SUPPLEMENTARY MATERIAL

File 2

VALIDATION OF CONTROL DATA BY COMPARISON BETWEEN BIOPSIES FROM DECEASED LIVERS (CL) AND BIOPSIES FROM CONTROL LIVING LIVER DONORS (CLLD)

The choice of control livers in our study was very difficult. Literature data demonstrate that the integrity of mRNA is scarcely affected by sudden death without agonal state. In contrast agonal state preceding death has a substantial effect on gene expression (Li *et al.* 2004). For this reason we have chosen as control in our study a set of 5 deceased livers from individuals who suffered sudden death without agonal state (Khaitovich *et al.*, 2005). Expression data of these samples are publicly available at Array Express repository, accession number E-AFMX-11.

E-AFMX-11 Tissue Samples

All individuals, 3 males and 2 females, suffered sudden death for reason other than their participation to the study and without any relation to the tissues used. Age was ranging from 27 to 29 and was unknown in 2 cases. Total RNAs, isolated from 200 mg of frozen tissues using the Trizol reagent, were of high and comparable quality as gauged by the ratio of 28S to 18S ribosomal RNAs estimated using the Agilent 2100 Bioanalyzer (range 1.4-1.6) (Khaitovich *et al.* 2005).

GSE12720 Tissue Samples

As the use of control livers from subjects who suffered sudden death can be questioned we have compared our results to a set of expression data coming from microarray analysis (same technology and experimental conditions) of six liver biopsies from living donors (3 males and 3 females, age ranging from 29 to 50, mean 40) and very recently available from GEO repository [4, 10].

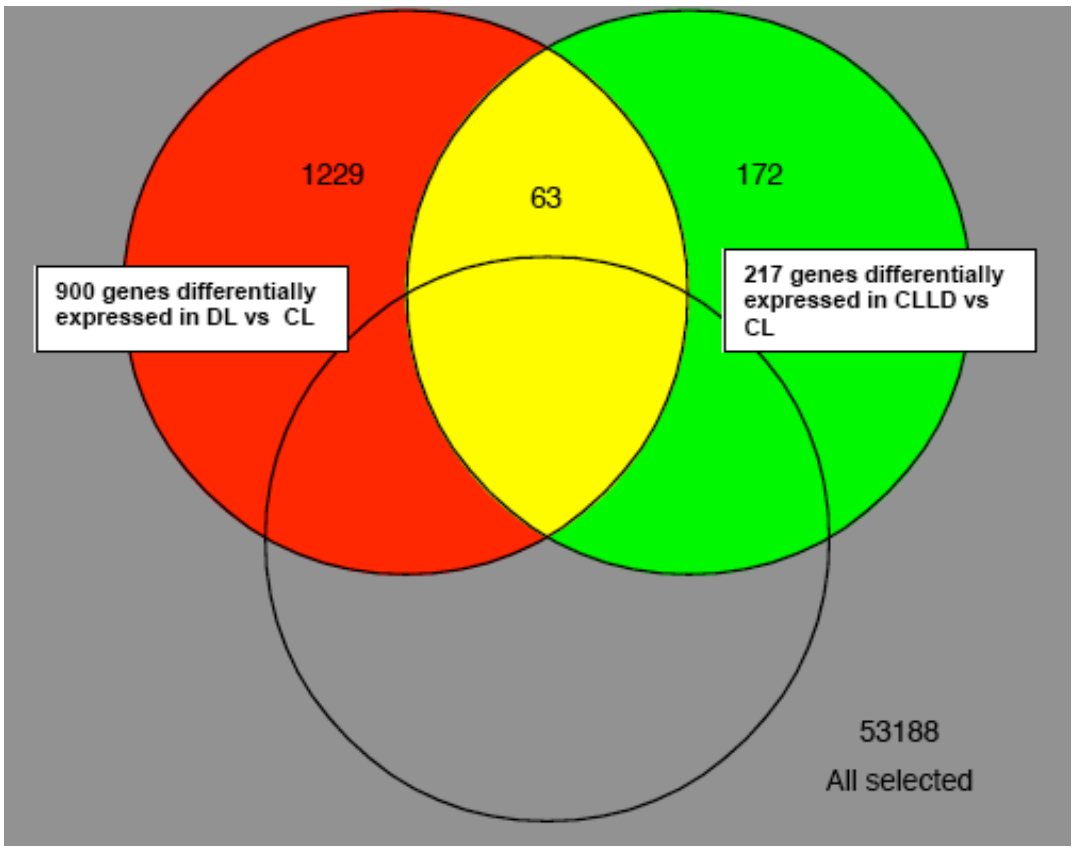
Biopsies were placed immediately in RNA Later (Ambion, Austin, TX) and frozen. Total RNA, extracted using Trizol (Invitrogen, Carlsbad, CA) was cleaned on RNEasy columns (Qiagen Inc, Valencia, CA). RNA integrity was confirmed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) (De Jonge *et al.* 2009).

All expression data were obtained by microarray analysis, using Affymetrix gene chip HG-U133 Plus 2. Hybridization data were normalized and quantified using Robust Multiarray Analysis (RMA) software (Wu *et al.* 2004). Expression data were considered reliable with a raw signal higher than 10.0. Raw expression data from RMA analysis were normalized per gene, dividing each measurement by the median of all measurements for that gene. Normalized data were log-transformed. Genes with a fold change > 1.75 were considered differentially expressed between CL and CLLD samples using the Student's t test with a p value <.01 and a Benjamini and Hoechberg false discovery rate <0.05.

RESULTS

Two hundred-seventeen genes were found differentially expressed between CL and CLLD samples with a fold change ranging from 1,75 to 10, mainly upregulated in CL. Only 20 genes show a fold change higher than 3. The VENN diagram in supplemental figure 2 shows that only 54 (63 probe sets) out of these genes are shared with the set of 900 genes differentially expressed between DL and CL groups. These genes, involved mainly in metabolic and catabolic processes, are described in Supplementary Table **1**.

Sixty-three probe sets, corresponding to 54 genes are overlapping, demonstrating that no more than the 6% of the genes dysregulated in the donor samples might be affected by the choice of control samples.



Supplementary Fig. (2). VENN Diagram showing the overlap of differentially expressed genes in CL *vs*. CLLD, compared to the list of 900 modulated genes, obtained between DL and CL.

Supplementary Table 1. List of 54 Genes already Dysregulated between CL and CLLD

Probe ID	Gene Name	Gene Bank	Description
209703_x_at	AAM-B	BC004492	DKFZP586A0522 protein
203682_s_at	ACAD2	NM_002225	isovaleryl Coenzyme A dehydrogenase
200727_s_at	ACTR2	AA699583	ARP2 actin-related protein 2 homolog (yeast)
220020_at	APP3	NM_022098	hypothetical protein LOC63929
236439_at	BCL6	AI733564	Homo sapiens transcribed sequence with weak similarity to protein pir:A40138 man
243820_at	BHLHB5	U80755	basic helix-loop-helix domain containing, class B, 5
48031_r_at	C5orf4	H93077	chromosome 5 open reading frame 4
225507_at	C6orf111	BF591408	chromosome 6 open reading frame 111
230375_at	C6orf111	AI936531	chromosome 6 open reading frame 111
223232_s_at	CGN	AI768894	cingulin
230180_at	DDX17	AA521056	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17
213998_s_at	DDX17	AW188131	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17
1559954_s_at	DDX42	AF147429	DEAD (Asp-Glu-Ala-Asp) box polypeptide 42
212107_s_at	DHX9	BE561014	DEAH (Asp-Glu-Ala-His) box polypeptide 9
233165_at	DIP	AJ242655	SH3 protein interacting with Nck, 90 kDa
213149_at	DLAT	AW299740	dihydrolipoamide S-acetyltransferase
236237_at	FLJ10980	AA526387	Homo sapiens transcribed sequences
1554514_at	FLJ20581	BC013753	hypothetical protein FLJ20581
204997_at	FLJ26652	NM_005276	glycerol-3-phosphate dehydrogenase 1 (soluble)
1555131_a_at	GIG13	BC026102	period homolog 3 (Drosophila)
214681_at	GK	AI830490	glycerol kinase
241955_at	HECTD1	BE243270	HECT domain containing 1
1557100_s_at	HECTD1	AL038005	HECT domain containing 1
242349_at	HECTD1	AW275658	HECT domain containing 1
226297_at	HIPK3	AV693403	homeodomain interacting protein kinase 3
211360_s_at	IP3R2	AB012610	inositol 1,4,5-triphosphate receptor, type 2
213446_s_at	IQGAP1	AI679073	IQ motif containing GTPase activating protein 1
224569_s_at	IRF2BP2	AW242432	interferon regulatory factor 2 binding protein 2
1552611_a_at	JAK1	AL555086	Janus kinase 1 (a protein tyrosine kinase)
212451_at	KIAA0256	N52532	KIAA0256 gene product
209254_at	KIAA0265	AI808625	KIAA0265 protein
231808_at	KRTAP4-7	AY007106	Homo sapiens, clone IMAGE:5302006, mRNA
212113_at	LOC552889	AI927479	Homo sapiens mRNA; cDNA DKFZp313P052 (from clone DKFZp313P052)
223766_at	LOC728407	AF130105	ARF GTPase-activating protein
239920_at	LOC728651	BF436302	upstream binding transcription factor, RNA polymerase I
242024_at	LUC7L2	T90999	Homo sapiens transcribed sequence with weak similarity to proteinens]
205187_at	MADH5	AF010601	MAD, mothers against decapentaplegic homolog 5 (Drosophila)
224568_x_at	MALAT1	AW005982	PRO1073 protein
223940_x_at	MALAT1	AF132202	PRO1073 protein
223578_x_at	MALAT1	AF113016	PRO1073 protein
213816_s_at	MET	AA005141	met proto-oncogene (hepatocyte growth factor receptor)
221661_at	NLT	AF210455	solute carrier family 22 (organic anion transporter), member 7
203242_s_at	PDLIM5	BG054550	LIM protein (similar to rat protein kinase C-binding enigma)
238419_at	PHLDB2	T68150	LL5 beta
1552622_s_at	POLR2J4	BQ613856	---

(Supplementary Table 1) Contd…..

Probe ID	Gene Name	Gene Bank	Description
201702_s_at	PPP1R10	AI492873	protein phosphatase 1, regulatory subunit 10
1560587_s_at	PRDX5	AI718223	peroxiredoxin 5
202482_x_at	RANBP1	AI862473	RAN binding protein 1
212332_at	RBL2	BF110947	retinoblastoma-like 2 (p130)
203227_s_at	SAS	NM_005981	sarcoma amplified sequence
202061_s_at	SEL1L	AI927770	sel-1 suppressor of lin-12-like (C. elegans)
244477_at	SLC12A3	AW292635	solute carrier family 12 (sodium/chloride transporters), member 3
217707_x_at	SMARCA2	AI535683	---
204104_at	SNAP45	NM_003083	small nuclear RNA activating complex, polypeptide 2, 45kDa
203609_s_at	SSDH	NM_001080	aldehyde dehydrogenase 5 family, member A1
222439_s_at	THRAP3	BE967048	thyroid hormone receptor-associated protein, 150 kDa subunit
232421_at	UBC	AV703311	scavenger receptor class B, member 1
243648_at	ZC3H11A	AA280627	Homo sapiens transcribed sequences
226689_at	ZCD2	AI749451	Homo sapiens mRNA; cDNA DKFZp686A1586 (from clone DKFZp686A1586)

DISCUSSION

A very large amount of research studies have been conducted on deceased tissues and also information included in all the databases reporting normal gene expression throughout the tissues (i.e. http://www.symatlas.gnf.org) is derived from autoptic studies.

We have compared livers from living and sudden dead subjects in order to validate the set of control livers that we used in this study. Even though 217 genes are differentially expressed between CLLD and CL samples, only 54 out of these genes (63 probe sets) are included in the set of 900 genes we found differentially expressed in DL *vs*. CL.

This means that these 54 genes might be affected by the choice of deceased tissues, as they are already dysregulated between CL and CLLD. Subtraction of these genes from the set of 900 genes differentially expressed does not affect at all either the global description or the considerations about injuries occurring in donor livers.

We conclude that the set of data used as control in our study is suitable to investigate dysregulated gene expression patterns already affecting donor grafts. We must also consider that biopsies from liver resections during surgery for neoplastic lesions or even benign pathologies are not completely immune by criticisms as they might be affected by therapy, organ inflammation, anaesthesia and so on.