in patients with acute GVHD. The former is used as a model of HLA-identical sibling HSCT, and the latter simulates F→M HLA-identical transplant, which is associated with higher risk of GVHD.

In the B6→129 model, the allogeneic immune reaction is mediated by a polyclonal T cell response against multiple antigens.\(^{(267,268)}\) This model mirrors the clinical setting of the majority of the HLA-matched HSCT, and was used to access the diversity of genetic signatures between the different organs. Since it has been described that H60 is an immunodominant miHA in the B6→129 strain combination,\(^{(197)}\) characterisation of the ratio of H60-specific CD8$^+$ T cells was performed for each tissue to evaluate its impact on the variety of transcription profiles.
In order to evaluate the extent to which distinct transcriptional profiles are a result of different TCR repertoire or reflect contrasting tissue antigen distribution, the MataHari T cell model was employed. In this model the donor-recipient mismatch is restricted to a single gender-specific ubiquitously expressed antigen, the male UTY peptide, and GVHD pathology is mediated by a monoclonal population of CD8\(^+\) T cells expressing a transgenic TCR specific for H-2D\(^b\):Uty.\(^{(118)}\)

To ensure a global approach to the characterisation of donor CD8\(^+\) T cell programs during GVHD development, a whole-transcriptome profiling platform was chosen. In contrast with classical 3’ expression microarrays which use the expression at the 3’ end to approximate expression of the entire gene and are thus unable to discriminate between alternatively spliced transcripts that have identical 3’ ends, whole-transcriptome approaches enable the detection not only of the level of expression, but also precisely what is being expressed, including alternative isoforms or genomic deletions. Affymetrix GeneChip® Mouse Gene 2.0 ST Array provides a comprehensive transcriptome coverage identifying over 30000 coding and non-coding transcripts.

In this first results chapter, classical methods of microarray analysis were employed to characterise the broad changes to the transcriptional profile of donor CD8\(^+\) T cells occurring during the development of acute GVHD.
1.1. **Brief characterisation of the minor histocompatibility antigen mismatched models**

**a) B6→129 model**

The B6→129 is a moderate to severe multiple mHA murine GVHD model, achieved by the transplant of C57BL/6 bone marrow and T cells to lethally irradiated 129/Sv recipients ([Chapter II – Materials & Methods, Figure 8](#)). The onset of the manifestations of GVHD was observed at day 3-4 and peaked at day 7 post-transplant ([Figure 11-A](#)). Intestinal involvement was one of the major features of this model, denoted by persistent diarrhoea. Consequently, a significant decrease in body weight was registered in the first week post-transplant ([Figure 11-B](#)), which was never recovered in the BM + T cells group, in clear contrast to the BM only control group. Other manifestations of acute GVHD recorded were hair ruffling, back hunching and lowered mobility. The high early mortality observed ([Figure 11-C](#)) reflects the severity and the rapid progression of the disease, which required the euthanasia of over half of the BM + T cell recipients by the end of the first week post-transplant.

Based on the GVHD kinetics, day 6 post-transplant was elected as the ideal time point for the collection of tissues samples for isolation of infiltrating lymphocytes. In accordance with the clinical features, the GVHD target organ with the highest proportion of donor derived CD8+ T cells from infiltrating
lymphocytes was the small intestine, in particular in the intraepithelial compartment, followed by the skin and the liver (Figure 12).

It has been reported that in a wide variety of multiple mHA mismatch GVHD models, the cytotoxic effectors generated were specific for a small fraction of the non-H-2 H antigens.\(^{(269)}\) Choi and colleagues\(^{(197,270)}\) have demonstrated that in the B6→129 model the H60 antigen is immunodominant, accounting for the specificity of a fifth of the expanded donor CD8\(^+\) T cells. To better characterise this effect and assess its impact on the diversity of the alloreactive immune response at each site, samples of lymphocytes isolated from the lymph nodes, spleen, blood, liver, small intestine and skin were analysed by flow cytometry using PE-conjugated pentamers for H60/H-2K\(^b\) complexes (Figure 13). Overall, at day 6 post-transplant, the frequency of H60 pentamer positive donor derived CD8\(^+\) T cells was less then 7%, with the exception of the gut lamina propria and the blood where it amounted to 9.4% and 11.7%, respectively (Figure 13-B and C). These results show that although H60 is a dominant mHA, it did not preclude the generation of CTL against other mHA; thus, gene expression profiling of such samples should still reflect the diversity of a polyclonal immune reaction against an array of antigens.

b) MataHari T cell model

The MataHari T cell model was characterised by Dr. Thomas Conlan, as part of his PhD project.\(^{(217)}\) Briefly, this is a model of a female into male mHA mismatch HSCT, in which a monoclonal population of transgenic CD8\(^+\) T cells that recognise the D\(^b\):Uty peptide (MataHari T cells) are transferred together
with bone marrow and polyclonal CD4\(^+\) T cells from wild-type C57BL/6 female into irradiated wild-type male C57BL/6 recipients (Chapter II – Materials & Methods, Figure 9). Since the HY antigen is a ubiquitously expressed male-specific antigen, the acute GVHD immune reaction only develops in the male recipients.

This model has been described as of mild to moderate severity, affecting predominantly the skin, lung, small intestine and the liver (Figure 14). Similarly to the B6→129 model, the onset of the clinical signs of GVHD is at day 4-5 post-transplant, and it progresses gradually for 2-3 weeks (Figure 15-A and B), with 50% mortality at 4 weeks post-transplant (Figure 15-C).

To maintain uniformity within the study, samples were collected at day 7 post-transplant. Interestingly, even though at this stage the clinical manifestations of acute GVHD were less severe than those observed in the B6→129 model, the majority of the organs showed a greater lymphocyte infiltration (Figure 16).
A systems immunology approach to GVHD

Chapter III

Figure 11. Characterisation of the B6→129 model: (A) GVHD score; (B) percentage of body weight change; (C) survival. Evaluation of the GVHD severity was performed by assessment of the subjects’ global wellbeing through a standardised clinical score (A), monitoring of the body weight (B), and appraisal of the mortality (C). Plots A and B represent the mean values and standard deviations measured at specific time-points over 7 weeks; plot C is a Kaplan-Meyer survival plot showing the survival rates over time (blue line: BM only recipient controls, n=6; red line: BM + T cells recipients, n=16). The B6→129 is a moderate to severe GVHD model, with an early onset of GVHD manifestations and high lethality after the first week post-transplant. BM, bone marrow.
Figure 12. Donor CD8⁺ T cell infiltration of the secondary lymphoid organs and GVHD target organs at day 6 post-transplant in the B6→129 model. (A) Gating strategy used to identify the donor CD8⁺ T cells: exclusion of doublets → exclusion of propidium iodide positive cells → exclusion of stroma cells → morphologic lymphocyte selection → donor CD8⁺ T cells selection based on expression of congenic markers. (B) Representative plots showing the percentage from lymphocytes of donor CD8⁺ T cells in each organ. (C) Summary data: mean values and standard deviations of the percentage from lymphocytes of donor CD8⁺ T cells are plotted for each tissue (n=4-11; data pooled from 4 independent experiments). LN, lymph nodes; LP, lamina propria; IEL, intraepithelial lymphocytes.
Figure 13. Frequency of H60-specific from tissue infiltrating donor CD8⁺ T cells in the B6→129 model. (A) Gating strategy used to identify the donor CD8⁺ T cells: exclusion of doublets → exclusion of propidium iodide positive cells → exclusion of stroma cells → morphologic lymphocyte selection → donor CD8⁺ T cells selection based on expression of congenic markers. (B) Representative plots showing the percentage H60-pentamer⁺ cells from donor CD8⁺ T cells, in each organ. (C) Summary data: mean values and standard deviations of the percentage of H60-pentamer⁺ cells from donor CD8⁺ T cells are plotted for each tissue (n=4-8; data pooled from 2 independent experiments). Although H60 behaves as a dominant mHA, it accounts for less that 10% of the diversity of donor CD8⁺ T cells that infiltrate most of the tissues. LN, lymph nodes; LP, lamina propria; IEL, intraepithelial lymphocytes.
Figure 14. Kinetics of MataHari T cell infiltration of the GVHD target organs.

*Courtesy of Dr Thomas Conlan.*\(^{(217)}\)
Figure 15. Characterisation of the MataHari T cell model: (A) GVHD score; (B) percentage of body weight change; (C) survival. General wellbeing of transplanted animals assessed by a standardised clinical score (A), body weight (B), and mortality (C) records were used to determine GVHD severity. Plots A and B represent the mean values and standard deviations measured at specific time-points over 7 weeks; plot C is a Kaplan-Meyer survival plot showing the survival rates over time (blue line: BM only recipient controls, n=3; red line: BM + T cells recipients, n=8). The MataHari T cell model is a mild to moderate GVHD model, when compared to the B6→129 model, characterised by a gradual progression of clinical manifestations over the course of 3 - 4 weeks.

*Courtesy of Dr Thomas Conlan.*

(217)
Figure 16. Donor CD8⁺ (MataHari) T cell infiltration of the secondary lymphoid organs and GVHD target organs at day 7 post-transplant. (A) Gating strategy used to identify the donor CD8⁺ T cells: exclusion of doublets → exclusion of propidium iodide positive cells → exclusion of stroma cells → morphologic lymphocyte selection → donor CD8⁺ T cells selection based on expression of congenic markers. (B) Representative plots showing the percentage from lymphocytes of donor CD8⁺ T cells in each organ. (C) Summary data: mean values and standard deviations of the percentage from lymphocytes of donor CD8⁺ T cells are plotted for each tissue (n=6; data pooled from 2 independent experiments). Organ infiltration by donor CD8⁺ T cell exhibits a similar pattern to that of the B6→129 model. LN, lymph nodes; BM, bone marrow; LP, lamina propria; IEL, intraepithelial lymphocytes.
1.2. Purification and microarray analysis of donor derived CD8\(^+\) T cells

All samples used in this project were collected at day 6 (for the B6 → 129 model) or day 7 (for the MataHari T cell model) post-transplant, processed and stored using optimized protocols designed to standardize the cell isolation methodology, reduce the processing time and maximize the cell sorting purity, in order to reproducibly yield sufficient quantities of cells to ultimately generate the required amount of high quality RNA.

To obtain precise expression profiles, positive selection of viable donor derived CD8\(^+\) T cells was performed through FACS sorting. Examples of the gating strategy employed for cell sorting in each model are illustrated in Figure 17: (1) exclusion of doublets, (2) exclusion of propidium iodide positive cells, (3) exclusion of stroma cells, (4) morphologic lymphocyte selection, (5) selection of donor CD8\(^+\) T cells based on the expression of congenic markers, (6) confirmation of CD8 positivity.

In order to ensure the comparability between groups, in particular in the secondary lymphoid organ samples, the donor CD8\(^+\) T cells were CFSE labelled prior to infusion into the recipients (both syngeneic and allogeneic). By only sorting CFSE low cells it was ensured that the microarray analysis was
restricted to cells that had undergone equivalent number of divisions (Figure 18-A).

In respect to the blood samples, staining for the α4β7 integrin (LPAM-1) permitted the differential sorting of cells imprinted for gut homing (CFSE<sub>low</sub>LPAM-1<sup>high</sup> vs CFSE<sub>low</sub>LPAM-1<sup>low</sup>) (Figure 18-B).

Purity of isolated populations was measured as the percentage CD8<sup>+</sup> cells bearing the donor congenic markers (Figure 19-A), which in both models was greater than 98% (Figure 19-B). Cell purification procedures were conducted at 4°C to minimize in vitro-induced changes in gene expression.

RNA samples extracted from the different groups of samples were labelled and hybridised to the Affymetrix GeneChip<sup>®</sup> Mouse Gene 2.0 ST array.
**B6 → 129 model**

**MataHari T cell model**

**Figure 17.** Strategy for high-purity FACS sorting of donor CD8⁺ T cells. Gating sequence used to select donor derived CD8⁺ T cells in each GVHD model: exclusion of doublets → exclusion of propidium iodide positive cells → exclusion of stroma cells → morphologic lymphocyte selection → selection of donor CD8⁺ T cells based on the expression of congenic markers → confirmation of CD8 positivity.

**Figure 18.** Staining strategy to isolate cells with equivalent number of divisions (A) and homing imprinting (B). (A) To ensure the comparability between groups, CSFE labelling of the CD8⁺ T cells prior to infusion into the recipients was used as a means to identify cells that had undergone similar number of divisions. The plot is a representative example of the gating strategy followed to exclusively select cells that had diluted the dye (grey area). CD8⁺ T cells collected from SLO of synBMT recipients displayed a bimodal staining (blue curve) in contrast to those isolated from alloBMT recipients (red curve) which were uniformly CFSE negative by day 7 post transplant. (B) Donor CD8⁺ T cells from the blood were subdivided into CFSE<sup>low</sup>LPAM-1<sup>high</sup> (cells imprinted for gut homing) and CFSE<sup>low</sup>LPAM-1<sup>low</sup> cells.
Figure 19. Sort quality check and evaluation of sample purity. (A) Gating strategy used to determine the purity of the isolated populations (viable cells, positive for CD8 and for the donor congenic markers). (B) Purity of the isolated populations (percentage of viable CD8⁺ cells bearing the donor congenic markers) selected for the RNA extraction used in the microarray study (B6→129 model: n=18; MataHari T cell model: n=45).
1.3. Exploratory data analysis

Unsupervised data analyses consist of exploratory techniques that do not require the incorporation of any prior knowledge and aim to identify groups with similar behaviours (i.e. expression profiles).\(^{(271)}\) Some commonly used mathematical tools that characterise the structure of the data include principal component analysis (PCA) and multidimensional scaling (MDS) for dimensionality reduction,\(^{(272,273)}\) as well as hierarchical clustering for relationship discovery.\(^{(274,275)}\)

The advantage of performing PCA on microarray data sets is that it allows visual assessment of the similarities and differences between samples in 2D or 3D plots, by virtue of reducing the dimensionality of the data while retaining most of its variation.\(^{(276)}\) This reduction is accomplished through the identification of the directions (principal components) along which the variation in the data is maximal, so that each sample can be represented by relatively few values instead of by thousands of variables.\(^{(276)}\)

a) B6→129 model

The two principle components of the gene expression profiles of the B6→129 data sets are plotted in Figure 20. The samples have been colour coded to reflect the three study groups (samples from healthy donors in light blue, syngeneic BMT recipient controls in dark blue and allogeneic BMT study subjects in red). Analysis of the bi-plots showed a greater diversity of gene
expression profiles in the alloBMT setting when compared to the syngeneic BMT controls or the donor samples, consistent with the underlying major differences in cellular developmental stage: donor – naïve; synBMT – antigen-inexperienced T cells undergoing homeostasis-driven proliferation; alloBMT – antigen-experienced effector T cells.

Clustering algorithms were used to establish groups that share similar characteristics. The result of these techniques can be depicted as dendrograms, tree like structures in which the nodes represent subsets of an expression data set joined to form groups according to their distance or similarity. As illustrated in Figure 21, hierarchical clustering of the samples using Pearson’s correlation as the dissimilarity distance metric with an average-linkage clustering method confirmed the existence of a great dissimilarity between the gene expression profiles of the alloBMT samples and the other groups. Moreover, within the alloBMT groups, a dichotomy between SLO and the GVHD target-organs was revealed. Evaluation of the validity and stability of the clusters using non-parametric bootstrap methodology supported these observations.

In the interest of resolving the similarities and dissimilarities within the alloBMT group and visualise the interrelationships between samples, multidimensional scaling was performed, as it conveys a superior fidelity in the representation of the distances between different cases, particularly for high-dimensional geometric objects. From Figure 22, it is clear the presence of a rift between the gene expression of T cells in the SLO and in the GVHD target organs.
b) MataHari T cell model

To evaluate if the diversity in transcriptional profiles observed in the B6→129 model could be due to a segregation of subpopulations of T cells according to their TCR repertoire, merely reflecting an asymmetric tissue antigen distribution, the same analysis was performed in the MataHari T cell model. If differences in antigen distribution were the main driver of the dissimilarities in gene expression among donor T cells isolated from different organs, it would be expected that in the MataHari T cell model the diversity in transcriptional profiles would be greatly reduced, as the antigen recognised by the MataHari T cells is uniformly expressed in all tissues.

The results, however, do not support this hypothesis. Both the hierarchical clustering (Figure 23) and the multidimensional scaling analysis (Figure 24) show a similar pattern to that observed in the B6→129 model. Although, in this model, the donor CD8⁺ T cell population is monoclonal, there still exists a clear difference between the transcription profiles of cells isolated from the SLO and the GVHD target organs. Interestingly, when assessing the variability within these two compartments, it appears that in the B6→129 model the individual tissue clusters are more discrete than in the MataHari T cell model, suggesting that, even though TCR repertoire variety is not required for the establishment of a SLO - GVHD target organ dichotomy, it could contribute to enhance the differences between tissues.

To further explore this concept, a direct comparison between the 2 data sets was performed, after correction for batch effects with the ‘ComBat’ function
of the ‘sva’ package for R.\(^{(245)}\) **Figure 25** depicts the average linkage agglomerative hierarchical clustering of all samples, with the underlying similarity matrix showing the individual Pearson’s correlation coefficients colour coded for easier interpretation. As can be observed, the separation between the three groups (donor + synBMT vs alloBMT – SLO vs alloBMT – GVHD target organs) is maintained, and the clustering of the samples appears to reflect the source tissue, rather than the GVHD model. This result validates the previous observations and confirms that the transcriptional diversity observed was largely independent of the TCR repertoire and antigen distribution.
Figure 20. Principal components analysis of the B6→129 data set. The plot shows the 2D projection of the first two principal components; samples were colour coded according to study groups (light blue: donors; dark blue: syngeneic BMT; red: allogeneic BMT). It is noteworthy a greater dispersion of the points in the alloBMT group when compared to the syngeneic BMT controls or the donor samples. PC1, 1st principal component; PC2, 2nd principal component.
**Figure 21.** Hierarchical clustering of the B6→129 data set, validated through non-parametric bootstrap methodology. Groups represent the average expression of the biological replicates. Three clusters are identified (red boxes), separating the donor + synBMT, from the alloBMT secondary lymphoid organs (SLO), from the alloBMT GVHD target organs. The approximate unbiased p-values (AU, in red) and the bootstrap probability (BP, in green) are shown above each node (identified by the grey numbers).
Figure 22. Multidimensional scaling analysis of the alloBMT samples from the B6→129 data set. A rift between the SLO and the GVHD target organs is apparent. CMD1, 1st coordinate of multidimensional scaling; CMD2, 2nd coordinate of multidimensional scaling; LN, lymph nodes.
Figure 23. Hierarchical clustering of the MataHari T cell data set, validated through non-parametric bootstrap methodology. Groups represent the average expression of the biological replicates. Similarly to the B6→129 data set, three clusters are identified (red boxes), separating the donor + synBMT, from the alloBMT secondary lymphoid organs (SLO), from the alloBMT GVHD target organs. The approximate unbiased p-values (AU, in red) and the bootstrap probability (BP, in green) are shown above each node (identified by the grey numbers).
Figure 24. Multidimensional scaling analysis of the alloBMT samples from the MataHari T cell data set. The dichotomy between SLO and GVHD target organs is maintained, however in the MataHari T cell model the difference is less pronounced with the liver samples being closer to the SLO than to the remaining GVHD target organs. CMD1, 1st coordinate of multidimensional scaling; CMD2, 2nd coordinate of multidimensional scaling; LN, lymph nodes.
Figure 25. Transcriptional profiles are independent of the TCR repertoire and do not reflect antigen distribution. Average linkage agglomerative hierarchical clustering of all samples from both experimental models and similarity matrix with colour coded Pearson’s correlation coefficients shows that sample clustering reflects the study groups (light grey: donor; dark grey: synBMT; black: alloBMT) and the tissues of origin, irrespective of GVHD model (circles: B6→129; squares: MataHari T cell).
1.4. SLO - GVHD target organ dichotomy reflects disparities in differentiation status of donor $T_{eff}$ cells

The magnitude of the divergence in gene expression between the SLO and the GVHD target organs was further evaluated in both models by performing gene set enrichment analysis (GSEA) comparing the transcriptional profiles of these two groups. GSEA is a method that identifies biologically meaningful, even if subtle, differences in gene expression, through the analysis of coordinated changes in sets of functionally related genes.\(^{(254)}\)

**Figure 26** summarises the results from GSEA comparing the GVHD target organs to the SLO in the B6$\rightarrow$129 model, using the the C5:BP (Gene Ontology – Biological Process) collection from the Molecular Signatures Database (MSigDB). The Gene Ontology (GO) terms enriched in the GVHD target organs were organised into an enrichment map using Cytoscape “Enrichment Map” plug-in.\(^{(281)}\) This network representation of the data clusters together mutually overlapping gene-sets, where the node size represents the number of genes in the gene-set; edge thickness is proportional to the overlap between gene-sets; and the enrichment p-value is mapped to the node colour as a white-to-red gradient. Analysis of the GO enrichment map reveals that the genes differentially expressed in the GVHD target organs in comparison to the SLO are involved mainly in processes of transcription and protein metabolism; signal transduction; cell component transport and localisation; regulation of cell
cycle, proliferation, development and apoptosis; and are linked to the establishment of the immune reaction.

In order to gain more insight into the pathways up-regulated in the GVHD target organs, a pair-wise comparison of the SLO and GVHD target organs was performed using the C2:CP:KEGG collection (Curated canonical pathways – KEGG PATHWAY), both on the B6→129 and the MataHari T cell models. A total of 154 pathways were identified as being enriched in the GVHD target organs, with an extensive overlap between the two models: 87.2% of pathways found in the MataHari T cell model were also present in the B6→129 model (Figure 27). The top 10 pathways common to both models, listed in Table 12, reflect a clear enrichment for pathways related to the effector mechanisms of adaptive immunity and to alloreactivity.

Moreover, the analysis of the expression profiles of a selection of genes frequently used to characterise the stages of T cell development\(^{(282-286)}\) (Figure 28, data shown only for the B6→129 model) revealed a coordinated change in the expression of a panoply of markers associated with CD8\(^+\) T cell effector function consistent with the transition from a naïve phenotype in the donor + synBMT group, to a post-TCR engagement activation status in the alloBMT group, which ranged from early to late effector phenotypes in the SLO and GVHD target organs, respectively. A similar trend was observed when the differentiation status of the donor CD8\(^+\) T cells in each tissue was evaluated by calculating the enrichment for published genetic signatures associated with Tc1/Tc17 polarisation,\(^{(287)}\) pathogenicity,\(^{(288)}\) and cytokine or effector response,\(^{(285)}\) using single-sample GSEA (Figure 29). Additionally, this
analysis revealed a wide range of enrichment scores across tissues and a dissonance between the polarisation of the T cells and their cytotoxicity; while the Tc1 and Tc17 transcriptional signatures are upregulated early on in the SLO, the maximum expression of effector molecules associated with tissue injury is only achieved in the GVHD target organs.
Figure 26. Transcriptional differences between SLO and GVHD target organs are biologically meaningful. Network visualization of GSEA based on the differentially expressed genes in the GVHD target organs vs the SLO, using EnrichmentMap (B6→129 data set). Enriched GO terms are depicted by red nodes, where colour represents the corresponding FDR-adjusted enrichment P-value (q-value), size is proportional to the number of genes in each node, edge thickness indicates overlap of genes between nodes, and the theme of each clusters is specified.
Figure 27. Venn diagram representing the overlap of enriched gene sets from the KEGG PATHWAY collection in the GVHD target organs vs the SLO, found in the B6→129 (circle) and the MataHari T cell (square) models. An extensive overlap between the two models is observed.

Table 12. Top 10 pathways enriched in GVHD target organs vs SLO common to both data sets (MSigDB C2:CP:KEGG collection).

<table>
<thead>
<tr>
<th>Pathway name</th>
<th>Size</th>
<th>NOM p-val</th>
<th>FDR q-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>T cell receptor signalling pathway</td>
<td>104</td>
<td>&lt;0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Endocytosis</td>
<td>171</td>
<td>&lt;0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>B cell receptor signalling pathway</td>
<td>71</td>
<td>&lt;0.0001</td>
<td>0.0016</td>
</tr>
<tr>
<td>MAPK signalling pathway</td>
<td>255</td>
<td>&lt;0.0001</td>
<td>0.0043</td>
</tr>
<tr>
<td>Natural killer cell mediated cytotoxicity</td>
<td>96</td>
<td>&lt;0.0001</td>
<td>0.0066</td>
</tr>
<tr>
<td>Chemokine signalling pathway</td>
<td>162</td>
<td>&lt;0.0001</td>
<td>0.0069</td>
</tr>
<tr>
<td>Lysosome</td>
<td>114</td>
<td>&lt;0.0001</td>
<td>0.0074</td>
</tr>
<tr>
<td>Cytokine-cytokine receptor interaction</td>
<td>225</td>
<td>&lt;0.0001</td>
<td>0.0220</td>
</tr>
<tr>
<td>Graft-versus-host disease</td>
<td>27</td>
<td>0.0051</td>
<td>0.0223</td>
</tr>
<tr>
<td>Allograft rejection</td>
<td>29</td>
<td>0.0056</td>
<td>0.0224</td>
</tr>
</tbody>
</table>
Figure 28. Heat map of selected genes commonly used to characterise T cell differentiation, including: transcription factors, chemokine/cytokine receptors, signalling molecules, integrins, activation markers, coregulatory molecules, and effector function molecules. Note the shift from a naïve phenotype in the donor + synBMT group, to an early activated status in the alloBMT SLO and a late effector phenotype in the GVHD target organs. Data from the B6→129 model. Sp, spleen; LN, lymph nodes; Bl, blood; Li, liver; SI, small intestine; Sk, skin.
Figure 29. Comparison of the differentiation status of donor CD8\(^+\) T cells across the tissue. Single-sample GSEA was used to determine the enrichment for gene sets associated with Tc1/Tc17 polarisation, pathogenicity, and cytokine or effector function.
1.5. Transcriptome diversity is neither due to early trafficking of effectors to the peripheral tissues nor to differential homing imprinting of donor-derived T cells

One possible explanation for the differences in transcriptional profiles described could be that at the time point the samples were collected the effector cells have already migrated out of the SLO towards the peripheral tissues, and those that remain in the LN and spleen represent the subpopulations of T cells that were either ineffectively primed or that are still going through the process of activation.

To test this hypothesis, a second set of experiments were performed using the MataHari model in which the HSCT recipients were treated with FTY720 (a sphingosine-1-phosphate antagonist) to trap T cells in the SLO (Figure 30). By starting the treatment at D+3 post-transplant (a time point before MataHari T cells are easily detectable in blood or GVHD organs) the lymphocyte egress from the lymph nodes was effectively blocked, as demonstrated by the marked reduction in circulating donor CD8+ T_eff cells in FTY720 treated animals in comparison to the controls (Figure 31).

Analysis of the transcriptional profiles of donor CD8+ T_eff cells isolated from the lymph nodes prior and post-treatment revealed that at D+3 donor T cells were in fact closer to the profiles of naïve input T cells, while at D+7 there was no significant difference between FTY720-treated subjects and
PBS-treated controls (Figure 32). The data therefore refute the hypothesis that the differences in gene expression between $T_{\text{eff}}$ cells isolated from SLO or GVHD target organs can be due to early egress of optimally activated $T_{\text{eff}}$ cells.

It has been demonstrated that tissue tropism imprinting is required for the induction and maintenance of acute GVHD. As such, the averaging of gene expression across several tissues could potentially conceal the convergence of transcriptional programs of $T_{\text{eff}}$ cells from anatomically related SLO and GVHD target organs. To investigate this, LN and blood samples from the MataHari T cell data set were subdivided into their two major constituents (mesenteric and peripheral LN; LPAM-1$^{\text{high}}$ and LPAM-1$^{\text{low}}$ blood T cells) and the transcriptional profiles of $T_{\text{eff}}$ cells derived from gut and skin, their respective draining LN, and GVHD target organ-imprinted peripheral blood subsets were compared. As depicted in Figure 33, the segregation SLO and GVHD target organ $T_{\text{eff}}$ cell profiles was maintained, with no evidence that local imprinting was linked to greater skin or gut-like effector differentiation.
Figure 30. Experimental design to evaluate the effect of SLO lymphocyte egress blockage on the transcriptional profile of donor CD8$^+$ T$_{eff}$ cells in the lymph nodes. BM, bone marrow; LN, lymph nodes.
Figure 31. Treatment with FTY720 blocks lymphocyte egress from the SLO. Blood samples from FTY720- and PBS-treated animals were collected at D+7 post-transplant. A marked decrease in circulating lymphocytes is observed, affecting particularly T cells (4.8 fold reduction in Thy-1.2⁺ cells vs 1.4 fold reduction in CD19⁺ cells), while neutrophil (Ly6G-Gr1⁺ cells) counts remain fairly unaffected.
Figure 32. Differences in gene expression between $T_{\text{eff}}$ cells are not due to early egress of optimally activated $T_{\text{eff}}$ cells. Multidimensional scaling analysis of the transcriptional profiles of donor CD8$^+$ T cells isolated from the LN at D+3 and D+7 post-transplant in animals treated with FTY720 or PBS showing that (1) at D+3 donor T cells in the lymph nodes are very similar to those from the donor and syngeneic BMT controls; (2) at D+7 there are no significant differences between FTY720- and PBS-treated samples. The data suggests that trafficking of effector cells to the periphery does not determine the differences observed between SLO and GVHD target organs. CMD1, 1st coordinate of multidimensional scaling; CMD2, 2nd coordinate of multidimensional scaling.
Figure 33. Diversity of transcription profiles is not explained by differences in homing imprinting of donor T cells. Multidimensional scaling analysis of the transcriptional profiles of donor CD8\(^+\) T\(_{\text{eff}}\) cells isolated from the gut and skin, their respective draining LN, and GVHD target organ-tropic peripheral blood subsets showed no preferential clustering of samples from anatomically related sources (mesenteric LN / LPAM-1\(_{\text{high}}\) blood subset / gut vs peripheral LN / LPAM-1\(_{\text{low}}\) / skin). LN, lymph nodes; CMD1, 1\(^{\text{st}}\) coordinate of multidimensional scaling; CMD2, 2\(^{\text{nd}}\) coordinate of multidimensional scaling.
1.6. Discussion

In this project two murine models of miHA-mismatch HSCT were chosen to simulate the most frequent type of HSCT in the clinic, the HLA-matched HSCT (B6→129 model), and the common F→M miHA mismatch associated with increased risk of GVHD development (MataHari T cell model). The first stage of animal model based translational research involves the validation of the experimental models used to characterise the incidence, severity and pathologic evidence of GVHD. Characterisation of the B6→129 model was performed in the beginning of the project in collaboration with Dr Thomas Conlan who also characterised the MataHari T cell model for his own PhD project. Both models were shown to reproduce the clinical features of acute GVHD, with similar kinetics and severity (Figure 11 and Figure 15), and with the same pattern of organ involvement (Figure 12 and Figure 16).

Previous studies have identified H60 as a dominant miHA in the B6→129 combination, inducing expansion of H60-specific CD8⁺ T cell clones that can amount up to ~20% of the total host CD8⁺ T cell population by D+10 post-HSCT. During the characterisation of the model, the contribution of the H60-specific CD8⁺ T cell subset was evaluated and it was determined that, although it represents a dominant population, at D+6 its frequency is usually less than half of that reported in the literature for D+10, and no significant difference was observed among the different organs (Figure 13). It is therefore
very unlikely that the dominance of the anti-H60 alloreactive response may reduce the diversity in transcription profiles in this model.

Indeed, exploratory analysis of the data set showed a great diversity in transcriptional profiles in the alloBMT group with a clear distinction between the SLO and the GVHD target organs. Moreover, in the GVHD target organs two main clusters were identified, one comprising the liver and the gut samples, the other containing the skin samples (Figure 21 and Figure 22). The same pattern was also observed in the MataHari T cell model, with significant diversity in the transcriptional profiles of donor T<sub>eff</sub> cells derived from the individual organs (Figure 23 and Figure 24).

Remarkably, although there are significant differences between the B6→129 and the MataHari T cell models (in the former the donor T cell population is polyclonal and respond to a multitude of miHA antigens, while in the latter the donor alloreactive response is mediated by a monoclonal population that recognise a single antigen), direct comparison of the transcriptional profiles of all the samples from both data sets revealed a great degree of correspondence between them (Figure 25). It is noteworthy that the samples clustered according to the experimental groups (donor, synBMT, alloBMT) and organ of origin rather than the experimental model, implying that the differences among tissues are more prominent at driving gene expression diversity than dissimilarities between experimental models. The consistent divergence between SLO and GVHD in both the polyclonal and monoclonal systems further supports the conclusion that this effect is independent of TCR repertoire and antigen distribution.
An interesting difference between the two models is the behaviour of the samples derived from the liver. While in the B6→129 model the liver samples segregate with the gut, in the MataHari T cell model their transcriptional profile appears to be closer to that of the SLO. Although in both models there is an extensive lymphocyte infiltration of the liver during the development of GVHD, differences in T_{eff} cell transcriptional profiles might be associated with strain specific factors that determine susceptibility to liver inflammation.\(^{292}\)

Analysis of the differentially expressed genes involved in the SLO-GVHD target organ dichotomy revealed fundamental differences in cellular programs, suggesting a greater level of effector differentiation in the GVHD target tissues (Figure 26 and Table 12). This unexpected level of heterogeneity in T_{eff} cell differentiation according to location was also evident from the pattern of expression of effector genes (Figure 28), which suggested a preferential Tc1/Tc17 differentiation program in the small intestine and in the skin T_{eff} cells, contrasting to Tc2 differentiation in the liver.

Lastly, the blockage of lymphocyte egress from the SLO and the comparison of the transcription profiles of GVHD target tissues with the corresponding draining lymph nodes and blood subset populations refuted the hypothesis that the tissue-associated variability in gene expression could be explained by early trafficking of effector cells or by differential homing imprinting of donor T cells.

Altogether, the data suggest that the diversity in transcriptional profiles of donor CD8\(^{+}\) T_{eff} cells observed after an allogeneic BMT, at the peak of the alloreactive immune reaction, does not reflect differences in antigen expression
between tissues nor an imbalanced distribution of particular clonal populations. The divide between the SLO and the GVHD target organs seems to be indicative of meaningful disparities in biological processes occurring at each site, and denotes the existence of a bias in the functional status of donor CD8$^+$ T$_{eff}$ cells which is dependent on the tissues they infiltrate.

Dissection of the data set and identification of biologically meaningful tissue-specific transcriptional signatures that may be applicable to other models of GVHD will require the application of unbiased methods of systems biology, which will be addressed in Section 2 of the results chapter.
2. Teff cell transcription profiles show compartmental specificity in acute GVHD

To obtain a more global view of the transcriptional changes in the donor CD8\(^+\) T cell response in the context of acute GVHD development, weighted gene correlation network analysis (WGCNA) was used to identify differences in the gene expression patterns at each site.

WGCNA has been described as "a step-wise data reduction technique, which (a) starts from the level of thousands of variables, (b) identifies biologically interesting modules based on a node significance measure, (c) represents the modules by their centroids (e.g., eigenvectors or intramodular hubs), (d) uses intramodular connectivity as quantitative measures of module membership, and (e) combines node significance and module membership measures for identifying significant hub nodes".\(^{(293)}\)

The discovery of highly correlated clusters of genes in an unsupervised manner is considered to be extremely useful computational approach in identifying modules of both co-expressed and co-regulated genes,\(^{(294-296)}\) which frequently represent conserved molecular “circuits” responsible for specific cellular functions.\(^{(297)}\)
2.1. Identification of gene clusters based on correlation and co-expression across tissues using WGCNA

In order to avoid any potential bias related to potential heterogeneity of the CD8\(^{+}\) T cell populations isolated from the different sites, the WGCNA algorithm was applied on the data set derived from the MataHari T cell model after exclusion of outlier samples and of tissue specific genes, as defined in the PaGenBase database.\(^{(247)}\)

The gene correlation network was constructed based on the pairwise Pearson correlations between all genes, by raising the co-expression similarity to a power \(\beta\) chosen using the scale-free topology criterion, while retaining the information on the correlation signs between genes. The soft threshold 8 was selected as being the smallest value of \(\beta\) with which approximate scale free topology is reached (Figure 34).

The correlation expression values were further used to construct a topological overlap matrix (TOM), which measured the proximity between neighbouring genes. Hierarchical clustering of the genes was performed using the dissimilarity values “1-TOM” as the distance measure. Modules were defined using a dynamic tree-cutting algorithm, merging those with a correlation above 0.75 (Figure 35). All uncorrelated genes were assigned to a grey module.
A total of 31 highly correlated modules of genes were identified (Table 13). To study the relationships among the found modules, the first principal component of each module (eigengenes) was determined and used as its representative profile. The modules similarity was quantified by eigengene correlation and displayed as a dendrogram and a heat map (Figure 36).

Three major groups of correlated eigengenes, meta-modules, can be recognised: (A) comprising modules 1 – 9; (B) comprising modules 10 – 19; and (C) comprising modules 20 – 31.

To understand the biological meaning of these findings, the correlation between the modules and the sample traits (experimental groups: donor, synBMT, alloBMT; alloBMT tissue groups: SLO, GVHD target organs; individual tissues: lymph nodes, spleen, bone, blood, liver, gut - IEL, gut - LP, skin - dermis, skin - epidermis) was performed (Figure 37).

Indeed, a positive correlation between the meta-modules and the experimental groups was observed: meta-module B related to donor the synBMT samples, while modules A and C related to alloBMT samples (SLO and GVHD target organs, respectively). Moreover, unique patterns of correlation were distinguishable at the tissue level; the modules with the highest significance ($p$-value $\leq 0.05$) being:

- lymph nodes – modules 5, 8 and 17;
- spleen – modules 2, 3, 4 and 6;
- blood – modules 1 and 20;
• bone marrow – modules 1, 3 and 4;
• liver – module 4;
• gut (LP) – modules 9, 13 and 27;
• gut (IEL) – modules 6, 29, 30 and 31;
• skin (dermis) – modules 22, 23, 25, 26 and 29;
• skin (epidermis) – modules 26, 28 and 29.

To better visualise the relationships between the modules, a eigengene network was generated using the modules as nodes and the inter-module adjacency measure as edges (Figure 38). This representation illustrates the diversity of the transcriptional signatures, highlighting the similarity between donor and the synBMT transcriptional profiles (green and blue) and the dichotomy between the alloBMT SLO (yellow) and GVHD target organs (orange) gene expression.
Figure 34. Selection of the power $\beta$ for gene correlation network generation. (A) Plot showing the scale-free fit index as a function of the soft-thresholding power. Soft threshold 8 is the lowest power to achieve a scale-free fit index above 0.90 (red line). (B) Log-log plot of the network connectivity using soft threshold power 8. The approximate straight line relationship (high $R^2$ value) shows approximate scale free topology.
Figure 35. Dendrogram obtained by hierarchical clustering of genes based on topological overlap, showing the assigned modules colour-coded.

Table 13. WGCNA modules.

<table>
<thead>
<tr>
<th>Mod.</th>
<th>Colour code</th>
<th>No. of genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>green</td>
<td>816</td>
</tr>
<tr>
<td>2</td>
<td>brown4</td>
<td>1113</td>
</tr>
<tr>
<td>3</td>
<td>lightcyan</td>
<td>615</td>
</tr>
<tr>
<td>4</td>
<td>lightcyan1</td>
<td>1569</td>
</tr>
<tr>
<td>5</td>
<td>skyblue3</td>
<td>117</td>
</tr>
<tr>
<td>6</td>
<td>floralwhite</td>
<td>151</td>
</tr>
<tr>
<td>7</td>
<td>orange</td>
<td>298</td>
</tr>
<tr>
<td>8</td>
<td>ivory</td>
<td>90</td>
</tr>
<tr>
<td>9</td>
<td>palevioletred3</td>
<td>327</td>
</tr>
<tr>
<td>10</td>
<td>darkorange</td>
<td>481</td>
</tr>
<tr>
<td>11</td>
<td>darkolivegreen</td>
<td>134</td>
</tr>
<tr>
<td>12</td>
<td>grey60</td>
<td>283</td>
</tr>
<tr>
<td>13</td>
<td>lightsteelblue1</td>
<td>103</td>
</tr>
<tr>
<td>14</td>
<td>blue</td>
<td>1100</td>
</tr>
<tr>
<td>15</td>
<td>yellowgreen</td>
<td>124</td>
</tr>
<tr>
<td>16</td>
<td>darkmagenta</td>
<td>131</td>
</tr>
<tr>
<td>17</td>
<td>darkturquoise</td>
<td>574</td>
</tr>
<tr>
<td>18</td>
<td>skyblue</td>
<td>179</td>
</tr>
<tr>
<td>19</td>
<td>steelblue</td>
<td>158</td>
</tr>
<tr>
<td>20</td>
<td>magenta</td>
<td>381</td>
</tr>
<tr>
<td>21</td>
<td>brown</td>
<td>913</td>
</tr>
<tr>
<td>22</td>
<td>tan</td>
<td>345</td>
</tr>
<tr>
<td>23</td>
<td>black</td>
<td>559</td>
</tr>
<tr>
<td>24</td>
<td>turquoise</td>
<td>1384</td>
</tr>
<tr>
<td>25</td>
<td>salmon</td>
<td>336</td>
</tr>
<tr>
<td>26</td>
<td>white</td>
<td>186</td>
</tr>
<tr>
<td>27</td>
<td>red</td>
<td>1520</td>
</tr>
<tr>
<td>28</td>
<td>bisque4</td>
<td>443</td>
</tr>
<tr>
<td>29</td>
<td>greenyellow</td>
<td>363</td>
</tr>
<tr>
<td>30</td>
<td>darkgrey</td>
<td>206</td>
</tr>
<tr>
<td>31</td>
<td>darkorange2</td>
<td>86</td>
</tr>
</tbody>
</table>
Figure 36. Visualization of the eigengene network representing the relationships among the modules. (A) Hierarchical clustering dendrogram of the eigengenes based on dissimilarity measure. (B) Heat map showing the correlation between eigengenes.
A systems immunology approach to GVHD

**Figure 37.** Heat map representation of the module-trait relationships. The table is colour-coded by correlation according to the colour label and each cell contains the corresponding p-value.
Figure 38. Eigengene network generated using the edge-weighted spring embedded layout algorithm (weight based on the inter-module adjacency measure, displayed as the relative distance between nodes; node area proportional to the size of the module; edge thickness proportional to the inter-module correlation coefficient; edge darkness proportional to the statistical significance of the correlation; edges with $p$-value > 0.05 not displayed).
2.2. WGCNA defined modules are preserved in an independent murine model of GVHD

Before advancing to the in-depth characterisation of the individual modules, the reproducibility of this modular organisation of the transcriptional profiles was tested by evaluating module preservation in an independent test data set (B6→129 model). For this purpose, the “modulePreservation” function from the WGCNA R package\(^\text{(261)}\) was used, which calculates a composite \(Z_{\text{summary}}\) statistic, taking into consideration both the network density and connectivity to determine whether the nodes remain highly connected in the test network, and if the connectivity pattern in the test network is similar to the reference network. Figure 39 summarises the results from this module preservation study.

According to the threshold guidelines recommended by Langfelder\(^\text{(261)}\) (\(Z_{\text{summary}} \geq 10\): “strongly preserved” modules; \(2 \leq Z_{\text{summary}} < 10\): “weak to moderately preserved” modules; \(Z_{\text{summary}} < 2\): “not preserved” modules), 19 out of the 31 modules were strongly preserved (modules 1, 2, 3, 4, 6, 7, 9, 13, 14, 15, 17, 23, 24, 26, 27, 28, 29, 30, 31) and only 1 module was not preserved (module 5).
Figure 39. Module preservation study. The plot shows the summary statistics $Z_{summary}$ as a function of the module size (blue line: $Z_{summary} = 10$; red line: $Z_{summary} = 2$).
2.3. Organ specific modules are conserved across species

To validate the modules further and assess their potential translational relevance, GSEA was run on two published and publically available data sets of gene expression profiling from peripheral blood CD3+ T cells of rhesus macaques (GSE73723) and humans (GSE73809) developing GVHD following allogeneic BMT,(298) using the strongly preserved modules as gene sets.

In rhesus macaques, module 1 was found to be strongly associated with the GVHD samples from untreated alloBMT recipients (Figure 40-A), while module 13 was associated with the synBMT control samples (Figure 40-B). Interestingly, the enrichment for module 1 was lost in the samples from alloBMT recipients treated with sirolimus (Figure 40-C) or with tacrolimus-methotrexate (Figure 40-D), reinforcing the relevance of this module in GVHD biology.

In the human data set, where samples from patients with and without GVHD were compared, there was a clear segregation of the SLO-associated modules (modules 1, 2, 3, 4, 6 and 17) which were enriched in the GVHD group (Figure 41), while the only gene set enriched in the non-GVHD group was module 24, a synBMT-associated module (Figure 42).

Since these two data sets were constructed using only blood samples, the preservation of the GVHD target organ-associated modules was not possible to determine in this analysis. The lack of published data examining the gene expression of donor T cells in the peripheral tissues both in animal models
and in humans highlighted an important gap in the current knowledge of GVHD. To explore this research niche, a collaboration with Dr Laura Jardine and Dr Mathew Collin from the Newcastle University Institute of Cellular Medicine was initiated with the objective of studying the transcriptional profile of CD8+ T cells isolated from the blood and skin of transplanted patients with GVHD. Being an ongoing project, the results included in this thesis are based only the analysis of the first 4 patients.

Due to the low number of cells isolated from the tissue samples, gene profiling was performed using the deep-sequencing technology RNA-seq. Calculation of read counts, normalization and differential expression analysis was performed using the Basespace Illumina tool Cufflinks Assembly & DE (https://basespace.illumina.com/apps/2582581/Cufflinks-Assembly-DE).

Given the small number of samples and the heterogeneity of the patients’ background (different acute GVHD severities, variable lengths of time from first GVHD signs and symptoms until sample collection, concurrent comorbidities and treatments) the strategy employed for the analysis of the data involved the computation of the single-sample GSEA (ssGSEA) for the 31 modules identified in the murine data set. This methodology is an extension of GSEA allowing the definition of an enrichment score that represents the degree of absolute enrichment of a gene set in each sample within a given data set.\(^{(265)}\) To establish the relative enrichment of the modules in each compartment, the NES were compared across the paired blood, dermis and epidermis samples for the individual patients.
Figure 43 is a heat map representation of this comparison. From its analysis it becomes evident the existence of a pattern in module enrichment that, in 3 of the 4 patients, successfully distinguished between the samples derived from the epidermis or from the blood. Although in patient A the pattern is not present, it is noteworthy that the epidermis-specific module 28 and the pan-GVHD target organ-associated module 29 are invariably overrepresented in the samples from the epidermis, in all patients.

Interestingly, the pattern of module enrichment in the dermis was considerably similar to that of the blood. Although this could suggest that, in humans, the dermis compartment is functionally very different to that of mice, it is also possible that this result could just have identified a pitfall in the processing methodology. In contrast to the epidermis, the dermis is a highly vascularised tissue; the impracticability of flushing the vasculature of the biopsy specimens may cause substantial contamination of the dermis samples with circulating lymphocytes, which could bias the analysis towards a more blood-like signature.

Altogether, these results corroborate that the modules found in the MataHari T cell model are preserved and conserved across species, which suggests they correspond to biologically meaningful sets of genes.
Figure 40. Validation of the murine modules on a data set from rhesus macaque. GSEA was performed to compare profiles of untreated alloBMT versus synBMT samples (A, B), sirolimus treated alloBMT versus synBMT samples (C), and tacrolimus-methotrexate treated alloBMT versus synBMT (D). Note the statistically significant enrichment for module 1 is observed in the untreated alloBMT samples (A), which is lost upon immunosuppressive treatment (C, D).
Figure 41. Validation of the murine modules on a data set from human samples.
Figure 41 (previous page). Validation of the murine modules on a data set from human samples. GSEA was performed to compare profiles of samples from transplanted patients with and without GVHD. (A – F) Statistically significant enriched modules in the GVHD samples. Note that all the enriched gene sets (1, 2, 3, 4, 6 and 17) are SLO-associated modules.

Figure 42. GSEA revealed statistically significant enrichment for module 24 in the samples from patients without GVHD.
Figure 43. GVHD target organ-associated modules accurately distinguish between samples isolated from the epidermis and samples isolated from the blood of transplanted patients with cutaneous acute GVHD. The heat maps show normalised enrichment score (NES) for each gene set determined by ssGSEA of blood, dermis and epidermis paired samples from 4 different patients.
2.4. Module annotation reveals biologically meaningful clusters of genes

To evaluate the functional relevance of coexpressed genes within the 19 strongly preserved modules, WebGestalt (WEB-based GEne SeT AnaLysis Toolkit)\textsuperscript{263,264} was used to examine the overrepresentation of the Gene Ontology (GO) categories (biological processes, molecular function, and cellular component) (\textit{Appendix - Supplementary Figures 2 - 20}) and KEGG pathways (\textit{Appendix - Supplementary Table 2}) within each module. The analysis of these results indicated that:

- modules 13, 14 and 15, associated both with donor and synBMT samples, were related with epigenetic regulation of gene expression (by chromatin organisation and histone modification), nucleic acid biosynthesis, and organisation and biogenesis of cellular components;
- module 7, associated with all alloBMT samples, was related with T cell activation and migration;
- module 17, associated with alloBMT LN samples, was mostly related with immune response modulation by Toll-like receptor and protein metabolism regulation by the mTOR pathway;
- modules 2, 3 and 4, associated with alloBMT spleen samples, were related with cellular component biogenesis and organisation, protein biosynthesis, and energy production through oxidative phosphorylation;
• module 1, associated with alloBMT bone marrow and blood samples, was related mainly with regulation of the cell cycle, nucleotide metabolism, and organelle organisation;

• module 29, associated with all GVHD target organs, was related primarily with cell differentiation and regulation of apoptosis, in response to TCR and MAPK signalling;

• modules 9, 13 and 27, associated with alloBMT gut (LP) samples, were related with epigenetic regulation of transcription (through chromatin organisation and histone modification) and regulation of protein metabolism by the ubiquitin-proteasome system;

• modules 6, 30 and 31, associated with alloBMT gut (IEL) samples, were related with T cell migration and activation;

• modules 23 and 26, associated with alloBMT skin (dermis) samples, were related with cell migration and differentiation, and oxidative phosphorylation;

• module 28, associated with alloBMT skin (epidermis) samples, were mainly related with regulation of the inflammatory response.

This modular representation of donor CD8$^+$ T eff cell gene expression further highlights the similarities between donor and synBMT samples, and emphasises the SLO-GVHD target organ transcriptome dichotomy revealed in the exploratory analysis of the data. In the steady state and after a synBMT, the donor T cells appear to be in quiescent state dominated by cell cycle regulation. However, in the context of an alloBMT a number of different cellular programs are activated, some of which are broadly expressed across all organs, while others are discretely associated with specific tissue compartments.
2.5. Identification of highly connected hub genes

According to the theory of module networks, the genes with the highest degree of connectivity within a module (known as hub genes) are interpreted as being representative of that module and are expected to be drivers required for core cellular pathways.

Module specific driver genes were identified using the eigengene-based connectivity, as a module membership measure, and their correlation with the sample traits, as a measure of gene significance (Appendix - Supplementary Table 3), and calculation of the over-enriched GO terms was performed using WebGestalt\(^{264}\) (Appendix - Supplementary Figures 21 - 39). Table 14 summarises these results, showing the top 10 driver genes in each module and the top 3 over-represented GO terms.

Integration of the information acquired from the module-trait relationship analysis, GO and KEGG pathway enrichment study and driver genes identification revealed biologically relevant tissue-specific features that reflect distinct stages of differentiation and activation, metabolic status, and cell fates of donor CD8\(^+\) T\(_{\text{eff}}\) cells.

Consistent with the current model of GVHD immunopathology,\(^{185}\) the pan-alloBMT associated gene set, module 7, includes key genes essential for T cell activation, expansion and migration, such as CD28 and CD40L, Tbx21 and Itga4. Moreover, the modules associated with SLO appear to mainly be
related to priming and proliferation cellular programs. Modules 1, 3 and 4, a common SLO transcriptional signature, are linked to cell cycle, DNA replication and oxidative metabolism, including genes like Cdc26, Cdk7, Ccnh, Atr, Hadh, Cs, Mdh1, Mdh2, and numerous others encoding proteins of the electron transport chain. Of relevance, module 17 appears to mark the contrast between SLO and GVHD target organs, being highly positively correlated with the LN but negatively correlated with both the gut and the skin. Interestingly, module 17 is mostly related to Toll-like receptor signalling (e.g. Tlr1, Tlr3, Tlr7, Irf7, Ifnar2, Irak4), stressing the importance of co-stimulation complementary to TCR-induced signals to enhance T_{eff} cell proliferation, survival and cytokine production in the lymphoid organs.

The GVHD target organ-associated modules, on the other hand, have an overwhelming number of genes linked to pathogenicity both through direct cytotoxicity and pro-inflammatory cytokine generation. In this context, module 29 could be viewed as a common transcriptional signature associated with all GVHD target organs, which is related to regulation of cell fate decisions through mitogen-activated protein kinases (MAPK) and TNF-alpha NF-kB signalling, and include several genes that determine Tc1 and Tc17 differentiation (e.g. Tnfr2, Traf1, Traf4, Gadd45b, Nr4a2, Tnfaip3, Rel, Fosl2, Kdm6b, Skil, Chd7). Taking into consideration the gut-associated module 31, further Tc17 polarisation is expected to occur in gastrointestinal tract as indicated by expression of additional Th17/Tc17 differentiation-related transcription factors, such as Ikar. In contrast, the epidermis-specific module 28 was predominantly linked to T cell activation (Il2ra, Tnfs9, Cd44), Th1-type pro-inflammatory
cytokine generation (Ifng, Csf1 and Csf2), glycolysis (Pkm2), Notch signalling (Rpbj, Furin), and survival (Bcl2l1).
### Table 14. Module specific driver genes summary.

<table>
<thead>
<tr>
<th>Mod.</th>
<th>Top 10 driver genes</th>
<th>top 3 over-represented GO terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Arhgap19, Hist2h2ac, Ube2c, Ccnb2, Kif18b, Ckap2, Cep55, Cdc20, Depdc1a, Prr11</td>
<td>GO:0007049 – cell cycle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0051301 – cell division</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0048285 – organelle fission</td>
</tr>
<tr>
<td>2</td>
<td>Gtpbp8, Psmg1, Bphl, Tomm70a, Eif3m, Znhit6, Thumpd3, Trmt5, Tfb1m, 2210016F16Rik</td>
<td>GO:0006996 – organelle organisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0051301 – cell division</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0008152 – metabolic process</td>
</tr>
<tr>
<td>3</td>
<td>Lyrm2, Cst7, Dtd1, Bloc1s1, Eif1b, Rpl37a, Pop5, Tor2a, Ostc, 2700060E02Rik</td>
<td>GO:0044237 – cellular metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0009060 – aerobic respiration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0007005 – mitochondrion organisation</td>
</tr>
<tr>
<td>4</td>
<td>Asap1, Pglyrp1, Nup37, Lclat1, Dera, Slamf6, Klhdc2, Tjp2, Pts, Nqo2</td>
<td>GO:0055114 – oxidation-reduction process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0006091 – generation of precursor metabolites and energy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0022900 – electron transport chain</td>
</tr>
<tr>
<td>6</td>
<td>Fdft1, Tacc2, Srgap3, Fhl2, Ilih5, Ctsb, Angptl2, Nmb, Xcl1, Bcl2a1a</td>
<td>GO:0035746 – granzyme A production</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0071613 – granzyme B production</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0032623 – IL-2 production</td>
</tr>
<tr>
<td>7</td>
<td>Zdhhc2, Nek7, Ptpn11, Cox17, Plek, S100a11, Farp1, Anxa2, Ccdc50, Pscr1</td>
<td>GO:0080134 – regulation of cell activation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0032101 – regulation of response to external stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0051174 – regulation of phosphorus metabolic process</td>
</tr>
<tr>
<td>9</td>
<td>Nab1, Adss, Zwint, Slu7, Ddx52, Ugp2, Cast, Tars, Gramd1c, Hat1</td>
<td>GO:0016070 – RNA metabolic process</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0006397 – mRNA processing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0006348 – chromatin silencing at telomere</td>
</tr>
<tr>
<td>13</td>
<td>Zmym2, Kmt2a, Usp32, Herc1, Bach2, Brwd1, Nipbl, Vps13d, Dynch1h1, Ly75</td>
<td>GO:0016569 – covalent chromatin modification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0016570 – histone modification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0006355 – regulation of transcription, DNA-dependent</td>
</tr>
<tr>
<td>14</td>
<td>Stat1, Nuak2, Mga4a, Arl4c, Atp1b1, Tbx2r, Syde1, Pik3ip1, Rapgef6, Galnt10</td>
<td>GO:0048583 – regulation of response to stimulus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0023051 – regulation of signaling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0010646 – regulation of cell communication</td>
</tr>
<tr>
<td>15</td>
<td>Rras2, G0s2, Eomes, S1c14a1, Ppargc1b, Utp20, Prmt3, Sh3bp1, Kat2a, Dkc1</td>
<td>GO:0022613 – ribonucleoprotein complex biogenesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0002302 – CD8-positive, alpha-beta T cell differentiation involved in immune response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GO:0034641 – cellular nitrogen compound metabolic process</td>
</tr>
</tbody>
</table>
Table 14 (continued)

<table>
<thead>
<tr>
<th>Mod.</th>
<th>Top 10 driver genes</th>
<th>top 3 over-represented GO terms</th>
</tr>
</thead>
</table>
| 17   | Slfn5, Samhd1, Herc6, Prkdc, Smchd1, Klhl6, Fchsd2, Fam208a, Atf7ip, Parp9 | GO:0050776 – regulation of immune response  
GO:0006952 – defense response  
GO:0032753 – positive regulation of IL-4 production |
| 23   | Cyt1, Asphd2, Mia, Tslp, Slc51a, Snord88a | |
| 24   | Tmem52, Krtap19-3, Ccdc85a, Amy1, Cntnap2, Pcdhb1, Nefl, Bpifa2, Rhou, D930020B18Rik | GO:0032536 – regulation of cell projection size |
| 26   | Clk1, Ddx5, Srrm2, Gabbr1, Malat1, Nxf1, Ptbp2, Leng8, Cdk16, 2810403A07Rik | GO:0043484 – regulation of RNA splicing  
GO:0000375 – RNA splicing, via transesterification reactions  
GO:0000398 – mRNA splicing, via spliceosome |
| 27   | Ramp3, Mpp7, Arhgap31, Socs5, Zbtb46, Mest, Tnfsf4, Rnf157, Ubtd2, Abcb1b | GO:0009966 – regulation of signal transduction  
GO:0016310 – phosphorylation  
GO:0036211 – protein modification process |
| 28   | Ifitm5, Cpd, Csf2, Ifng, Nek6, Tnfsf9, Tesk1, Csf1, Ccl20, Cd44 | GO:0032946 – positive regulation of mononuclear cell proliferation  
GO:0043067 – regulation of programmed cell death  
GO:0048584 – positive regulation of response to stimulus |
| 29   | Dusp5, Traf1, Tnfaip3, Tgif1, Cog10b, Rgs1, Nfkbiz, Dusp4, Smim3, Gpr132 | GO:0043067 – regulation of programmed cell death  
GO:0006355 – regulation of transcription, DNA-dependent  
GO:0019219 – regulation of nucleobase-containing compound metabolic process |
| 30   | Arf6, Sh3bgtr3, Bak1, Sec61a1, Mapk1ip1 | GO:0010638 – positive regulation of organelle organization |
| 31   | Slc16a10, Ikzf3, Asb2, Cd38, P2rx7, Chn2, Osgin1, 1810011H11Rik | GO:0046651 – lymphocyte proliferation  
GO:0032845 – negative regulation of homeostatic process |

(*) No enrichment for GO terms.
2.6. Discussion

The development of new and more powerful computational resources in the last years has fuelled a surge of bioinformatics methods that have enabled researchers to analyse data sets of increasing complexity. As a result, systems biology has flourished and has provided valuable insight into the dynamic and complex nature of the immune system.

One of the basic principles underlying this holistic approach is the perception that the whole of living organisms are more than the sum of their parts. In the context of genomics, this aphorism translates into the observation that the understanding of cellular programs is much more complex than the identification of individual genes implicated in any given pathway, it requires the study of the network of relationships established at the molecular and cellular levels.

In this respect, correlation networks tools have proved to be much more useful than classical statistical methods to analyse large, high-dimensional data sets.\(^{(299)}\) WGCNA, in particular, has been successfully employed to analyse gene expression data in various contexts, i.e. brain cancer,\(^{(231)}\) yeast cell cycle,\(^{(300)}\) mouse genetics,\(^{(301-304)}\) primate brain tissue,\(^{(305-307)}\) diabetes,\(^{(308)}\) chronic fatigue patients\(^{(309)}\) and plants.\(^{(310)}\) There is now ample evidence that gene correlation identifies clusters of co-regulated or co-expressed genes that frequently act as members of a biological pathway. For instance, this methodology was instrumental in the evolution of our understanding of
macrophage differentiation extending the concept of macrophage activation from M1 versus M2 polarisation to a spectrum model, characterized by a variety of programs linked to chronic inflammation granulomatous diseases.\(^{(311)}\)

By using WGCNA to explore the MataHar T cells data set, and assemble a signed network of highly correlated genes across the different study groups and tissues, 31 modules of positively correlated genes were identified (Figure 35 and Table 13). These groups of genes were shown to have distinct patterns of expression depending on the site from which samples were collected (Figure 37) and represented a robust network of genes involved in GVHD pathology, as the majority of them were strongly preserved in an independent murine GVHD model data set (Figure 39).

Moreover, the modules with the strongest preservation statistics \(Z_{\text{summary}} \geq 10\) were shown to accurately distinguish between blood CD3\(^+\) T cell samples from alloBMT or synBMT recipients, and more importantly, to segregate those with GVHD, both in a non-human primate model of GVHD and with patient samples (Figure 40 – Figure 42). The paucity of published data regarding donor T cell gene expression at the sites of tissue injury led to the development of a collaborative study with the Newcastle University Institute of Cellular Medicine aimed at characterising the differences in CD8\(^+\) T cell effector programs isolated from the blood vs the skin of transplanted patients with GVHD. Even though at this time only a limited number of samples were available for analysis, the results obtained with them support the efficacy of this modular organisation in identifying tissue-specific groups of genes over-represented in GVHD target organs (Figure 43). Nevertheless, further
validation of the GVHD target organs-associated modules is required, in particular for modules associated with gastrointestinal pathology.

The characterisation of these highly preserved modules (Table 14, Appendix - Supplementary Table 3) provided further evidence supporting the biological relevance of the diversity of donor CD8\(^+\) T\(_{\text{eff}}\) cells’ transcriptional profiles during the development of GVHD. The SLO-associated modules were found to mainly represent cellular programs linked with priming and proliferation, including cardinal pathways of T cell activation (e.g. TLR and RIG-I signalling), oxidative metabolic activity (e.g. fatty acid oxidation), and cell cycle regulation (e.g. CDK1 and aurora kinase pathways). The GVHD target organs-associated modules, on the other hand, were mostly related to effector function related pathways, highlighting their role in determining target organ damage. Interestingly, these modules seem to define a stepwise differentiation program which comprises a nuclear set of instructions linked to early Tc1/Tc17 polarisation, common to all GVHD target organs, that are further enhanced by tissue-specific cellular programs.

These observations suggest that the peripheral tissues play autonomous and dominant roles in determining T\(_{\text{eff}}\) cell function. Both in the gut and in the skin, a progression is observed in their subcompartments, from cellular programs centred on migration in response to chemokine signalling, in the lamina propria and dermis, to programs centred on cytotoxic function, in the intra-epithelial compartment and epidermis.

Interestingly, this difference in subcompartments is most striking in the skin, where a large number of genes encoding multiple pro-inflammatory
cytokines (i.e. Ifng, Il2, Il3, Il13, Il17a, Csf1, Csf2) and cytokine receptors (e.g. Il2ra, Il1r1) are highly correlated with T\textsubscript{eff} cells in the epidermis, as part of module 28.

One possible explanation for this phenomenon could be that as the T\textsubscript{eff} cells migrate through the tissues they receive environmental cues which shape their post-activation differentiation path. Such a model would imply the requirement of \textit{in situ} interactions between the donor CD8\textsuperscript{+} T cells and tissue resident cell populations which would “license” them to achieve their full effector potential.
3. Donor CD8$^+$ T$\text{eff}$ pathogenicity in the skin is regulated \emph{in situ} by epidermal Langerhans cells

To better understand the cellular mechanisms underlying the divergence of transcriptional profiles among the GVHD target organs, and in particular between subcompartments within each organ, the remainder of the study was focused on the skin.

The skin’s immune niche is composed of two main compartments: the dermis, populated by macrophages, dendritic cells, mast cells, \(\gamma\delta\) T cells and \(\alpha\beta\) T cells; and the epidermis, where Langerhans cells (LC), dendritic epidermal \(\gamma\delta\) T cells (DETC) and tissue memory CD8$^+$ T cells (T$\text{RM}$) are the only resident immune cell populations.$^{312,313}$ The maintenance of this cellular network and its tight regulation are crucial for an effective cutaneous immune surveillance, as made evident from the increased incidence and severity of cutaneous infections and malignancies in immunocompromised patients, and by the development of a variety of skin disorders when the skin defence mechanisms are misdirected.$^{314}$

In the context of alloBMT, the skin’s microenvironment is greatly disrupted. As a consequence of the conditioning regimen, most of the host skin immune populations are depleted and quickly replaced by donor derived ones. Notably, the Langerhans cells are the exception: being radioreistant and
capable of proliferating locally and maintaining homeostasis independently of the bone marrow, Langerhans cells persist long after BMT and full donor chimerism is only achieved in the presence of donor T cells.\textsuperscript{(315-317)}

The characterisation of the changes in the skin immune network following alloBMT has been one of the focus of Dr Bennett and Prof Chakraverty’s lab. Studies conducted by Dr Thomas Conlan, Dr Cara Lomas and Dr Ivana Ferrer in the minor histocompatibility antigen mismatched model of acute GVHD (MataHari T cell model) have shown that skin infiltrating donor CD8\(^+\) T cells migrate through the dermis and invade the epidermis, localising preferentially along the basement membrane, in close proximity to the Langerhans cells (\textbf{Figure 44}).

The analysis of the donor CD8\(^+\) T cells epidermal infiltration and of the Langerhans cells replacement kinetics revealed that even though the ratio between host and donor LC is reversed in the first 3 weeks following donor CD8+ T cell entrance into the skin, a considerable number of host derived LC still populate the epidermis at that point (\textbf{Figure 45}).

Moreover, Dr Thomas Conlan’s work in the Langerin-DTR system showed a unique association between LC and skin GVHD development.\textsuperscript{(217)} Expressing the simian diphtheria toxin (DT) receptor downstream of the Langerin gene promotor, these transgenic mice become selectively depleted of epidermal LC after treatment with DT. Using the Langerin-DTR mice as recipients for alloBMT, Dr Conlan was able to establish that in the absence of LC the severity of cutaneous GVHD was greatly abated, which was directly correlated with a significant reduction in the number of donor CD8\(^+\) T cells.
infiltrating the epidermis (Figure 46). This effect was observed both after systemic or localised LC depletion (Figure 47), suggesting that GVHD immunopathology modulation by LC occurred in situ.
Figure 44. Immunofluorescence staining of the skin of alloBMT recipients with cutaneous acute GVHD (MataHari T cell model). (A) Cross section of the dermal-epidermal transition showing the preferential localisation of donor CD8$^+$ T cells (green; arrows) along the basement membrane (blue). (B) Epidermal sheet showing the close proximity of donor CD8$^+$ T cells (green; arrows) and LC (cyan; arrow heads).

Courtesy of Dr. Cara Lomas and Dr Ivana Ferrer.
Figure 45. Host-derived LC replacement kinetics after alloBMT. Although the ratio between host and donor LC is reversed in the first 3 weeks following donor CD8+ T cell entrance into the skin (A), a considerable number of host-derived LC still populate the epidermis at that point (B).

Unpublished data, courtesy of Dr Ivana Ferrer.
Figure 46. Effect of systemic host LC depletion on cutaneous acute GVHD. Comparison between the histology scores (A) and donor T cell infiltration of the epidermis (B) of LC replete (black circles) and LC depleted (white squares) alloBMT recipients. In the absence of LC, the severity of cutaneous GVHD is reduced, which was accompanied by a significant contraction in the number of epidermal donor T cells (PBS treated: n=9; DT treated: n=8).

Courtesy of Dr Thomas Conlan.

Figure 47. Effect of localised host LC depletion on donor CD8+ T cell infiltration of the epidermis. (A) Langerin-DTR alloBMT recipients received PBS to the right ear and DT to the left ear by intradermal injections at D-21. (B) At D+7 post-transplant, accumulation of donor CD8+ T eff cells was significantly reduced in the ear that had been depleted of LC (n=9 in each group).

Courtesy of Dr Thomas Conlan.
3.1. LC depletion reduces the differences between dermal and epidermal donor CD8⁺ Teff cells transcriptional profiles

To further investigate the role of host LC in the development of GVHD, the adapted MataHari T cell → Langerin-DTR model was used, in which MataHari CD8⁺ T cells were transferred together with female wild-type C57BL/6 bone marrow and polyclonal CD4⁺ T cells into lethally irradiated Langerin-DTR male recipients, treated either with PBS or DT prior to the transplant. At D+7 post-transplant, the donor derived CD8⁺ T cells were isolated from the peripheral lymph nodes, dermis and epidermis of the recipients; the samples’ RNA was purified and a microarray analysis was performed using the Affymetrix GeneChip® Mouse Gene 2.0 ST array, as previously described.

As an exploratory analysis of the dataset, the effect of LC depletion was assessed by performing pairwise comparisons of the transcriptional profiles of the six groups of samples. As illustrated by the correlation matrix depicted in Figure 48, DT treatment had no significant effect on the lymph nodes or on the dermis, but induced a change in the gene expression of Teff cells which infiltrate the epidermis, reducing the disparity between the two skin compartments. This observation is consistent with the knowledge that treatment with DT is nontoxic to mice and only affects cells expressing the simian DT receptor.
Figure 48. Correlation matrix comparing the transcription profiles of donor CD8+ T\textsubscript{eff} cells isolated from the lymph nodes, dermis and epidermis of LC depleted (DT) and LC replete (PBS) alloBMT recipients, at D+7 post-transplant. Groups represent the average expression of the biological replicates. The colour code reflects the Pearson correlation coefficient for each comparison. LC depletion affects mainly the transcriptional profile of epidermal T cells, reducing the differences between the two skin compartments.
3.2. In the absence of LC, epidermal donor T cells fail to fully up-regulate the effector cell programs

In order to better characterise the changes induced by DT treatment, the enrichment for the gene sets defined in the KEGG PATHWAY collection was determined. By comparing dermis vs lymph nodes and epidermis vs dermis through GSEA, it was possible to identify the shifts in cellular programs as cells transited from one compartment to another, and to evaluate the impact that LC depletion produced on these processes (Appendix - Supplementary Table 4).

In Figure 49 the normalised enrichment score (NES) for the 91 gene sets (rows) significantly over- or under-represented in each pair of tissues were compared according to treatment (columns). The results confirmed that DT treatment had a negligible effect on the array of pathways that were up- and down-regulated when donor CD8$^+$ T cells migrated from the lymph nodes to the dermis, and evidenced that a variety of cellular programs that are normally triggered upon transition from the dermis to epidermis failed to be activated in the absence of LC.

Examination of the gene sets that were differentially enriched in the epidermis in the presence of LC revealed that the genetic programs were related to: cellular processes (such as cell growth, death, transport and catabolism), signal transduction, genetic information processing (mainly replication and repair, transcription and translation), energy and nucleic acid
metabolism; and were linked to immune system function and immune disease, namely GVHD (Figure 50, Appendix - Supplementary Table 4).
A systems immunology approach to GVHD

Figure 49. Heat map showing the enrichment for the 91 gene sets significantly over- or under-represented in the dermis vs the lymph nodes and in the epidermis vs the dermis, with or without LC. The colour code reflects the normalised enrichment score (NES) for each gene set. Note that the majority of pathways up-regulated by T cells in a LC replete epidermis are absent when LC have been depleted.
**Figure 50.** Classification of the gene sets that were differentially enriched in the epidermis in the presence of LC according to KEGG PATHWAY mapping.
3.3. Donor CD8\(^+\) T cells expression of the epidermal specific transcriptional signature is regulated by host LC

To test the concept that LC were implicated in defining an epidermal specific transcriptional signature, the expression pattern of the genes from module 28, previously identified by WGCNA in the MataHari T cell model as being highly correlated with the epidermis, was compared between PBS and DT treated groups. As depicted in the heat map in Figure 51, the 443 genes that form this module were consistently highly expressed in the LC replete epidermis samples (PBS treated), however, in the absence of LC (DT treated), this pattern was broadly disrupted and only small clusters of genes were expressed to the same extent as in the controls. This reduction in the level of expression was most profound for the module’s hub genes (Figure 52) denoting their central role in driving the expression of the remaining genes.

The analysis of the overrepresentation of the GO categories and KEGG pathways for the set of module 28 genes that failed to be upregulated in the epidermis in the DT treated group (Appendix – Supplementary Figure 40 and Supplementary Table 5) indicated that the main biological processes affected by the absence of LC were related to regulation of leukocyte proliferation / apoptosis and development of the immune response, involving cytokine-cytokine receptor interactions, and signalling through Jak-STAT, MAPK, TLR and TCR pathways.
Figure 51. Heat map of the 443 genes that form module 28, comparing the transcriptional profiles of the lymph nodes, dermis and epidermis, in the presence (PBS) or absence (DT) of LC. The colour code reflects the Z-score of expression. LC depletion leads to failure in T cell upregulation of this epidermal-specific gene signature.

Figure 52. Module 28 top 100 gene network, where the circles represent the genes and the lines represent the connectivity between the genes. The colour code reflects the log2 fold change (log2 FC) in expression level in the epidermis of PBS treated vs DT treated groups. Note the extreme down-regulation of most of module’s hub genes in the absence of LC.
3.4. LC dictate donor T cell accumulation, survival, effector function and $T_{RM}$ differentiation potential in the epidermis

To validate these results, a second set of experiments in the MataHari T cell → Langerin-DTR model was conducted to evaluate by multicolour flow cytometry the proliferation, effector function and survival of skin infiltrating donor CD8$^+$ T cells in the presence or absence of LC. AlloBMT recipients treated with either PBS or DT at D-21 were euthanized at D+7 and samples from the lymph nodes, dermis and epidermis were collected (Figure 53). The proliferative index of donor CD8$^+$ T cells was determined by EdU staining (alloBMT recipients were given EdU 25 mg/kg iv 12h prior to sacrifice); survival was assessed by quantification of caspase-3 activity; ex vivo cytotoxic effector function was measured by staining for IFN-γ production, without further re-stimulation (Figure 54 and Figure 55).

Although no difference was detected on the proliferation index of epidermal T cells (Figure 54-A and Figure 55-A), it was invariably found that a higher proportion of these cells were apoptotic (PBS: 5.9%± 0.7 vs DT: 7.8% ± 0.4; mean ± SD) (Figure 54-B and Figure 55-B). Moreover, in the absence of LC, T$_{eff}$ cells produced significantly lower amounts of IFN-γ (PBS: 20.2% ± 5.6 vs DT: 7.2% ± 3.6; mean ± SD) (Figure 54-C and Figure 55-C).

To investigate the long term effects of host LC depletion on donor CD8$^+$ T$_{eff}$ cells’ capacity to differentiate into tissue-resident memory T cells ($T_{RM}$), a
long term experiment was conducted in collaboration with Dr Ivana Ferrer and Ms Sophie Ward in which PBS or DT treated transplanted Langerin-DTR mice (MataHari T cell—Langerin-DTR model) were sacrificed at weeks 1, 2, 3 and 4 post-transplant and skin samples collected for study of the epidermal kinetics of donor derived CD8$^+$ $\text{TRM}$ cells based on CD69 and CD103 expression (Figure 56). Interestingly, the results showed that in the absence of host LC the accumulation of donor CD8$^+$ $\text{TRM}$ cells in the epidermis was impaired, an effect that gained momentum with time and by week 4 post-transplant, the difference between the percentage of epidermal CD69$^+$ CD103$^+$ MataHari T cells in the two groups achieved statistical significance (PBS: 49.9% ± 9.9 vs DT: 37% ± 10; mean ± SD) (Figure 57).
Figure 53. Experimental design to evaluate the effect of LC depletion on proliferation, effector function and survival of skin infiltrating donor CD8⁺ T cells.
Figure 54. Assessment of the proliferation (A), survival (B) and effector function (C) of donor CD8⁺ T cells by FACS staining. Samples from the lymph nodes, dermis and epidermis were stained for EdU incorporation, active caspase-3 and IFN-γ. The gating strategy used to identify donor CD8⁺ T cells is shown, as well as representative plots for each group and condition.
Figure 55. Summary data referring to the study of the effect of LC depletion upon proliferation (A), survival (B) and effector function (C) of donor CD8$^{+}$ T eff cells. Although no difference in proliferation was observed between groups, in the absence of LC, the percentage of donor CD8$^{+}$ T eff cells was consistently higher and the production of INF-γ was significantly reduced. PBS: n=6; DT: n=5 (except for EdU staining of dermis and epidermis where PBS: n=3); data pooled from 2 independent experiments.
Figure 56. Experimental design to evaluate the effect of LC depletion donor CD8+ T cells capacity to differentiate into epidermal T_{RM} cells.
Figure 57. Kinetics of donor derived CD8+ epidermal T_{RM} cells in the presence or absence of LC. (A) Gating strategy used to identify donor CD8+ T cells in the epidermis. Representative plots of CD103 and CD69 staining are shown for each time point and condition. (B) Summary data. In the absence of LC, a lower proportion of donor T cells express a T_{RM} cell phenotype. For each time point, PBS: n=7, DT: n=5; data pooled from 2 independent experiments.
3.5. Responsiveness to IFN-γ by LC and Notch pathway signalling are central in the host LC – donor T cell crosstalk

Altogether, the gene expression profiling and the protein expression data support the hypothesis that a crosstalk between LC and donor CD8⁺ Teff cells occurs in the epidermis which regulates T cell effector function and fate, and ultimately determines the severity of the GVHD immune reaction. Consistent with this concept, in experiments performed by Dr Thomas Conlan which compared the transcription profile of LC after allogeneic versus syngeneic BMT controls, it was observed a significant enrichment for type I and type II IFN responsive genes (Table 15).\(^{(217)}\)

One possible interpretation of these data is that a feed-forward loop involving LC responsiveness to IFN-γ is required to be established for LC to induce cognate activation, expansion and survival of incoming T cells. Moreover, in accordance with the current understanding of the relevance of Notch signalling to T cell alloreactivity,\(^{(318-320)}\) in particular to the production of IFN-γ,\(^{(321)}\) the present work identified Rbpj and Furin, two key regulators of Notch pathway, as part of the group of genes that drives the expression of module 28.

In order to test this hypothesis, two sets of experiments (Figure 58 and Figure 62) were performed using the MataHari T cell model that evaluated the
effect of interrupting these two signalling pathways upon the accumulation, survival and IFN-γ production of donor CD8⁺ T cells in the skin.

To evaluate the requirement of LC responsiveness to IFN-γ, Ifngr1⁻/⁻ or WT recipients were transplanted following the MataHari T cell model, with the exception that T cells were activated in vitro with concanavalin A and IL-7 prior to injection into the recipients, in order to overcome any priming deficit induced by absence of IFN-γ receptor signalling in the SLO (Figure 58). At D+7 post-transplant samples from skin and spleen were collected and analysed by flow cytometry (Figure 59 - Figure 61).

Both the host LC and the donor CD8⁺ Teff cells were analysed in this study. It was observed that although IFN-γ signalling did not significantly affect LC expression levels of costimulatory molecules, such as CD40, CD70 or CD86, when unable to detect IFN-γ signalling, LC failed to upregulate the expression of MHC class I H-2Db molecules (H-2Db MFI – WT: 1753 ± 77 vs Ifngr1⁻/⁻: 604 ± 306; mean ± SD) (Figure 59). As a result, the accumulation of donor CD8⁺ Teff cells was reduced specifically in the epidermis of Ifngr1⁻/⁻ recipients (WT: 20 x 10³ cells/g ± 7 x 10³ vs Ifngr1⁻/⁻: 4.7 x 10³ cells/g ± 4 x 10³; mean ± SD), even though the expansion of donor CD8⁺ Teff cells in the spleen was similar in both groups and no difference in the absolute numbers of tissue infiltrating T cells was noted in the dermis (Figure 60 and Figure 61-A). Remarkably, absence of IFN-γ receptor signalling in host LC neither increased T cell apoptosis (Figure 60 and Figure 61-B) nor reduced T cell production of IFN-γ (Figure 60 and Figure 61-C).
The impact of Notch signalling blockade on the development of cutaneous GVHD was assessed through the treatment of alloBMT recipients with LY411575, a potent inhibitor of the presenilin-dependent gamma secretase complex responsible for the release of the Notch Intracellular Domain (NICD) after ligand binding to the Notch receptor. For the purpose of minimising the interference with T cell priming events, the treatment was given only from D+5 onwards, which corresponded to the onset of skin infiltration by donor T cells (Figure 62). As before, at D+7 post-transplant samples from skin and spleen were collected and analysed by flow cytometry.

As illustrated in Figure 63, Notch signalling blockade induced a decrease in IFN-γ production by donor CD8^+ T eff cells in the skin, particularly in the epidermis (DMSO: 23.8% ± 9.5 vs LY411575: 11.2% ± 2.5; mean ± SD), whereas it had no significant effect in the spleen or in the gut. These findings reinforce the idea that donor T cell pathogenicity is determined by distinct tissue specific / selective signals which regulate T cell effector function.
Table 15. Top 10 type I and type II IFN responsive genes upregulated by LC after alloBMT.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Description</th>
<th>Log2 Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>iigp1</td>
<td>Interferon inducible GTPase 1</td>
<td>25.66</td>
</tr>
<tr>
<td>Cxcl9</td>
<td>Chemokine (C-X-C motif) ligand 9</td>
<td>20.76</td>
</tr>
<tr>
<td>Ifi202b</td>
<td>Interferon activated gene 202B</td>
<td>13.70</td>
</tr>
<tr>
<td>Ifi205</td>
<td>Interferon activated gene 205</td>
<td>11.94</td>
</tr>
<tr>
<td>Gbp2</td>
<td>Guanylate binding protein 2</td>
<td>9.89</td>
</tr>
<tr>
<td>Cish</td>
<td>Cytokine inducible SH2-containing protein</td>
<td>8.36</td>
</tr>
<tr>
<td>Cxcl10</td>
<td>Chemokine (C-X-C motif) ligand 10</td>
<td>7.95</td>
</tr>
<tr>
<td>Ccl22</td>
<td>Chemokine (C-C motif) ligand 22</td>
<td>6.49</td>
</tr>
<tr>
<td>Ccr7</td>
<td>Chemokine (C-C motif) receptor 7</td>
<td>5.76</td>
</tr>
<tr>
<td>Oasl2</td>
<td>2'-5' oligoadenylate synthetase-like 2</td>
<td>5.48</td>
</tr>
</tbody>
</table>

*Courtesy of Dr Thomas Conlan*
Figure 58. Experimental design to evaluate the requirement of LC responsiveness to IFN-γ for donor CD8⁺ T cell accumulation in the epidermis.
Figure 59. Comparison of alloBMT driven changes in LC expression of MHC class I H-2Db (A), CD40 (B), CD70 (C) and CD86 (D). Gating strategy used to identify LC, representative histograms of cell surface staining and summary data are shown. When unable to detect IFN-γ signalling (Ifngr1−/−), LC fail to up-regulate MHC class I. WT female (synBMT control): n=5; WT male (alloBMT): n=5; Ifngr1−/− (alloBMT): n=7; data pooled from 2 independent experiments.
Figure 60. Study of the survival (A) and effector function (B) of donor CD8+ T cells isolated from alloBMT recipients incapable of detecting IFN-γ signalling (Ifngr1–/–) vs WT. Samples from the spleen, dermis and epidermis were stained active caspase-3 and IFN-γ. The gating strategy used to identify donor CD8+ T cells and representative plots for each group and condition are shown.
Figure 61. Summary data referring to the study of the effect of LC responsiveness to IFN-γ signalling upon accumulation (A), survival (B) and effector function (C) of donor CD8⁺ T effector cells. Note the significant reduction in tissue infiltrating donor T cells specifically in the epidermis of IIfngr1⁻/⁻ recipients. No differences in donor CD8⁺ T cell survival or effector function was observed between the groups. WT male: n=7; IIfngr1⁺⁺: n=8; data pooled from 2 independent experiments.
Figure 62. Experimental design to evaluate the effect of Notch signalling blockade on donor CD8^+ T cell effector function in the epidermis.
Figure 63. Study of the effect of Notch signalling blockade on donor CD8⁺ T cell effector function in the spleen, dermis and epidermis. (A) Gating strategy used to identify donor CD8⁺ T cells and representative plots for each tissue and condition. (B) Summary data. Treatment with LY411575 dramatically reduced donor CD8⁺ T cell IFN-γ production in the epidermis, with no statistically significant effect on the other tissues. DMSO (control): n=8; LY411575: n=7; data pooled from 3 independent experiments.
3.6. Discussion

Prior work in various infection models has documented the enhancement of the effector function of T cells when they enter the peripheral tissues. McGill et al. and Dolfi et al. have shown that after initial activation in the LN, influenza-specific CD8+ T cells require additional antigen-dependent interactions with DC in the lungs for the establishment of an effective protective response.\(^{222-224}\)

Likewise, in a model of LCMV meningitis, Kang et al. observed that cognate peptide–MHC I interactions occurring in the meninges influenced cytotoxic T cell division in situ;\(^{322}\) in the skin, McLachlan et al. showed that CD11\(^{hi}\) dermal dendritic cells’ antigen presentation in situ drives IFN-\(\gamma\) production by cutaneous T\(_{eff}\) cells.\(^{225}\)

It is conceivable that this “two-hit” model of the immune response is also applicable in the context of GVHD and it is supported by observation of a dichotomy between the transcriptional profiles of T\(_{eff}\) cells in SLO and GVHD target organs, which reflected an up-regulation of effector function genes and pathways in the peripheral tissues. Moreover, in the particular case of the skin, this study identified distinct transcriptional signatures for T\(_{eff}\) cells isolated from the dermis or the epidermis.

Dr Thomas Conlan had demonstrated using the Langerin-DTR system that the severity of cutaneous acute GVHD was greatly reduced when host LC were depleted.\(^{217}\) The advantage of using the Langerin-DTR system is that by treating the recipients with DT at D-21, selective depletion of host LC is
achieved as the repopulation kinetics of LC is considerably longer than other Langerin expressing cell populations. Following on this work, the same model was used to test the effects of LC depletion on the donor CD8$^+$ T cells at the gene expression level.

DT treatment did not alter significantly T$_{eff}$ cell transcriptome in the LN and dermis, yet it produced a profound shift in the gene expression of T$_{eff}$ cells in the epidermis (Figure 48 and Figure 49). This suggests that the changes induced by LC depletion were not due to the absence of migratory LC compromising priming events, but rather a result of the disruption of the epidermal niche. Interestingly the main effect observed was the abatement of the differences between the skin’s two compartments: in a LC replete skin, the dermis-to-epidermis transition of T$_{eff}$ cells was characterised by the up-regulation of a set of genes, identified by WGCNA as module 28, which failed to occur when LC had been depleted (Figure 51).

Consistent with the previous reports that LC depletion was associated with abrogation of cutaneous GVHD, comparison of the cellular programs activated in the epidermis of PBS or DT treated recipients revealed that host LC increased donor CD8$^+$ T$_{eff}$ cell pathogenicity through cognate antigen recognition, co-stimulatory signalling and cytokine-cytokine receptor interactions, thus promoting their accumulation and survival, and enhancing their effector function.

These results were validated by FACS in experiments which confirmed that the reduction in cutaneous GVHD severity was not only due to having lower numbers of donor T$_{eff}$ cells in the skin, in the absence of LC, T cells infiltrating
the epidermis were clearly more pro-apoptotic, produced lower levels of effector molecules such as IFN-γ (Figure 55), and had a reduced potential of differentiation into T_{RM} cells (Figure 57).

Two main features were identified as being central to this LC–T_{eff} cell crosstalk: LC responsiveness to IFN-γ, required for DC maturation\(^{(325)}\) and up-regulation of MHC class I,\(^{(326)}\) and signalling through the Notch pathway, recognised as a critical regulator of allogeneic T cell responses in GVHD.\(^{(319,321,327)}\)

Indeed, in the absence of IFN-γ receptor signalling, LC failed to up-regulate H-2Db expression and T_{eff} cell accumulation in the epidermis was reduced (Figure 60 and Figure 61). It could be argued that the defect in T_{eff} cell accumulation observed in the epidermis was a result of impaired expansion of donor T cells in a system unresponsive to IFN-γ signalling. However, no differences in the absolute numbers of donor T cells were detected in the spleen and dermis, suggesting that in vitro ConA/IL-7 activation of donor T cells was successful in overcoming any priming deficits inherent to the Ifngr1\(^{-/-}\) system.

It has been previously reported that Notch signalling blockade reduces the severity of GVHD by inducing a hyporesponsive program in T cells.\(^{(318,319,321)}\) However, these studies were performed either by using genetic models or prolonged pharmacological and antibody-mediated treatments which produce a sustained suppression of Notch signalling affecting not only T_{eff} cells function but all developmental stages. The alternative approach adopted here (treatment with a GSI only from D+5 onwards) was aimed at preserving the priming events which usually occur in the first 3 days after T cell transfer and
allow for a normal infiltration and early accumulation of donor T cells in the peripheral tissues. Moreover, in these studies only the lymphoid organs (spleen and lymph nodes) were evaluated and the differences in IFN-γ production do not accurately reflect a decrease in T cell effector function *in vivo* as all measurements were performed after re-stimulation either with anti-CD3/anti-CD28 or phorbol 12-myristate 13-acetate and ionomycin. By performing direct *ex vivo* measurements (without re-stimulation) of intracellular IFN-γ in donor CD8+ T<sub>eff</sub> cells from the various sites, it was possible to evaluate the true effect of Notch signalling blockade in each compartment. With this protocol (Figure 63), it became evident that the production of IFN-γ donor CD8+ T<sub>eff</sub> cells is only detectable in the GVHD target organs. Interestingly, the GSI treatment did not impact all the compartments evenly: while there was a consistent reduction in the levels of IFN-γ in the skin, particularly in the epidermis, in the gut the production of IFN-γ remained mostly unaffected. These results are consistent with the proposed modular architecture of donor T<sub>eff</sub> cells gene expression during GVHD development, in which independent programs would be autonomously regulated. The epidermis specificity of the GSI treatment effect upon IFN-γ production, mirroring the changes in effector function observed in donor CD8+ T cells induced by LC depletion, furthers the hypothesis that, in the skin, Notch signalling is a potential dominant pathway involved in the host LC – donor T cell cross talk.

Altogether, the data demonstrate that, in the skin, GVHD is defined by tissue-autonomous instructions dictated by T<sub>eff</sub> cell migration to the epidermis and interaction with LC *in situ*, and is regulated by Notch-dependent signals to enhance local cytokine production by T<sub>eff</sub> cells and Notch-independent signals
to promote T cell accumulation and resistance to apoptosis. Although the results strongly suggest that LC are the central participants in this cross-talk with donor CD8$^+$ T cells, to fully characterise this mechanism, further work is required in models allowing for inducible and cell selective inhibition of these signalling pathways.
IV. Discussion
1.1 Systems immunology – a comprehensive approach to acute GVHD study

The aim of this project was to undertake an unbiased and systems-wide methodology to defining the cellular mechanisms underlying tissue-specific pathology in GVHD. The primary hypothesis being tested was that GVHD target tissues exert dominant, idiosyncratic roles in regulating effector T cell functions.

The approach taken, involving gene expression profiling of donor CD8^+ T cells isolated at key time points of the alloreactive response development (before T cell transfer – naïve donor T cells; D+3 – early post-priming; D+6/D+7 – peak of peripheral tissue T cell infiltration) and from a variety of tissues (secondary lymphoid organs, blood, and GVHD target organs), permitted the characterisation of the transcriptional response of donor T cells during the evolution of GVHD.

For this purpose, two well characterised murine models of GVHD were used which reproduce the clinical features and histological abnormalities found in patients with acute GVHD. The models were selected for their clinical relevance, as the B6 → 129 model,\(^{(267,268)}\) being a multiple minor histocompatibility antigen mismatched BMT model, is similar to a standard HLA-identical sibling SCT,\(^{(328)}\) and the MataHari T cell model,\(^{(118)}\) being a single
A systems immunology approach to GVHD

Chapter IV

minor histocompatibility antigen mismatched BMT model, simulates female→male HLA-identical SCT which is associated with higher risk of GVHD development.\(^{(329,330)}\)

Exploratory analysis of the data revealed a remarkable heterogeneity in the gene expression of donor CD8\(^+\) T\(_{\text{eff}}\) cells. At D+7 post-transplant, a clear dichotomy between the transcriptional profiles of SLO and GVHD target organs was identified, underlining a distinct T\(_{\text{eff}}\) cell differentiation pattern according to their location (Figure 22 and Figure 24). These differences were found to be independent of the TCR repertoire and of antigen distribution, and reflected an enrichment for effector function related genes in the peripheral tissues (Figure 26 and Figure 28).

Such findings challenge the SLO-centric classical scheme of acute GVHD development,\(^{(331,332)}\) which envisages the alloreactive response as a sequence of events: (1) priming of the immune response, entailing APC activation and maturation; (2) donor T cell activation, proliferation and differentiation; (3) migration of T\(_{\text{eff}}\) cells and other effector leukocytes to GVHD target tissues; and (4) target tissue destruction by T\(_{\text{eff}}\) cells.\(^{(81,90,91)}\) According to this model, the full process of effector differentiation would be determined by priming, so T cells entering the peripheral tissues would already be hard-wired to cause tissue damage. However, contrary to what would have been anticipated, a great diversity in T\(_{\text{eff}}\) cell transcriptional profiles was observed among the tissues which could not be explained by an early egress of effector cells from the SLO to the GVHD target organs (Figure 32), and were not indicative of homing imprinting idiosyncrasies (Figure 33).
Thus far, transcriptome data analysis in GVHD has been focused on the blood and SLO, and has been based on gene expression level comparison between groups followed by description of the gene ontology of the transcripts found to be significantly differentially expressed.\(^{298,333-335}\) However, this methodology underestimates the differences in the system, disregarding any change that does not meet the preconceived thresholds. Moreover, by failing to recognise patterns of correlated gene expression, it lacks insight into important biological aspects of the disease pathogenesis.

WGCNA is a powerful analytical tool which groups genes into modules on the basis of their coexpression similarities.\(^{228}\) This systems biology methodology has been proven effective in identifying functionally related genes, hence enabling a more systematic and global interpretation of gene expression data.\(^{230,231}\)

In the present study, 31 modules of genes were identified by WGCNA and, remarkably, their value in representing gene expression changes occurring during alloreactive response development in independent models of GVHD was largely validated. Not only were the majority of the modules shown to be preserved in a second murine model of GVHD (Figure 39), but, more importantly, the ones with the strongest preservation statistics accurately segregated synBMT from alloBMT samples (Figure 40), and identified those with GVHD (Figure 41), in published data sets from Rhesus macaques and human samples.\(^{298}\) Of note, a statistically significant association was found between several SLO-specific modules and blood samples from GVHD.
affected subjects. In contrast, no GVHD target organ-specific modules were enriched in the blood from either rhesus macaques or humans.

Due to the paucity of published transcriptional data on T_{eff} cells infiltrating the peripheral tissues, validation of the GVHD target organ-specific modules required a collaboration with Dr Laura Jardine and Dr Mathew Collin from the Newcastle University Institute of Cellular Medicine who provided paired samples of CD8^{+} T cells isolated from blood and skin biopsies of patients with GVHD. Likewise, SLO-specific modules were found to be enriched mainly in the blood samples, while the GVHD target organ-specific modules were over-represented in the epidermis (Figure 43).
1.2 A modular conception of the cellular programs underpinning GVHD development

Analysis of SLO- and GVHD target organ-specific modules has broadened the understanding of GVHD biology, providing valuable insight into cellular programs reported to be associated with GVHD in previous studies, and aiding to the mapping of the alloreactive response.

For instance, module 1, associated with the blood and bone marrow, was related to the regulation of the cell cycle, and included Cdk1, Aurka and Aurkb as driver genes, lending support to the concept proposed by Li et al. that cyclin-dependent kinases are required for alloreactive antigen-specific T cell clonal expansion,\(^{336}\) and to Furlan et al. observation that aurora kinase control of T cell turnover is dysregulated during GVHD.\(^{298}\) Furthermore, the LN associated module 17 was linked to Toll like receptor and retinoic acid-inducible gene 1 signalling, two extensively studied type I IFN inducing pathways recognised in the literature as playing an important role in the pathogenesis of GVHD.\(^{104,337}\) Interestingly, modules 3 and 4, which are mostly related to oxidative phosphorylation and fatty acid metabolism, were positively correlated with T\(_{\text{eff}}\) cells from the spleen, bone marrow and lymph nodes, but negatively correlated with T\(_{\text{eff}}\) cells from the gut and skin, suggesting that the dependence
of alloreactive T cells upon fatty acid oxidation described by Byersdorfer et al. may be limited to the SLO.\(^{338}\)

In contrast to the priming and proliferation programs that characterised the SLO, GVHD target organ-associated modules were mostly involved in pro-inflammatory cytokine generation. Notably, the pan-GVHD target organ associated module 29 included multiple transcription factor genes that regulate Th17/Tc17 fate, e.g. *Fosl2, Kdm6b, Skil* and *Chd7*.\(^{339-341}\) Additional Th17/Tc17-related transcription factors, such as *Ikzf3*, were identified as driver genes of module 31, reinforcing the role of Tc17 differentiation program during GVHD, particularly in the gastrointestinal tract.\(^{287}\) On the other hand, in the skin, the epidermis-specific module 28 driver genes were predominantly linked to T cell activation (*Il2ra, Tnfs9, Cd44*), Th1-type pro-inflammatory cytokine generation (*Ifng, Csf1* and *Csf2*), glycolysis (*Pkm2*), Notch signalling (*Rpβj, Furin*), and survival (*Bcl2l1*).

The diversity of transcription profiles and the organ specificity of cellular programs in GVHD uncovered by this novel approach suggests that T cell function is driven by compartment-specific interactions occurring at each site. However, further work is required to understand the regulatory mechanisms responsible for driving this spatial diversity of effector T cell programs. One attractive research avenue would be to integrate the gene expression data obtained in this study with multi-omic data from public repositories to define gene regulatory networks and develop working models capable of explaining and predicting gene expression, and simulating the effect of specific changes *in silico*.\(^{342}\) For instance, the data from published ChIP-seq studies now
available as libraries\textsuperscript{(343,344)} containing information on transcription factors, their genomic binding sites and DNA binding profiles, could be used to probe the gene modules and identify the transcriptional networks that regulate each program. Additionally, it would be of great value to analyse the DNA and histone modifications associated with each compartment and study their relationship with the patterns of program activation/silencing observed, considering the emergent information concerning the characterisation of the epigenomic landscapes in primary cells and tissues.\textsuperscript{(345)}
1.3 A “two-hit” model of cutaneous acute GVHD development

To investigate the mechanisms of induction of the tissue-specific programs, the study was focused on the skin as it displayed one of the most striking differences between compartments, and its associated modules, in particular module 28, were also enriched in samples collected from human GVHD patients.

The key histopathological findings that characterise acute cutaneous GVHD include a sparse to extensive lymphocytic infiltrate which commonly involves the epidermis and adnexal structures, accompanied by hydric changes to the basement membrane and keratinocyte necrosis. Similar changes were observed in the MataHari T cell model of GVHD, with donor CD8+ T eff cells spanning the dermis-epidermis interface, in close proximity to the radio-resistant host-derived LC (unpublished data).

It has long been recognised that host APC are required for the induction of acute GVHD, while donor derived APC are mostly involved in perpetuating tissue injury. However, in the particular case of the skin, this subject is still controversial: while Merad et al. showed that host LC were sufficient to cause skin GVHD, Li et al. demonstrated that LC were in fact
not required to induce GVHD, as other subsets of APC were equally able to prime T cells.\(^{(215)}\) However, in both studies the observations reported were based on experimental models in which the normal skin immune niche had been extensively remodelled. In the former, by using MHC-mismatched donor → host BM chimeras as recipients for a second GVHD-inducing transplant, Merad \textit{et al.} created a state in which the only residual recipient APC at the time of the second transplant were the LC. On the other hand, Li \textit{et al.} used a transgenic model in which LC are constitutively depleted (Langerin-DTA system), disregarding any changes that this determined to the skin immune baseline.

In contrast, the Langerin-DTR model developed by Dr Clare Bennett allows for the selective depletion of LC when recipients are treated with DT at least 21 days prior to transplantation.\(^{(216)}\) By allowing 3 weeks of rest, all other populations of Langerin\(^+\) cells, including the Langerin\(^+\) dermal DC, are regenerated and have reached steady state equilibrium at the time of BMT.\(^{(323,324)}\) Using these transgenic mice as recipients of minor histocompatibility antigen mismatch BMT (MataHari T cell model), it was demonstrated that host LC were indeed required for donor CD8\(^+\) T\(_{\text{eff}}\) cell accumulation in the epidermis, while host Langerin\(^+\) dermal dendritic cells were found to be redundant.\(^{(217)}\) Moreover, it was observed that the severity of cutaneous GVHD was greatly reduced when LC were absent, even when only localised LC depletion was performed.\(^{(217)}\)

These results suggest that \textit{in situ} interactions between donor CD8\(^+\) T\(_{\text{eff}}\) cells and host LC were pivotal for the regulation of tissue damage in skin GVHD.
This concept that T cell effector function is enhanced in the peripheral tissues through antigen-dependent interactions with resident DC conforms to the “two-hit” model of the immune response which has already been well-documented in several studies on inflammation and infection.\(^{(222-225,322)}\) It was thus hypothesised that the differences detected in the transcriptional profiles of \(T_{\text{eff}}\) cells isolated from the two skin compartments, dermis and epidermis, were dictated by LC.

Indeed, with LC depletion the contrast between the gene expression profile of dermal and epidermal donor CD8\(^+\) \(T_{\text{eff}}\) cells was attenuated. It was found that migration of \(T_{\text{eff}}\) cells from the dermis to the epidermis enhanced their pathogenicity and was linked to the expression of a broad effector program, characterized by cell cycle commitment, mRNA synthesis and increased metabolic activity (Figure 50). However, in the absence of LC, donor CD8\(^+\) T cells infiltrating the epidermis were rendered incapable of up-regulating the full array of effector genes (Figure 51), had a reduced survival (Figure 54), and showed impaired differentiation into resident memory cells (Figure 57).

Consistent with the hypothesis of LC–\(T_{\text{eff}}\) cell crosstalk, a complementarity of signalling pathways was predicted from the gene expression profiles of the two cell types: \(Ifng\) was a driver gene of module 28 in \(T_{\text{eff}}\) cells; and multiple IFN-\(\gamma\) responsive genes were up-regulated in host LC in the context of alloBMT. Despite the caveats of the model used (\(Ifngr1^{-/-}\) mice), pilot experiments showed that in the absence of IFN-\(\gamma\) receptor signalling, LC failed to up-regulate H-2Db expression and \(T_{\text{eff}}\) cell accumulation in the epidermis was significantly reduced, even though donor T cell expansion in the
SLO and infiltration of other peripheral tissues was not affected (**Figure 59 and Figure 61**).

Furthermore, in line with the results from various groups which reported Notch signalling as being central in the development of alloreactive T cells,\(^{318-320}\) particularly for IFN-γ production,\(^ {321}\) two key regulators of the Notch pathway (*Rbpj* and *Furin*) were identified as possible drivers of module 28 expression in the epidermis. These observations were validated by Notch signalling blockade experiments through treatment with a GSI, which induced a significant reduction of IFN-γ production in the skin (**Figure 63**). By delaying GSI treatment until D+5, preservation of priming events and migration of T cells to the peripheral tissues was achieved, as no differences in donor CD8\(^+\) T\(_{eff}\) cell accumulation was observed in any of the studied tissues.

It is recognised that the models used to test the requirement of host IFN-γ responsiveness and Notch signalling are non-specific, and that further work is needed to ascertain their relevance in LC regulation of donor T\(_{eff}\) cell pathogenicity. However, given that the results consistently mirror the changes observed when LC are depleted, it is proposed that these are instrumental pathways central to the LC–T\(_{eff}\) cell crosstalk.
1.4 Study limitations & Future work

Although the paramount role of experimental models in the progress of our understanding of immune system complexity is undeniable, and in providing important insight into the pathophysiology of immune mediated diseases, caution must be taken when interpreting the results obtained in animal studies. In the particular case of GVHD, to draw conclusions from murine models and translate them into the scenario of clinical GVHD, it is necessary to consider 5 potential caveats: (90,191)

a) Differences in conditioning regimens used to prepare the recipient. While in humans pre-transplant conditioning regimens consist mainly of chemotherapy, with or without low rate-delivered fractioned TBI, in mouse models of GVHD irradiation alone is typically used with large fraction doses and high dose rates. This disparity in conditioning regimens results in important differences in the extent of the tissue damage and proinflammatory reaction, which may influence the GVHD phenotype. In all BMT experiments performed in this project, a split dose of irradiation was delivered to the recipients with a 48h interval between fractions. Moreover, characterisation of the GVHD phenotype was performed in both models to ascertain that it mirrored the clinical and histological features observed in human patients.
b) **Donor–recipient genetic and immunologic disparities.**

In humans, the majority of HSTC performed are matched by high-resolution HLA typing at the allele level for several MHC loci, but mismatched for miHAs. To reproduce these conditions, various MHC- and/or miHA-disparate models have been created through the combination of different inbred mouse strains. It has been shown that different combinations of strains result in a variety of GVHD phenotypes which depend greatly on the MHC type and immunodominant target antigens. Such differences are known to determine the TCR repertoire and polarisation of donor T cells.\(^{(351)}\) Awareness of these strain-specific variations is critical for the comparison of results between models. Moreover, mice and humans differ in many aspects of the immune system, not only in terms of cell populations present at each niche (e.g. the mouse skin is primarily populated by DETC, whereas in humans α/β T cells predominate), but also in the expression of co-stimulatory molecules and chemokine receptors (e.g. CD28 is expressed by nearly all mouse CD4\(^+\) and CD8\(^+\) T cells, while in humans it is restricted to only 80% of CD4\(^+\) and 50% of CD8\(^+\) T cells). Additionally, the non-haematopoietic compartment also displays important species-specificities (e.g. in contrast to mice, human endothelial cells express both MHC-I and MHC-II molecules constitutively, and are thus able to present antigen both to CD8\(^+\) and CD4\(^+\) T cells), which make some observations made in animal models difficult to translate into the human system.\(^{(352)}\)
c) **Source of donor immune cells.**

The graft for HSCT in humans is most commonly derived from mobilized stem cell products containing circulating donor immune cells. On the other hand, typically, in murine models spleen cells and/or lymph node cells are delivered together with the BM graft to ensure a sufficient dose of T cells for the induction of GVHD. Not only immune-cell populations from different sources may have different trafficking capacities, but, most importantly, differences in cell composition can have a great influence on GVHD phenotype. To mitigate these differences, in the BMT protocol used, the BM graft was solely enriched with purified CD4\(^+\) T cells and CD8\(^+\) T cells, in a 2:1 ratio.

d) **Differences in the microbial baseline.**

The kinetics and severity of GVHD, in particular intestinal GVHD, has been reported to be affected by the enteric pathogens.\(^{(347)}\) Notably, both the composition and the diversity of the enteric microbiome of humans and rodents are very different, as animals used in experimental models are, unlike humans, housed under specific-pathogen free conditions since birth. Moreover, transplanted patients are frequently required to undergo antibiotic prophylaxis and/or treatment which can alter significantly the enteric microbiome.

e) **Age of the donors and recipients.**

While the majority of HSCT are performed on adult humans, murine HSCT studies tendentiously use young adult mice, as older age in mice is associated with increased sensitivity to radiation, and reduced efficiency
of immune reconstitution after transplant. For this project, all transplant recipients were 10 - 20 weeks old, and those used as donors were 8 - 16 weeks old.

Regarding the systems immunology approach, although WGCNA is a powerful tool to analyse microarray data sets, it requires an adequate pre-processing and normalization of the expression data and it can be biased when dealing with technical artifacts, tissue contamination or poor experimental design.\(^{348}\) In the case of this project, it was necessary to exclude tissue specific genes from the analysis. Despite the efforts to avoid preconceptions, it is recognised that by relying on gene ontology annotations or pathway information to analyse the modules, the characterisation of the cellular programs is incomplete and may show a bias towards what is already known.\(^{349}\)

Although the SLO-associated modules were readily validated in a non-human model of GVHD and with human samples from transplanted patients with GVHD, testing of the value of GVHD target organ-associated modules was hindered by the lack of studies focusing on the peripheral tissues. The analysis performed in collaboration with the Newcastle University Institute of Cellular Medicine permitted confirmation that these modules were effective at distinguishing between CD8\(^+\) T cells isolated from the blood and skin of GVHD patients, but lacked the power to determine their correlation with individual compartments. Future work should therefore include follow-up studies designed with larger sample sizes comprising specimens from other GVHD target organs.
In terms of the characterisation of the tissue-specific cues that control donor T cell effector function, this work provides compelling evidence that in the skin LC play a central role in regulating the development of cutaneous GVHD. Although the experiments performed in the Ifngr1\(^{-/}\) model and with GSI treatment are insufficient to demonstrate the requirement of IFN-γ and Notch signalling in the LC–T\(_{\text{eff}}\) cell crosstalk, this work provides the framework for future studies in which these signalling pathways may be inhibited in inducible and cell selective systems.

Moreover, a better understanding of LC biology and their interactions with the remaining elements of the skin immune network will grant important insight into the mechanisms regulating the homeostasis of the skin’s immune niche. A prominent subject of interest that would merit exploring is their effect on the skin’s microbiota, as there is increasing data supporting a correlation between the risk of GVHD and the patient’s microbiome diversity.(100,188,350)

Lastly, the classical prophylactic approaches to acute GVHD, such as treatment with methotrexate and cyclosporine/tacrolimus, has proved to be of great value in the prevention of GVHD. However, prolonged non-specific immunosuppression can impair the GVL response and hamper the reconstitution of the immune system, increasing the risk of relapse of the underlying neoplastic disease and of infectious complications. By demonstrating that GVHD is defined by tissue-autonomous regulation of effector T cells, the data presented here establishes the rationale for future studies aimed at identifying tissue-specific molecular targets which may translate into the prevention of the development of severe GVHD in the clinic.
V. Conclusion
Conclusion

Current knowledge of GVHD biology is based on research mainly focused on the events occurring in the SLO, using classical analytical methods that employ preconceived thresholds. Their capacity to identify novel and biologically meaningful pathways and cellular programs involved in GVHD pathogenesis has thus been limited, and the translational clinical benefit has been reduced. The systems immunology methodology explored in this thesis represents an innovative approach to the study of GVHD that not only takes into account the variability in donor T cell expression programs determined by their location (SLO versus GVHD target organs), but also characterises the cellular programs expressed at each site based on the analysis of a weighed correlation network of genes.

It was demonstrated that (1) the transcriptional profiles of donor CD8$^+$ T$_{eff}$ cells in peripheral tissues (skin, gut, liver) were very distinct from those in the lymphoid organs (lymph node, spleen, bone marrow and blood), (2) these profiles diverged sharply between the different GVHD target organs, and (3) also differed between individual sub-compartments of single organs. This phenomenon was independent of the TCR repertoire, and did not reflect variations in antigen distribution, suggesting that T cell effector programs are driven by compartment-specific interactions occurring at each site.
In the particular case of the skin, LC–T\textsubscript{eff} cell crosstalk was shown to be central to the development of cutaneous acute GVHD, where responsiveness to IFN-\(\gamma\) was required for LC maturation, while Notch-dependent signals enhanced local cytokine production by T\textsubscript{eff} cells and Notch-independent signals to promoted T cell accumulation and resistance to apoptosis.

Collectively, these data demonstrate that GVHD is defined by tissue-autonomous regulation of effector T cells and provides a rationale for the development of precision therapies directed at blocking GVHD in individual tissues, so that by avoiding global immune suppression, these approaches may permit preservation of GVT effects.
VI. Acknowledgements
Acknowledgements

First and foremost, I want to express my sincere gratitude to my supervisors Professor Ronjon Chakraverty, Dr Clare Bennett and Professor Stephen Mackinnon for accepting me as their PhD student and giving me the opportunity to work in their laboratory; and to Bloodwise and Fundação para a Ciência e Tecnologia (FCT Portugal), for their financial support, without which this work would not have been possible. I am particularly thankful to Professor Chakraverty and Dr Bennett for the frequent thought-provoking discussions which so often led to new research avenues that were critical for the success of my project. Their continuous support, patience, motivation, and immense knowledge have been inspirational.

My gratitude goes out to all past members of the lab who welcomed me with open arms and helped with the transition from the clinics to the lab, but specially to Janani Sivakumaran for her warm-hearted enthusiasm for science and for setting the example of high standards research, to Thomas Conlan for having the patience to teach me and help me when I started in the lab, to Cara Lomas for her cheerful companionship, and to Sophie Ward for all the help during the long days of experiments. I extend a warm thank you to everyone at the Institute of Immunity and Transplantation, from Prof Chakraverty and
Dr Bennett’s research group and beyond, for all the input, help, discussions and friendship.

I would also like to pay a special tribute to Prof João Forjaz Lacerda, Dr Maria João Costa and Dr Maria de Lurdes Guerra, my mentors during my Haematology residency, for their guidance, friendship and incentive to pursue a physician-scientist career.

Finally, I must thank my family: my parents and my sister, for always believing in me and encouraging me to go further, even when it meant moving to a different country; my parents-in-law, for their unconditional support and accepting to give Snoopy a new home until we can be re-joined; and most of all, my wife Rita, for her unfaltering love and for being by my side all the way through this long journey.
VII. References
References


10. Fliedner TM, Körbling M, Calvo W, Bruch C, Herbst E. Cryopreservation of blood mononuclear leukocytes and stem cells suspended in a large fluid
A systems immunology approach to GVHD

References


A systems immunology approach to GVHD

References


140. New JY, Li B, Koh WP, Ng HK, Tan SY, Yap EH, et al. T cell infiltration and chemokine expression: Relevance to the disease localization in
A systems immunology approach to GVHD

References


transplantation-induced intestinal cell death and reduces GVHD. Blood 2015, Jul 23;126(4):437-44.


224. Dolfi DV, Duttagupta PA, Boesteanu AC, Mueller YM, Oliai CH, Borowski AB, Katsikis PD. Dendritic cells and CD28 costimulation are required to
A systems immunology approach to GVHD


266. [computer program]. GraphPad Software; USA.


A systems immunology approach to GVHD

References


257


A systems immunology approach to GVHD

Appendix

VIII. Appendix
Appendix A

CD containing:

- .pdf files of the Supplementary Figures
- .xlsx files of the Supplementary Tables

List of Supplementary Figures

Supplementary Figure 1. Evaluation of internal reproducibility: correlation between the pairs of biological replicates for each tissue.

Supplementary Figure 2. Direct acyclic graph for enriched GO categories for module 1 genes.

Supplementary Figure 3. Direct acyclic graph for enriched GO categories for module 2 genes.

Supplementary Figure 4. Direct acyclic graph for enriched GO categories for module 3 genes.
Supplementary Figure 5. Direct acyclic graph for enriched GO categories for module 4 genes.

Supplementary Figure 6. Direct acyclic graph for enriched GO categories for module 6 genes.

Supplementary Figure 7. Direct acyclic graph for enriched GO categories for module 7 genes.

Supplementary Figure 8. Direct acyclic graph for enriched GO categories for module 9 genes.

Supplementary Figure 9. Direct acyclic graph for enriched GO categories for module 13 genes.

Supplementary Figure 10. Direct acyclic graph for enriched GO categories for module 14 genes.

Supplementary Figure 11. Direct acyclic graph for enriched GO categories for module 15 genes.

Supplementary Figure 12. Direct acyclic graph for enriched GO categories for module 17 genes.

Supplementary Figure 13. Direct acyclic graph for enriched GO categories for module 23 genes.

Supplementary Figure 14. Direct acyclic graph for enriched GO categories for module 24 genes.
**Supplementary Figure 15.** Direct acyclic graph for enriched GO categories for module 26 genes.

**Supplementary Figure 16.** Direct acyclic graph for enriched GO categories for module 27 genes.

**Supplementary Figure 17.** Direct acyclic graph for enriched GO categories for module 28 genes.

**Supplementary Figure 18.** Direct acyclic graph for enriched GO categories for module 29 genes.

**Supplementary Figure 19.** Direct acyclic graph for enriched GO categories for module 30 genes.

**Supplementary Figure 20.** Direct acyclic graph for enriched GO categories for module 31 genes.

**Supplementary Figure 21.** Direct acyclic graph for enriched GO categories for module 1 driver genes.

**Supplementary Figure 22.** Direct acyclic graph for enriched GO categories for module 2 driver genes.

**Supplementary Figure 23.** Direct acyclic graph for enriched GO categories for module 3 driver genes.

**Supplementary Figure 24.** Direct acyclic graph for enriched GO categories for module 4 driver genes.
**Supplementary Figure 25.** Direct acyclic graph for enriched GO categories for module 6 driver genes.

**Supplementary Figure 26.** Direct acyclic graph for enriched GO categories for module 7 driver genes.

**Supplementary Figure 27.** Direct acyclic graph for enriched GO categories for module 9 driver genes.

**Supplementary Figure 28.** Direct acyclic graph for enriched GO categories for module 13 driver genes.

**Supplementary Figure 29.** Direct acyclic graph for enriched GO categories for module 14 driver genes.

**Supplementary Figure 30.** Direct acyclic graph for enriched GO categories for module 15 driver genes.

**Supplementary Figure 31.** Direct acyclic graph for enriched GO categories for module 17 driver genes.

**Supplementary Figure 32.** Direct acyclic graph for enriched GO categories for module 23 driver genes.

**Supplementary Figure 33.** Direct acyclic graph for enriched GO categories for module 24 driver genes.

**Supplementary Figure 34.** Direct acyclic graph for enriched GO categories for module 26 driver genes.
Supplementary Figure 35. Direct acyclic graph for enriched GO categories for module 27 driver genes.

Supplementary Figure 36. Direct acyclic graph for enriched GO categories for module 28 driver genes.

Supplementary Figure 37. Direct acyclic graph for enriched GO categories for module 29 driver genes.

Supplementary Figure 38. Direct acyclic graph for enriched GO categories for module 30 driver genes.

Supplementary Figure 39. Direct acyclic graph for enriched GO categories for module 31 driver genes.

Supplementary Figure 40. Direct acyclic graph for enriched GO categories for the genes in module 28 that failed to be upregulated in the epidermis in the DT treated group.

List of Supplementary Tables

Supplementary Table 1. List of the liver, small intestine and skin specific genes excluded from the analysis.

Supplementary Table 2. List of overrepresented KEGG pathways in the 19 highly preserved modules.
Supplementary Table 3. List of the driver genes for each of the 19 highly preserved modules.

Supplementary Table 4. List of KEGG pathways enriched in the dermis vs lymph nodes and in the epidermis vs dermis, in the presence (PBS) or absence (DT) of LC.

Supplementary Table 5. List of KEGG pathways overrepresented in the genes that failed to be up-regulated in the epidermis in the absence of LC.
Appendix B

During my PhD, I contributed with experimental work that led to the publishing of four scientific articles: