Bioinformatic Analysis of Transcriptional Regulation by Nur77 in Central Nervous System and Immune System

Montserrat Olivares Costa (molivares1@uc.cl)
Pontificia Universidad Católica de Chile https://orcid.org/0000-0002-2182-2164

Fernando Faunes
Universidad Andrés Bello: Universidad Andres Bello

María Estela Andrés
Pontifical Catholic University of Chile: Pontificia Universidad Catolica de Chile

Research note

Keywords: Transcription factor, gene expression, Nur77, Databases

DOI: https://doi.org/10.21203/rs.3.rs-505939/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Objective

The objectives of this work were to find genes regulated by Nur77 in neurons and to evaluate the possible common role of this transcription factor in neurons and lymphatic cells using published experimentally generated databases of ChIP-Seq and a microarray. We also characterized Nur77 binding throughout the genome.

Results

We identified 113 Nur77 target genes in neuronal stem cells and 116 in neuronal cells. Cell adhesion and anchoring processes emerged as regulated by Nur77 in neurons and lymphatic cells. We found 9 common genes regulated by Nur77. Finally, we described a significant distribution of Nur77 binding sites in strong enhancers and active promoters. This work is a first step to understand the role of Nur77 and its common targets in neurons and immune cells.

Introduction

Nur77 (also known as NGFI-B, TR3, and NR4A1) (1–3), is a transcription factor encoded by an immediate-early gene, and an orphan member of the nuclear receptor superfamily (4,5). Nur77 has been largely studied in the immune system for its function inducing apoptosis of auto-reactive immature lymphocytes T and modulating the inflammatory response (6,7). In the Central Nervous System (CNS), Nur77 is widely expressed, particularly in brain nuclei that receive dopaminergic and noradrenergic neurotransmission, which regulates the expression of Nur77 (8). Brain pathologies characterized by an imbalance of catecholamine neurotransmission, such as anxiety, addiction, schizophrenia, and Parkinson disease are associated with changes in the expression of Nur77(9–11). Despite the growing amount of information about Nur77 in CNS, the lack of knowledge of Nur77 target genes in neurons hinders the study of its function. Here we analyzed public and experimentally generated databases to provide a list of Nur77-directly regulated genes in neurons.

Given that Nur77 is involved in the response to synaptic stimulation and control of metabolism in immune and nervous cells (12–15), we asked if Nur77 regulates the expression of a set of common genes in the nervous and immune systems. Our analysis revealed a group of genes that binds Nur77 on their promoters and cell adhesion and anchoring functions emerged as regulated by Nur77 in both cell types. We also characterized Nur77 binding sites throughout the genome, finding a significant distribution in strong enhancers and active promoters, reinforcing the function of Nur77 as a transcriptional activator. The work presented here is an approach pretending to guide the experimental focus regarding Nur77 investigation, solving in part the problem of lack of knowledge of Nur77 target genes and presenting new functions that can be attributed to this transcription factor in both the immune and nervous systems.

Methods
Data acquisition

We downloaded the ChIP-Seq peaks from Gene Expression Omnibus (GEO) corresponding to EGFP-Nur77 in K562 (GSE31363) (16) and endogenous Nur77 in NSC (GSM1603270) and NC (GSM1603273) (17). The microarray data was obtained from the work published by Chen et al. in 2014, (GEO Accession; GSE76805) (12). We used ENSEMBL gene annotation for human (hg19) and mouse (mm10) which were directly extracted as TxDb objects though GenomicFeatures and TxDb. Mmusculus.UCSC.mm10.ensGene R packages. Additional genomic annotations of different chromatin states and functional regions were extracted from UCSC Table browser, including the K562 Genome Segmentation by ChromHMM and Ensembl Regulatory Build.

Nur77 binding site characterization

The overlap between the different ChIP-Seq peaks and genomic features were calculated with the GenomicRanges R package using the function ndOverlaps. We intersected the genomic coordinates with K562 Genome Segmentation by ChromHMM and calculated the overlapping enrichment, by normalizing by total coverage of each chromatin state across the genome.

To obtain Nur77 target genes, we used the annotation of TSS and proximal promoters provided by the Ensembl Regulatory build for human (hg19) and mouse (mm10). The annotated TSS and proximal promoters were assigned to transcripts ID from Ensembl annotation if they were located within 2000 nt upstream and 500 nt downstream from a transcript start coordinate. Then, the function findOverlaps from GenomicRanges was used to find Nur77 ChIP-Seq peaks that overlapped with TSS or proximal promoters from the Ensembl Regulatory build. We used the biomaRt R package to find the common gene symbol associated with each of the Ensembl IDs that were assigned to each TSS and proximal promoter that overlapped with a Nur77 ChIP-Seq peak.

Microarray analyses

We computed the differentially expression results from the microarray data published by Chen et al. in 2014 (12), to filter genes that have an adjusted p-value lower than 0.05 and an absolute value of log2 fold change greater than 1. We plotted the results using ggplot2 and highlighting the names of genes for which a Nur77 peak were detected in their promoters.

Ontology analysis

We use the Gene Ontology Consortium server (www.geneontology.org) (18,19) and processed using default parameters with PANTHER analysis tool (20).

Results And Discussion

Nur77 target genes in neurons
We considered as Nur77 target genes, those genes that bind Nur77 in a window of -2kb to +500 bp from TSS in their promoters, according to our re-analysis of ChIP-Seq data of endogenous Nur77 of mouse neural stem cells (NSC) and NSC differentiated to neurons (NC) (17). To select genes in which the binding of Nur77 influences transcription, we used a microarray of mRNA from mouse hippocampal pyramidal neurons overexpressing Nur77 (12) and selected genes with a significant change of expression after Nur77 overexpression (Fig. 1A). We identified 113 Nur77 target genes in NSC (Table S.1) and 16 out these 113 genes changed their expression with a fold-change ≥ 2 after Nur77 overexpression (Fig. 1B). In NC, we identified 116 Nur77 target genes (Table S.2) and 17 out of these 116 genes changed their expression with a fold-change ≥ 2 after Nur77 overexpression (Fig. 1C). We found 53 Nur77 target genes common to NSC and NC, suggesting that these genes maintain Nur77 binding to their promoter during neuronal differentiation (Table S.3).

**Binding of Nur77 on promoters of genes in the immune system and the central nervous system**

To find out whether Nur77 exerts a similar function in the nervous and immune systems and its conservation in human and mouse, we compared the binding peaks from two high quality experimentally generated ChIP-Seq databases: 1.- ChIP-Seq from the ENCODE project of overexpressing EGFP-Nur77 in the human chronic myelogenous leukemia cell line K562 (16) and 2.- ChIP-Seq of mouse NC (17) (Fig. 2A). Despite the differences between these two ChIP-Seq protocols, we found 271 genes with Nur77 common binding on their promoter (-2kb to +500bp from the TSS) in immune and neuronal cells (Fig. 2A, Table S.4).

GO analysis of these 271 genes (Fig. 2B and Table S.5) showed a significant enrichment of Nur77 target genes in ribonucleoprotein complex binding (fold of enrichment 5.2), cadherin bindings (fold of enrichment 3.94), cell adhesion molecule binding (fold of enrichment 3.05) and protein domain specific binding (fold of enrichment 2.66) (Fig. 2B). In the cellular component classification, Nur77 target genes were mainly enriched in categories of adhesion and junction: focal adhesion (fold of enrichment 3.45), cell-substrate adherent junction (fold of enrichment 3.43), cell-substrate junction (fold of enrichment 3.39), adherens junction (fold of enrichment 3.02), and anchoring junction (fold of enrichment 2.93). Two categories of nuclear localization were enriched: nuclear speck (fold of enrichment 3.33) and nuclear body (fold of enrichment 2.86) (Fig. 2B). Many proteins encoded by Nur77 target genes, which are common to the nervous and immune system, are ribonucleoproteins and adhesion molecules. This fact was strengthened by the enrichment of Nur77 target genes in nuclear bodies and areas of adherents and anchoring junctions (Fig. 2C).

In the biological process classification, Nur77 target genes were enriched in regulation of Endoplasmic-Reticulum-Associated protein Degradation (ERAD) pathway (fold of enrichment 12.91) and regulation of response to endoplasmic reticulum stress (fold of enrichment 8.01). Three GO terms related to interleukin signaling were enriched: interleukin-12-mediated signaling pathway (fold of enrichment 10.11) cellular
response to interleukin-12 (fold of enrichment 9.68) and response to interleukin-12 (fold of enrichment 9.49). Nur77 target genes were also enriched in the regulation of protein autophosphorylation (fold of enrichment 9.49) and cell aging (8.22) GO terms (Fig. 2B).

Interestingly, 9 out of the 271 genes were also found in the set of genes that significantly changed their expression after Nur77 overexpression in neurons: AGAP3, BIRC5, DYM, ITGB3, KIF21B, MORN5, RREB1, STRIP2 and WEE1, suggesting that these genes are also regulated in the immune system. Previous evidence supports that Nur77 controls the expression of BIRC5 gene (21), validating our results. Further studies are required to fully validate Nur77 control over these genes, both in the nervous and immune system.

**Characterization of Nur77 binding sites throughout the genome.**

To further characterize the binding profile of EGFP-Nur77, we compared the enrichment of ChIP-Seq peaks across 15 chromatin states defined by Ernest et al. for K562 cell line with the peaks of Nur77 (22). We calculated the overlap enrichment across the EGFP-Nur77 peaks and chromatin states for K562 cell line, obtaining a significant enrichment of Nur77 peaks across chromatin states associated with transcription (Fig. 3A). The analysis of the coordinates of EGFP-Nur77 peaks, reported by ENCODE, with respect to the nearest transcription start site (TSS), revealed a high frequency of Nur77 binding between +1000 and −1000 nucleotides from TSS (Fig. 3B).

An enrichment analysis using the human chromatin segmentation model generated by Ernst, (2011) showed significant enrichment of Nur77 peaks in chromatin states associated with transcriptional regulatory regions, particularly in strong enhancers and active promoters (Fig. 3C). Two of chromatin states are described as strong enhancers (states 4 and 5), which differentiate in the occurrence of specific chromatin marks and distance to the TSS. Strong enhancer state number 4 presented a higher occurrence of histone 3 lysine 4 trimethylation (H3K4me3) and histone 3 lysine 9 acetylation (H3K9ac) and was closer to TSS than Strong enhancer state number 5. Nur77 was enriched in both chromatin states described as strong enhancers, exhibiting a log2 enrichment greater than 4 in chromatin state 4 (Fig. 3C). High enrichment of Nur77 binding was also observed in transcriptional transition states of chromatin. These areas presented similar characteristics to transcriptional elongation areas, but with an increased presence of H4K20me1 and H3K4me1 and more sensitive to DNAse (22), suggesting an intermediate state between the promoter activation and effective elongation. In contrast, we found a negative enrichment for Nur77 binding in transcriptional elongation areas. Nur77 was poorly enriched in states numbers 6 and 7, both described as weak enhancers which differ in the occurrence of H3K4me2 and DNase sensitivity (22). Finally, negative enrichment was observed in heterochromatin regions, indicating the absence of Nur77 in inactive areas of the chromatin (Fig. 3C).
Altogether these data indicate that Nur77 is mostly associated with active sites of chromatin, concordant with the role of Nur77 as a transcriptional activator, also confirmed by the high presence of Nur77 in the TSS. Our data also shows that Nur77 binds transcriptional transition areas, suggesting that Nur77 is present in the promoters of its target genes independently of their state (active or inactive). On the other hand, the data suggests that the presence of Nur77 in enhancers would be limited to the active state.

In conclusion, our data analyses show that Nur77 is bound to the promoter of its target genes independently of their transcriptional state (weak, poised or active). In addition, our data suggests that the presence of Nur77 in enhancers is limited to the active state. We propose that Nur77 is always present in promoters but only binds to enhancers when it is upregulated, modulating thus transcription in response to stimuli.

Our analysis showed a strong participation of Nur77 target genes in anchoring and adhesion functions, which is consistent with the previously described roles of Nur77 in the modulation of neurite growth in neurons (13), and in the immune response (14,15), both processes that require interaction between cells and with the extracellular matrix.

Finally, genes found in this work as common targets of Nur77 in the nervous and immune systems are new and undescribed targets of this transcription factor. The work presented here is an approach pretending to guide the experimental focus regarding Nur77 investigation, solving in part the problem of lack of knowledge about Nur77 target genes and presenting new functions that can be attributed to this transcription factor in both the immune and nervous systems.

**Limitations**

The work presented here is a re-analysis of previously validated databases. However, differences in protocols or the overexpression of Nur77 could generate biases in the analyses. To be sure that genes described here are really modulated by Nur77, we were very restrictive in the selection process, this could lead to an underrepresentation of all genes regulated by Nur77 in neurons.

**Abbreviations**

ChIP-Seq: Chromatin immunoprecipitation followed by sequencing; CNS: Central nervous system; NC: Neuronal cells; NSC: Neuronal stem cells; GO: Gene ontology; TSS: Transcription start site.

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**
Not applicable

Availability of data and materials

The datasets analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository, https://www.ncbi.nlm.nih.gov/geo/. Access codes are detailed in the manuscript.

All data generated during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by Fondecyt grants 1150200 and 1191152. MO was supported by a Conicyt doctoral fellowship number 21140438

Authors' contributions

MO, Design of the study, analysis and wrote the paper. FF, Discussed the results and wrote the paper. EA, Supervised the research, discussed the results, and wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

We thank Dr. Guillermo Parada for carrying out the bioinformatic analysis presented in this article.

References

1. Hazel TG, Nathans D, Lau LF. A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily. Proc Natl Acad Sci U S A. 1988.
2. Milbrandt J. Nerve growth factor induces a gene homologous to the glucocorticoid receptor gene. Neuron. 1988.
3. Chawnshang C, Kokontis J, Shutsung L, Yijan C. Isolation and characterization of human TR3 receptor: A member of steroid receptor superfamily. J Steroid Biochem. 1989.
4. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, et al. The nuclear receptor superfamily: The second decade. Cell. 1995.
5. Flaig R, Greschik H, Peluso-Iltis C, Moras D. Structural basis for the cell-specific activities of the NGFI-B and the Nurr1 ligand-binding domain. J Biol Chem. 2005.
6. Woronicz JD, Calnan B, Ngo V, Winoto A. Requirement for the orphan steroid receptor Nur77 in apoptosis of T-cell hybridomas. Nature. 1994;
7. McMorrow JP, Murphy EP. Inflammation: A role for NR4A orphan nuclear receptors. Biochemical Society Transactions. 2011.

8. Campos-Melo D, Galleguillos D, Sánchez N, Gysling K, Andrés ME. Nur transcription factors in stress and addiction. Frontiers in Molecular Neuroscience. 2013.

9. Xing G, Zhang L, Russell S, Post R. Reduction of dopamine-related transcription factors Nurr1 and NGFI-B in the prefrontal cortex in schizophrenia and bipolar disorders. Schizophr Res. 2006.

10. Lévesque D, Rouillard C. Nur77 and retinoid X receptors: crucial factors in dopamine-related neuroadaptation. Trends in Neurosciences. Trends neurosci; 2007.

11. Rouillard C, Baillargeon J, Paquet B, St-Hilaire M, Maheux J, Lévesque C, et al. Genetic disruption of the nuclear receptor Nur77 (Nr4a1) in rat reduces dopamine cell loss and L-Dopa-induced dyskinesia in experimental Parkinson’s disease. Exp Neurol. 2018.

12. Chen Y, Wang Y, Ertürk A, Kallop D, Jiang Z, Weimer RM, et al. Activity-induced Nr4a1 regulates spine density and distribution pattern of excitatory synapses in pyramidal neurons. Neuron. 2014.

13. Jeanneteau F, Barrère C, Vos M, De Vries CJM, Rouillard C, Levesque D, et al. The stress-induced transcription factor NR4a1 adjusts mitochondrial function and synapse number in prefrontal cortex. J Neurosci. 2018.

14. Cunningham NR, Artim SC, Fornadel CM, Sellars MC, Edmonson SG, Scott G, et al. Immature CD4 + CD8 + Thymocytes and Mature T Cells Regulate Nur77 Distinctly in Response to TCR Stimulation. J Immunol. 2006.

15. Moran AE, Holzapfel KL, Xing Y, Cunningham NR, Maltzman JS, Punt J, et al. T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse. J Exp Med. 2011.

16. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012.

17. Terranova C, Narla ST, Lee YW, Bard J, Parikh A, Stachowiak EK, et al. Global developmental gene programing involves a nuclear form of fibroblast growth factor receptor-1 (FGFR1). PLoS One. 2015.

18. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: Tool for the unification of biology. Nature Genetics. 2000.

19. Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, et al. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. 2019.

20. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019.

21. Yu C, Cui S, Zong C, Gao W, Xu T, Gao P, et al. The orphan nuclear receptor NR4A1 protects pancreatic β-cells from endoplasmic reticulum (ER) stress-mediated apoptosis. J Biol Chem. 2015.

22. Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011.

23. Hoffmann R. A wiki for the life sciences where authorship matters. Nature Genetics. Nat Genet. 2008.
Figure 1

Nur77 target genes in neurons. A) Pipeline of the experimental procedure of public databases (12,17) and the analysis to select Nur77 target genes. B,C) Volcano plots representing genes that change their expression when Nur77 is overexpressed in pyramidal neurons (blue), genes that also have a peak of binding of Nur77 in their promoter (red dots) and TSS (orange dots) according to the ChIP-Seq form NSC (B), and from NC (C). Positive numbers indicate overexpression and negative numbers down-regulation. The dotted lines demarcate Log2 Fold of Change = ±1.*
Figure 2

GO analysis of common target genes for Nur77 among the immune and nervous systems A) Pipeline of experimental procedure of public databases, and corresponding GO analysis. B) Enriched GO terms (more than 2.5 fold of change). C) The number of genes classified in each GO terms category. Molecular function classification in green bars, cellular component classification in blue bars and biological process classification in red bars. Cell adhesion molecule binding category includes cadherin binding GO term.
genes. Adherens and anchoring junctions category includes genes of adherens junction, anchoring junction, focal adhesion, cell-substrate adherent junction, and cell-substrate junction GO terms. Nuclear body category includes genes of nuclear speck GO term. Regulation of response to endoplasmic reticulum stress includes genes of regulation of ERAD pathway GO term. Response to interleukin-12 category includes genes of interleukin-12-mediated signaling pathways, cellular response to interleukin-12 and response to interleukin-12 GO terms (Table S.5).*
Nur77 binds to active promoters and strong enhancers in K562 cells. Analysis of publicly available database from the anti-EGFP ChIP-Seq experiment in K562 cell line overexpressing EGFP-Nur77 (16). A) Pipeline of experimental procedure of public database and bioinformatics analysis. B) Abundance of Nur77 binding respecting to the nearest TSS location. C) Enrichment of Nur77 binding sites across different chromatin regions defined by Ernst et al. in 2011 (22). In red, log2 of enrichment ≥ 2. In blue, log2 of enrichment ≤ -2. *

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- TableS.1.xlsx
- TableS.2.xlsx
- TableS.4.xlsx
- TableS.5.xlsx
- TableS3.xlsx