Data in Brief

Global transcriptome analysis of T-competent progenitors in the bone marrow

Vionnie W.C. Yu, David T. Scadden

* Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA, USA
b Harvard Stem Cell Institute, Cambridge, MA, USA
c Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

**ABSTRACT**

T cells are known to develop in the thymus. However, molecular events that control the transition from hematopoietic progenitor cells in the bone marrow to T precursor cells seeded in the thymus remained poorly defined. Our recent report showed that osteocalcin (Ocn)-expressing bone cells in the bone marrow have major impact on T cell immunity by regulating T progenitor development in the bone marrow (Yu et al., 2015)[1]. Selective endogenous depletion of Ocn+ cells by inducible diphtheria toxin receptor expression (OcnCre;iDTR) led to reduction of T-competent common lymphoid progenitors (Ly6D− CLPs) in the bone marrow and loss of T cells in the thymus. Expression of the Notch ligand DLL4 by Ocn+ cells in the bone marrow ensures the production of Ly6D− CLPs, and expression of chemotactic molecules CCR7 and PSGL1 to enable subsequent thymic seeding. These data indicate that specific mesenchymal cells in bone marrow provide key molecular drivers enforcing thymus-seeding progenitor generation and thereby directly link skeletal biology to the production of T cell based adaptive immunity. Here we present the transcriptome profiles of Ly6D− CLPs derived from Ocn+ cells deleted mice (OcnCre+;iDTR) compared to those derived from control littermates (OcnCre−;iDTR). These data are publically available from NCBI Gene Expression Omnibus (GEO) with the accession number GSE66102.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66102.

2. Experimental design, materials and methods

2.1. Mouse models

To specifically delete mature osteolineage cells, the OC-cre strain (from Dr. Thomas L. Clemens, Johns Hopkins University, MD, USA) [2], which expressed Cre under the Osteocalcin promoter was crossed to the iDTR strain [3]. Temporally controlled cell ablation was achieved upon injection of diphtheria toxin into the OcnCre+/−;iDTR strain [1]. To study the effect of Ocn+ cells on hematopoiesis, OcnCre+/−;iDTR injected with diphtheria toxin or OcnCre+;iDTR injected with PBS served as controls while OcnCre−;iDTR injected with diphtheria toxin were used as mutants. For most experiments, 25 ng diphtheria toxin in PBS/g of body weight was injected daily into both control and mutant animals for 28 days from 4 weeks of age to achieve an acute deletion of specific osteolineage populations. Mice were harvested for analysis the next day after the last dose of diphtheria injection. For all experiments, littermates were used as controls. All animal usage and procedures performed were approved by the Institutional Animal Care and Use Committee of Massachusetts General Hospital.
3. Flow cytometry

Before sacrifice, peripheral blood was collected from each mouse and subjected to complete blood cell count. For each mouse, tibiae, femurs, iliac crests, spines, ulnae, radii, and humeri were collected for bone marrow cells. In addition, spleen and thymus were also collected for lymphocyte staining. Changes in hematopoietic populations were quantified by flow cytometry. Bone marrow cells harvested from each animal were counted. We routinely stain 5 × 10^7 cells per sample for the hematopoietic stem population, and 1 × 10^7 cells per sample for each progenitor and mature population. Lineage cocktail consists of biotinylated B220, CD3e, CD4, CD8a, CD19, CD11b, Gr1, Ter119, CD11c, and NK1.1 antibodies. Fluorescence conjugated to streptavidin was used to recognize lineage cocktail. The following antibody combinations were used to recognize Ly6D^− CLP (Lineage-Pacific Orange, Sca-Pacific Blue, cKit-APC-Cy7, CD127-PE-Cy7, Thy1.2-FITC, Ly6D-APC).

4. Microarrays

Ly6D^− CLPs were sorted from 6 to 8 weeks old OcnCre;iDTR mutants and controls using the FACSAria following daily DT treatment for 4 weeks. RNA was extracted using the TRizol® (Invitrogen) according to the manufacturer’s instructions. Samples were processed by the NuGen Ovation V2 laboratory Process in the microarray core facility of Dana-Farber Cancer Institute. Briefly, first stand cDNA was prepared from total RNA using a DNA/RNA chimeric primer and a reverse transcriptase. The resulting double stranded cDNA with a unique heteroduplex at the 5’ end of the antisense strand were amplified using SPIATM amplification, a repeated process of SPIATM DNA/RNA primer hybridization, DNA replication, strand displacement and RNA cleavage which resulted in a rapid accumulation of cDNA with sequence complementary to the original RNA. The SPIATM amplified cDNA was purified using the Zymo Research DNA Clean & Concentrator™ system. The purified cDNA was fragmented through a chemical and enzymatic process and labeled via enzymatic attachment of a biotin-labeled nucleotide to the 3’-hydroxyl end of the fragmented cDNA. The biotinylated cDNA was added to a hybridization solution containing several biotinylated control oligonucleotides (for quality control), and hybridized to the Mouse430A microarray chip overnight at 45 °C. The chips were then transferred to a fluidics instrument that performs washes to remove cDNA that has not hybridized to its complementary oligonucleotide probe. The bound cDNA was then fluorescently labeled using phycoerythrin-conjugated streptavidin (SAPE); additional fluor were then added using biotinylated anti-streptavidin antibody and additional SAPE. Each cDNA bound at its complementary oligonucleotide was excited using a confocal laser scanner, and the positions and intensities of the fluorescent emissions were captured. These measures provide the basis of subsequent biostatistical analysis.

5. Quality assessment

Standard QA/QC analyses involved chip analysis with the assayQualityMetrics BioConductor package and found no significant quality issues with any of the chips, as determined by (among other methods) visual inspection, intensity distributions or RNA degradation plots. The data was background corrected and normalized with RMA (Robust Multichip Average) using the “affy” BioConductor package. Values in the data matrix represent log2 normalized intensity values.

6. Differential expression quantification and classification

The selected database comprises 6 expression measurements of 41,345 genes, 3 in the first condition and 3 in the second. The recorded expression values ranged between 4.023 and 10,198.29. Supervised learning or class prediction methods were used for molecular classification and pattern recognition. This analysis involves selecting the features (genes) most correlated with a phenotypic distinction of interest. These features or “marker genes” are biologically interesting in themselves but they can also be used as the input of a classification algorithm that uses existing “labeled” samples to build a model to predict the labels for future samples. Genes correlated with a binary class distinction, for example a morphological or clinical phenotype, is directly identified and selected by using a “distance” metric, for example t-test statistic = (mA − mB) / (sA + sB) [m and s are the means and std. dev. per class] or Signal to noise ratio = (mA − mB) / (sA + sB) [m and s are the means and std. dev. per class]. Our analysis generated

![Fig. 1. Heat map illustration of 30 differential genes among the 400 genes identified when comparing the transcriptome of T competent progenitors (Ly6D^− CLPs) isolated from Ocn^− deleted mice compared to control littermates. Red represents overexpression whereas blue means lower expression.](image)
Table 1
List of 30 genes overexpressed in mutant Ly6D− CLPs compared to control Ly6D− CLPs.

| Accession | T-test | P value | FDR  | Fold change | Description |
|-----------|--------|---------|------|-------------|-------------|
| 17523486  | 26.2253| 0.0156  | 1.1068| 1.1068      | Zfp105      |
| 17208361  | 19.9969| 0.0001  | 1.0156| 1.0156      |             |
| 17390695  | 15.5519| 0.0025  | 1.0156| 1.0156      | Sprr2k      |
| 17264448  | 13.6163| 0.0032  | 1.0156| 1.0156      | Arhgef15    |
| 17512177  | 13.3854| 0.0028  | 1.0156| 1.0156      | ENSMUST00000153288 |
| 17391347  | 13.4355| 0.0002  | 1.0156| 1.0156      | Nqhp1       |
| 17497417  | 12.2032| 0.0005  | 1.0156| 1.0156      | Olf710      |
| 17544210  | 12.0084| 0.0015  | 1.0156| 1.0156      |             |
| 17374443  | 11.084 | 0.002   | 1.0156| 1.0156      |             |
| 17533124  | 10.7079| 0.0006  | 1.0156| 1.0156      | Sisph9      |
| 17305243  | 9.913  | 0.0097  | 1.0156| 1.0156      | Sftpd       |
| 17254041  | 9.4857 | 0.0008  | 1.0156| 1.0156      | Cid2        |
| 17296891  | 9.2744 | 0.0077  | 1.0156| 1.0156      |             |
| 17307801  | 9.0039 | 0.0012  | 1.0156| 1.0156      | Gulo        |
| 17430674  | 8.9262 | 0.0024  | 1.0156| 1.0156      | ENSMUST00000152128 |
| 17312223  | 8.9021 | 0.0009  | 1.0156| 1.0156      | BC025446    |
| 17548581  | 8.7592 | 0.0016  | 1.0156| 1.0156      | Gm3257      |
| 17219374  | 8.6936 | 0.001   | 1.0156| 1.0156      | Alyre2f     |
| 17302297  | 8.456  | 0.0048  | 1.0156| 1.0156      | Pcdh17      |
| 17466729  | 8.4263 | 0.0056  | 1.0156| 1.0156      | 4921507P07Rik |
| 17475498  | 8.3725 | 0.0085  | 1.0156| 1.0156      | Cyp2F2      |
| 17520871  | 8.0847 | 0.0026  | 1.0156| 1.0156      | Acpl2       |
| 17432985  | 8.0258 | 0.0022  | 1.0156| 1.0156      | ENSMUST00000138492 |
| 17467548  | 7.9069 | 0.01    | 1.0156| 1.0156      | IgkL6-14    |
| 17232499  | 7.8339 | 0.002   | 1.0156| 1.0156      | Rsph4a      |
| 17550502  | 7.6586 | 0.0035  | 1.0156| 1.0156      | 17.94637    |
| 17286115  | 7.6359 | 0.0087  | 1.0156| 1.0156      | Phe3d1      |
| 17423885  | 7.5189 | 0.0032  | 1.0156| 1.0156      | Gm136       |
| 1743901   | 7.5064 | 0.0151  | 1.0156| 1.0156      | Tais1r3     |
| 17407496  | 7.2659 | 0.011   | 1.0156| 1.0156      | Ccr1        |

400 selected differential genes. Fig. 1 is a color map illustrating 30 genes overexpressed in mutant Ly6D− CLPs compared to control Ly6D− CLPs. Table 1 shows the mean, standard deviation, score, and fold change of the 30 genes.

Acknowledgments

The authors thank Vance Morgan at the Partners HealthCare Center for Personalized Genetic Medicine for coordinating microarray experiment and analysis. This work was supported by MGH Federal Share of the Program Income under C06 CA059267, Proton Therapy Research and Treatment Center (V.W.C.Y.), BD Biosciences Stem Cell Grant (V.W.C.Y.), Bullock-Wellman Fellowship Award (V.W.C.Y.), Tosteson & Fund for Medical Discovery Fellowship (V.W.C.Y.), and NHLBI grants HL097794 (D.T.S.), HL044851 (D.T.S.), and HL100402 (D.T.S). The authors declare no competing financial interests.

References

[1] V.W. Yu, et al., Specific bone cells produce DLL4 to generate thymus-seeding progenitors from bone marrow. J. Exp. Med. 212 (2015) 759–774.
[2] M. Zhang, et al., Osteoblast-specific knockout of the insulin-like growth factor (IGF) receptor gene reveals an essential role of IGF signaling in bone matrix mineralization. J. Biol. Chem. 277 (2002) 44005–44012.
[3] T. Buch, et al., A Cre-inducible diphtheria toxin receptor mediates cell lineage ablation after toxin administration. Nat. Methods 2 (2005) 419–426.