Data in Brief 3 (2015) 40–46

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Selective transcriptional regulation by Myc: Experimental design and computational analysis of high-throughput sequencing data

Mattia Pelizzola [a,*], Marco J. Morelli [a], Arianna Sabò [a], Theresia R. Kress [a], Stefano de Pretis [a], Bruno Amati [a,b,**]

[a] Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia (IIT), Via Adamello 16, 20139 Milan, Italy
[b] Department of Experimental Oncology, European Institute of Oncology (IEO), Via Adamello 16, 20139 Milan, Italy

ARTICLE INFO

Article history:
Received 29 October 2014
Received in revised form
2 February 2015
Accepted 2 February 2015
Available online 12 February 2015

Keywords:
B-cell lymphoma
Transcription
Post-translational histone modifications

ABSTRACT

The gene expression programs regulated by the Myc transcription factor were evaluated by integrated genome-wide profiling of Myc binding sites, chromatin marks and RNA expression in several biological models. Our results indicate that Myc directly drives selective transcriptional regulation, which in certain physiological conditions may indirectly lead to RNA amplification. Here, we illustrate in detail the experimental design concerning the high-throughput sequencing data associated with our study (Sabò et al., Nature. (2014) 511:488–492) and the R scripts used for their computational analysis.

* Corresponding author.
** Corresponding author at: Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia (IIT), Via Adamello 16, 20139 Milan, Italy.
E-mail addresses: mattia.pelizzola@iit.it (M. Pelizzola), bruno.amati@iit.it (B. Amati).

http://dx.doi.org/10.1016/j.dib.2015.02.003

Specifications Table

Organism/cell line/tissue	*Human (P493-6 B cells), mouse (Eμ-myc B cells, 3T9 fibroblasts)*
Sex	*Not applicable*
Sequencer or array type	*Illumina Hi-Seq 2000*
Data format	*Raw and analyzed*
Experimental factors	***Eμ-myc**: wild type B-cells (Control, "C"), Eμ-myc transgenic B-cells not yet transformed (Pre-tumoral, "P"), and lymphoma cells (Tumor, "T")* ***3T9.Serum**: 3T9 fibroblasts serum starved (t0 h) or released for 1 (t1 h) or 2 (t2 h) hours* ***3T9.mycER**: MycER-infected 3T9 fibroblasts untreated (0 hOHT) or treated for different periods of times with OHT (4, 8, 16 hOHT) to activate the MycER chimera.* ***P493**: P493-6 cells treated with Tetracycline (Myc transgene repressed) for 72 h (t0) and then released in fresh medium without Tetracycline (allowing expression of the transgene) for 1 h (t1 h), 24 h (t24 h) or several passages ("High Myc"). P493-6 cells treated with Tetracycline (Myc transgene repressed) plus OHT (endogenous Myc activated: "Low Myc")*
Experimental features	*Cells with the indicated genotype/treatment were used for ChIP-Seq (for Myc, RNAPII, H3K4me3, H3K4me1, H3K27ac), totRNA-Seq, 4sU-RNA-Seq or DNAse-Seq experiments, as reported*
Consent	*n/a*
Sample source location	*Milan, Italy*

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE51011.

2. Experimental design

The Gene Expression Omnibus (GEO) Series GSE51011 contains 94 high-throughput sequencing samples associated with the Sabò et al. study [1]. These samples cover different –omics datasets (ChIP-Seq for transcription factors and post-translational histone modifications, RNA-Seq, 4sU-Seq and DNAseI-Seq) produced in different organisms and biological systems. Some of the samples have to be used as references for other samples: for example, as inputs in the ChIP-seq peak calling procedure, or as baselines for the identification of differentially expressed genes. To help navigating through these data we collected the most relevant associated metadata in Table 1.

The different biological systems used in the study allowed the analysis of the effects of modulation of Myc levels in vitro and in vivo, both at physiological and pathological levels. In the Eμ-*myc* model [2], Myc overexpression was achieved in vivo specifically in the mouse B-cell compartment, where it causes lymphoma development. This model system gave access to primary wild type B-cells (Control, "C"), Eμ-*myc* transgenic B-cells not yet transformed (Pre-tumoral, "P"), and lymphoma tumoral cells (Tumor, "T"). Modulation of Myc expression in human B-cells was obtained in a time-controlled manner in vitro in the cell line P493-6 [3], harboring a tet-regulated Myc transgene. A line of mouse 3T9 fibroblasts was also used in which endogenous c-*myc* was modulated from low basal levels (in conditions of serum starvation) to mitogen-induced levels (upon serum stimulation). In the same cells, we expressed a conditionally active MycER chimaera, allowing us to induce active Myc at supra-physiological levels through administration of OHT to the culture medium.

3. Data analysis: Source code design and installation

In addition to the methods in the original publication [1], the source code used for the computational analysis of the high-throughput sequencing data is available as supplemental material of this manuscript.