Supplementary Information

Early transcriptional changes induced by Wnt/β-catenin signaling in hippocampal neurons

Eduardo Pérez-Palma[1*], Víctor Andrade[1*], Mario O. Caracci[1], Bernabé I. Bustos[1], Camilo Villaman[1], Matías A. Medina[1], Miguel E. Ávila[1,2], Giorgia D. Ugarte[1] and Giancarlo V. De Ferrari[1] (gdeferrari@unab.cl).

This file contains:

- Supplementary Methods.
- Supplementary Table 1. Summary statistics of the samples sequenced in this study.
- Supplementary Table 2. Summary of sequence alignment using Bioconductor.
- Supplementary Figure 1. Experimental workflow procedures.
- Supplementary Figure 2. Read normalization analyses plots for RNASeq samples.
- Supplementary Figure 3. Overlap of early Wnt/β-catenin target genes.

Please note that Supplementary Tables 3 and 4 are presented online.

- Supplementary Table 3 (online). Differentially expressed genes in hippocampal neurons upon Wnt3a treatment (4 h).
- Supplementary Table 4 (online). Sub-networks components.

Supplementary Methods

RNASeq summary statistics and alignment to the rat reference genome. After whole-transcriptome sequencing of the 6 samples used in the study (3 Wnt treatment conditions and 3 controls), we obtained an average number of 100.62 million of pair end reads with 50.28% of GC content, 49.72% of AT content, a RNA integrity Number (RIN) of 8.5 and a sequencing quality (Q30) of 90.5% (Supplementary Table 1). The difference in the amount of sequencing reads between the third replicate (Ctrl3 and Wnt3) and the other 2 conditions was due technical reasons (higher depth obtained), nevertheless further quality controls (QC) including read count normalization across samples helped to reduce this difference in order overcome bias in the following functional analyses. After single-end read alignment with the rat reference genome (Rnor6) using R Bioconductor packages RSubread and Limma, we obtained an average number of 43.9 (83.81%) and 43.5 million (83.29%) of successfully mapped reads respectively (R1 and R2), resulting in an average mapping rate of 83.55% (Supplementary Table 2).

Quality control procedure of RNASeq samples. After sequence alignment, we followed the DESeq pipeline in order to perform quality controls to examine and subsequently reduce the heterogeneity present between experimental replicates. We performed a density plot to examine the distribution of reads between all samples, showing 2 different peaks without DESEq normalization process (Supplementary Figure 2a). After normalization, the peaks joint together showing a more homogeneous distribution of read densities across samples (Supplementary Figure 2b). A second approach was done using a boxplot distribution of reads, where before normalizing we can see a heterogeneous variance in sample W2 (wider box) and a shift in C3 and W3, due to the higher number of sequencing reads (Supplementary Figure 2c). After normalization, the variance between samples shrinks, indicating a reduction in heterogeneity (Supplementary Figure 2d). In order to see the effect of normalization in the differential expression results, we constructed a p value distribution plot (Supplementary Figure 2e), observing, as expected because of the short-term stimulation experiments, a conservative pattern where a minimum number of tests are showing statistical significance ($p < 0.05$).

Supplementary Table 1. Summary statistics of the samples sequenced in this study.

Identifier	Sample ID	Total Reads (Mb)	GC (%)	AT (%)	RIN	Q30 (%)
C1	Rn_C4h-25-3	56.79	50.08	49,92	8.2	89,39
C2	Rn_C4h-23-4	53.24	50.21	49,78	7.9	89,39
C3	Rn_C4h-22-12	189.71	50.50	49,50	10.0	92,36
W1	Rn_W4h-25-3	54.35	50.12	49,88	7.8	89,32
W2	Rn_W4h-23-4	55.58	50.31	49,68	7.2	89,12
W3	Rn_W4h-22-12	194.05	50.44	49,55	10.0	92,56

Supplementary Table 2. Summary of sequence alignment using Bioconductor.

Identifier	Sample ID	Input R1 (reads)	Mapped (reads)	Mapped (%)	Input R2 (reads)	Mapped (reads)	Mapped (%)	Mapping rate (%)
C1	Rn_C4h-25-3	28,396,511	24,073,712	84.78	28,396,511	24,001,972	84.52	84.65
C2	Rn_C4h-23-4	26,617,832	23,496,274	88.27	26,617,832	23,409,699	87.95	88.11
C3	Rn_C4h-22-12	93,919,684	80,368,624	85.57	93,919,684	79,539,537	84.69	85.14
W1	Rn_W4h-25-3	27,174,572	23,968,025	88.20	27,174,572	23,888,677	87.91	88.05
W2	Rn_W4h-23-4	27,792,483	23,953,882	86.19	27,792,483	23,834,804	85.76	85.97
W3	Rn_W4h-22-12	125,578,182	87,683,880	69.82	125,578,182	86,543,807	68.92	69.37

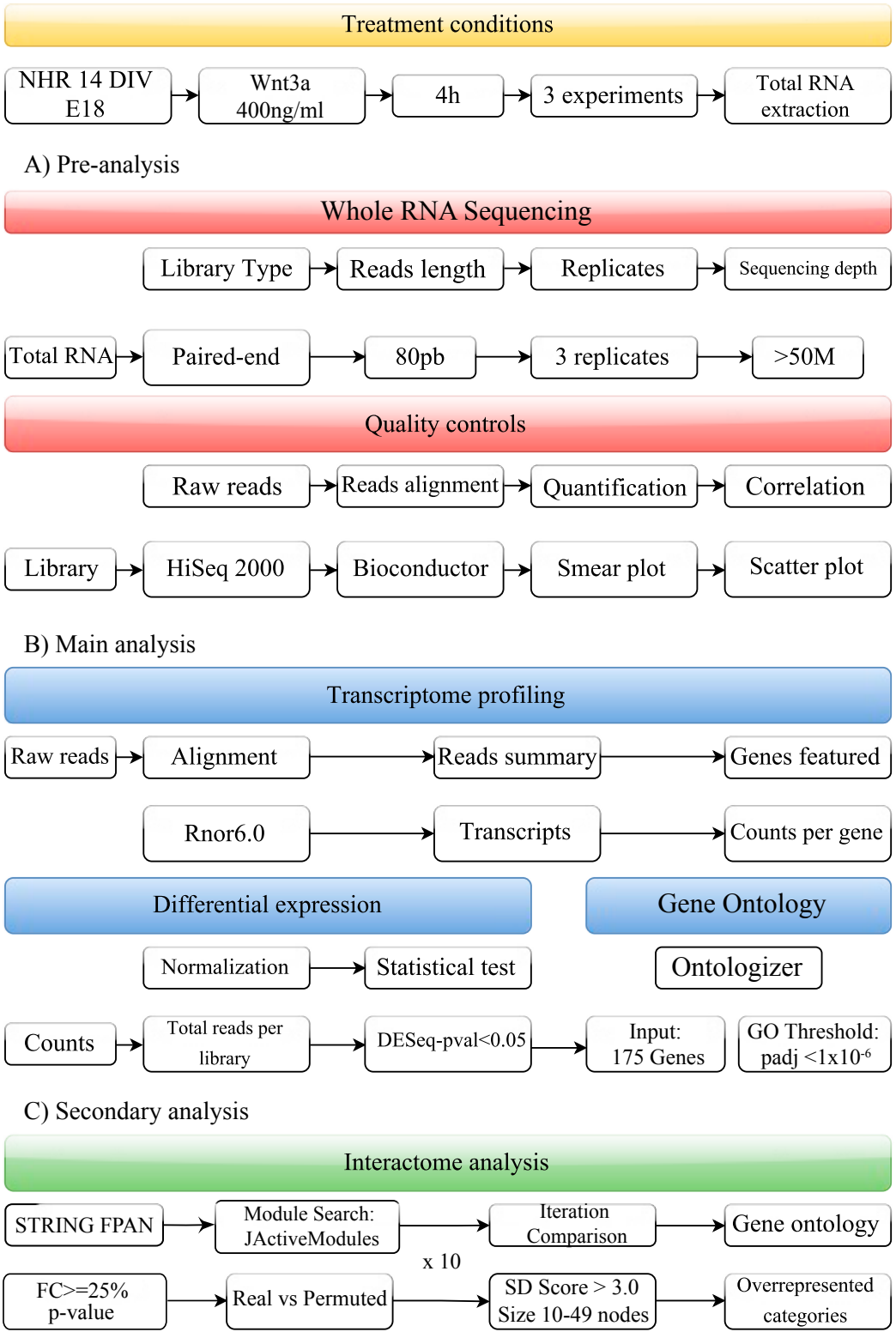
Supplementary Figure Legends.

Supplementary Figure 1. Experimental workflow. Each color box represents the 4 main steps of the study-highlighting key (white boxes) in a sequential order (arrow direction). 1) Treatment Conditions (Yellow): rat hippocampal neurons (RHN) were treated, in triplicate with Wnt3a (400 ng/ml), followed by total RNA extraction after 4h of incubation; 2) Experimental design for the whole RNA sequencing (Red): After libraries preparation with the TruSeq mRNA kit, fragments we sequenced at high throughput with the HISeq2000 sequencing, obtaining an average number of 100.62M of 80bp length paired end raw reads; 3) Transcriptome Profiling, Differential Expression and Biological networks analyses (Blue): Raw reads quality controls, alignment to the rat reference genome and posterior read variance normalization procedures were followed by differential expression analysis, resulting in a list of significant genes ($p < 0.05$). Then a gene ontology analysis (GO) was done using the Ontologizer software in order to select overrepresented ontological categories; 4) Interactome Analysis (Green): Analysis of all up-regulated gene expression signals in a rat network context (FPAN) alongside a Permuted analysis as a negative control. Module search was performed with JActive Modules over the FPAN to identify activated subnet-works. Comparison with Permuted analysis allowed the identification of true significant modules. Finally, a GO enrichment analysis was performed for the identified modules.

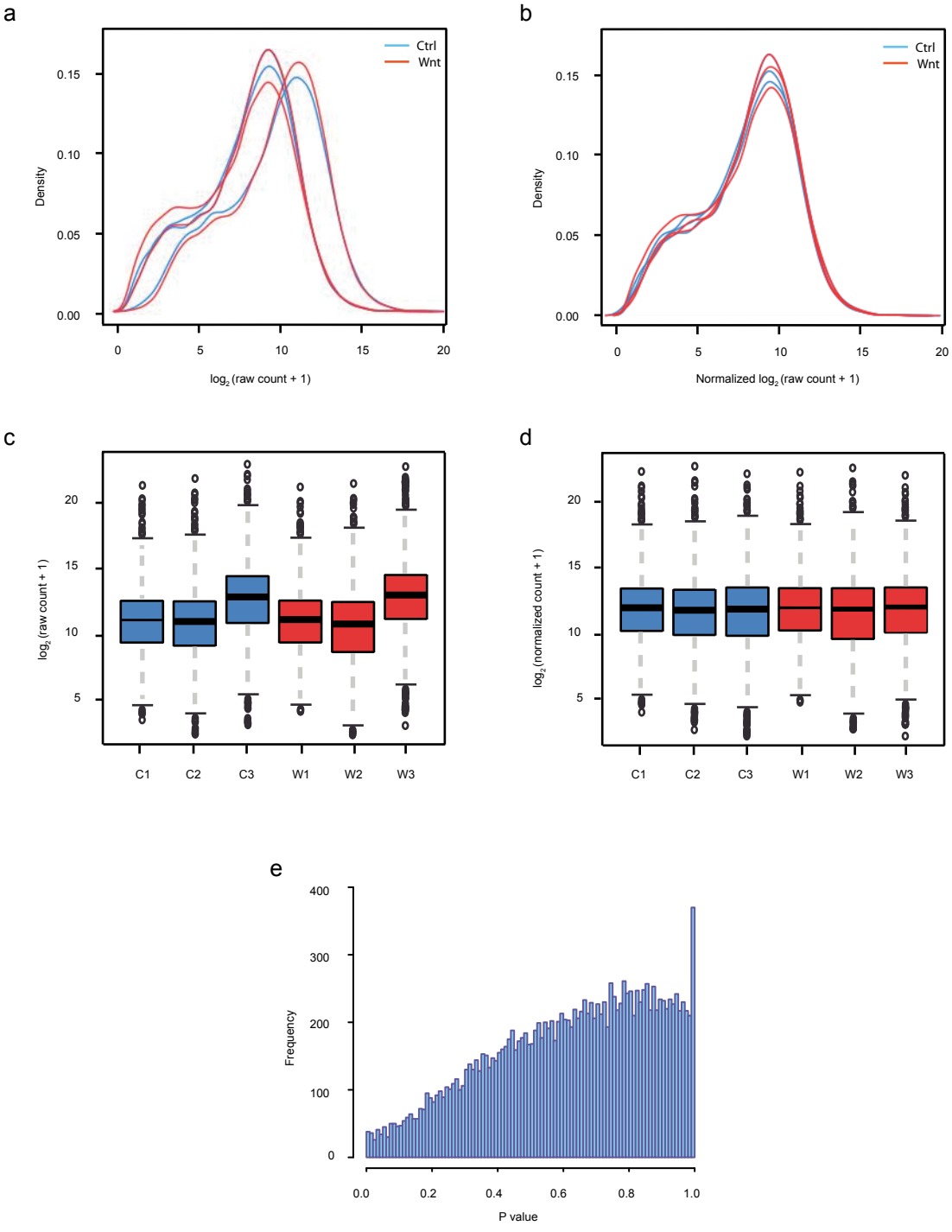
Supplementary Figure 2. Read normalization analyses plots for RNASeq samples. Read counts density plots showing read distribution across all the samples before (a) and after (b) normalization. Vertical axis shows read count densities for all samples and the horizontal axis shows log2 +1 read counts. (c-d) Box plot distribution of reads before (c) and after (d) normalization. Vertical axis shows log2 +1 read counts and each box plot shows the variance distribution for each sample. (e) P value distribution histogram showing differential expression significance tests after the normalization procedure. Vertical axis shows frequency of the p values in the differentially expression gene set.

Supplementary Figure 3. Overlap of early Wnt/β-catenin target genes. (a) Venn diagram depicting the overlap of our nominally significant genes ("This study", red circle, n=170) and curated direct Wnt/β-catenin target genes found in The Wnt Homepage (violet circle n=75) (http://web.stanford.edu/group/nusselab/ cgi-bin/wnt/target_genes) and Non-canonical Wnt signaling components extracted from the REACTOME database (blue circle, n=143) ("beta catenin independent Wnt signaling"; DOI: 10.3180/REACT_172694.1). (b) Venn diagram depicting the overlap between: i) Predicted Wnt/β-catenin target genes described by Hodar et al., 2010 (green circle, n=89); ii) Wnt1 transcriptional network observed in neural progenitor cells from Wexler et al 2011 (yellow circle, n = 294); iii) Genes with putative LEF1/TCF targets reported in Wisniewska et al., 2012 (cyan circle, n=428); and iv) Our nominally significant genes (red circle, n=170). For each comparison, details of the overlapping genes among the lists are provided in a color legend. Gray shading depicts the genes common for all lists assayed in the respective diagram. (c) Venn Diagram describing the overlap between our nominally significant genes ("This study", red circle, n=170) and the "second transcriptional wave" defined by the known target genes that could be activated indirectly through the 28 initial transcription factors (TFs) nominally expressed in our study (gray circle, n=99). "Second wave transcription" genes were retrieved from the curated database of human transcriptional targets TRRUST (http://www.grnpedia.org/trrust/) for 15 (AHR, ARX, DACH1, FOXO1, FOXO3, GATA2, GBX2, ID2, ID4, MSX2, NFAT5, NR4A2, TBX2, TBX3 and TFAP2C) out 28 TFs for which target information was available.

Supplementary Fig. 1



Supplementary Fig. 2



Supplementary Fig. 3