

Multiplex enhancer-reporter assays uncover unsophisticated TP53 enhancer logic

Annelien Verfaillie¹, Dmitry Svetlichnyy¹, Hana Imrichova, Kristofer Davie¹, Mark Fiers²,
Zeynep Kalender Atak¹, Gert Hulselmans¹, Hana Imrichova¹, Valerie Christiaens¹, and Stein
Aerts^{1#}

¹Laboratory of Computational Biology, Center for Human Genetics, University of Leuven, Leuven,
Belgium

²VIB Center for the Biology of Disease, Leuven, Belgium

correspondence to: stein.aerts@med.kuleuven.be

Supplemental Tables

Supplemental Table S1 – Final set of CHEQ-seq region. Table containing all 18399 CHEQ-seq regions with the following additional information:

Region information: the chromosomal position, corresponding barcode sequence, barcode expression (\log_2FC) and significance ($padj$). Overlap with designed regions: name of designed region, chromosomal position of designed region, bp overlap, and percentage overlap. Overlap with all called TP53 ChIP peaks (in-house data): peak name, chromosomal position of peak, peak score, peak summit location, and percentage overlap

Supplemental Table S2 – List of ChIP-peaks associated with the CHEQ-seq subsets.

Table containing ChIP peaks assigned to their CHEQ-seq subsets. The following information is provided:

Peak information: name, chromosomal position, expression and significance and percentage overlapped by region. Region information: chromosomal position, region name and barcodes

Subset name

Supplemental Table S3 –Final set of STARR-seq region. Table containing all 59584 STARR-seq regions with the following additional information

Region information: the name, chromosomal position, differential expression (\log_2FC) and significance ($padj$). Overlap with designed regions: name of designed region, chromosomal position of designed region and percentage overlap. Overlap with all called TP53 ChIP peaks (in-house data): peak name, chromosomal position of peak and percentage overlap

Supplemental Table S4 – List of ChIP-peaks associated with the STARR-seq subsets.

Table containing ChIP peaks assigned to their STARR-seq subsets. The following information is provided:

Peak name, chromosomal position, expression and significance, percentage overlapped by region, peakscore, region name and subset name

Supplemental Table S5 – Genes associated to CHEQ-seq subsets. Indirect and direct peaks were assigned the closest gene within 20kb using GREAT. The peak name, gene name and distance to the gene are provided.

Supplemental Table S6 - Gene Ontology results of directly bound peaks. 671 directly bound peaks and 2963 indirect peaks from TP53 ChIP-seq in MCF7 were assigned their closest gene within 20kb and analyzed using GREAT for their Gene Ontology and pathway enrichment. No TP53 related terms could be found associated to the indirect peaks.

Term Name	Binom Raw P-Value	Binom FDR Q-Val	Binom Fold Enrichment	Binom Observed Region Hits	Binom Region Set Coverage	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment	Hyper Observed Gene Hits	Hyper Total Genes	Hyper Gene Set Coverage
GO Biological Process											
cellular response to stress	1.16E-14	1.22E-10	4.7023	38	5.66%	1	5.77E-06	3.225	35	1145	20.47%
cellular response to DNA damage stimulus	2.77E-13	1.44E-09	6.3569	26	3.87%	2	9.39E-05	3.965	23	612	13.45%
negative regulation of cell cycle process	9.69E-10	3.37E-06	9.1403	14	2.09%	3	3.68E-03	5.8613	12	216	7.02%
Panther Pathway											
p53 pathway	5.51E-11	8.38E-09	18.1196	11	1.64%	1	7.91E-06	12.1734	9	78	5.26%
p53 pathway feedback loops 2	9.38E-08	7.13E-06	19.9228	7	1.04%	2	4.93E-03	11.7226	5	45	2.92%
MSigDB Pathway											
Direct p53 effectors	3.53E-22	4.66E-19	21.6385	22	3.28%	1	1.08E-12	13.8617	18	137	10.53%
p53 signaling pathway	1.71E-17	1.13E-14	29.378	15	2.24%	2	1.08E-09	18.6182	12	68	7.02%
Validated transcriptional targets of TAp63 isoforms	8.88E-12	3.91E-09	17.8299	12	1.79%	3	1.50E-05	15.925	8	53	4.68%
p73 transcription factor network	4.07E-11	1.34E-08	18.6531	11	1.64%	4	3.28E-03	9.3484	7	79	4.09%
p53 Signaling Pathway	1.16E-09	3.07E-07	60.714	6	0.89%	5	3.43E-03	26.3757	4	16	2.34%
Cell Cycle: G2/M Checkpoint	3.73E-08	8.20E-06	33.6908	6	0.89%	6	1.57E-02	17.5838	4	24	2.34%
MSigDB Perturbations											
Genes up-regulated in B lymphocytes at 6 h after exposure to 10 Gy dose of ionizing radiation.	1.46E-28	4.90E-25	23.0139	28	4.17%	1	1.68E-18	15.6301	24	162	14.04%
Genes up-regulated in B lymphocytes at 2 h after exposure to 10 Gy dose of ionizing radiation.	1.23E-20	2.07E-17	39.9108	16	2.38%	2	5.32E-11	25.3207	12	50	7.02%
Genes displaying an ionizing radiation response in the human skin cell samples.	5.22E-20	5.86E-17	30.946	17	2.53%	4	1.39E-10	18.7882	13	73	7.60%
Genes significantly (FDR < 10%) up-regulated in IMR-90 cells (fibroblast) in response to direct irradiation.	7.62E-20	6.41E-17	23.0397	19	2.83%	3	7.17E-11	15.0718	15	105	8.77%
Genes up-regulated in NT2/D1 cells (embryonal carcinoma) in response to treatment with cisplatin	2.78E-17	1.87E-14	41.7597	13	1.94%	5	3.42E-09	25.1197	10	42	5.85%
Genes up-regulated in the human skin cells at 4 h after exposure to 5 Gy dose of ionizing radiation.	7.69E-17	4.31E-14	38.5393	13	1.94%	7	1.59E-07	21.1006	9	45	5.26%
Genes whose promoters contain TP53 [GeneID=7157] response elements.	3.41E-16	1.64E-13	76.3422	10	1.49%	6	6.12E-08	43.4424	7	17	4.09%
Primary up-regulated targets of TP53 [GeneID=7157] in the H1299 (lung cancer) cell line.	2.69E-14	1.13E-11	29.4981	12	1.79%	9	2.65E-05	14.5521	8	58	4.68%
Genes discriminating TP53 [GeneID=7157] status across various genotoxic stress agents.	9.53E-11	3.21E-08	54.8159	7	1.04%	10	9.74E-05	32.9697	5	16	2.92%

Supplemental Table S7 – Overview of the collected publicly available TP53 ChIP seq data.

	cell type			experimental conditions			publication information					
	name	type	tissue of origin	p53 stimulation	effect	time	GEO accession	author and PMID	mapped reads / million	total peaks	direct	indirect
IR_p53	HCT116	colorectal carcinoma	colon	ionizing radiation	DNA damage	8h	GSE60267	Desantis et al (PMID 25996291)	11.9	830	160	670
SFU_IMR90	IMR90	normal (fibroblast)	lung	5-fluorouracil	DNA damage	6h	GSE31558	Botcheva et al. (PMID 22127205)	5.2	1046	551	495
IR_LCL	LCL	lymphoma	B lymphocyte (blood)	ionizing radiation	DNA damage	24h	GSE46991	Zeron-Medina et al. (PMID 24120139)	20	1078	707	371
SFU_HCT116	HCT116	cancer cell line	colon	5-fluorouracil	DNA damage	12h	GSE58506	Sanchez et al. (PMID 25524025)	50	2797	660	2137
ActD_U2OS	U2-OS	osteosarcoma	bone	actinomycin D	inhibition of transcription	24h	GSE22186	Smeenk et al. (PMID 21394211)	5.2	3319	1734	1585
SFU_MCF7	MCF7	adenocarcinoma	breast	5-fluorouracil	DNA damage	8h	SRP007261	Nikulenkov et al. (PMID 22790872)	6.7	4389	2838	1551
nutlin_IMR90	IMR90	normal (fibroblast)	lung	nutlin-3a	p53 stabilization	6h	GSE58740	Sammons et al. (PMID 25391375)	79	4399	2318	2081
Eto_U2OS	U2-OS	osteosarcoma	bone	etoposide	inhibition of transcription	24h	GSE22186	Smeenk et al. (PMID 21394211)	5.5	5184	3156	2028
Adria_Skin	HFk	foreskin keratinocytes	skin	adriamycin	genotoxic	24h	GSE56640	McDade et al. (PMID 24823795)	13.8	5976	2402	3574
doxo_colorec	SW480	colorectal adenocarcinoma	colon	doxorubicin	DNA damage	16h	GSE67181	Hüntten et al. (PMID 26183718)	19.4	6346	4697	1649
RITA_MCF7	MCF7	adenocarcinoma	breast	RITA	p53 stabilization	8h	SRP007261	Nikulenkov et al. (PMID 22790872)	7.3	6363	3917	2446
Cipl_skin	HFk	foreskin keratinocytes	skin	cisplatin	genotoxic	24h	GSE56640	McDade et al. (PMID 24823795)	11.1	10883	5806	5077
nutlin_MCF7	MCF7	adenocarcinoma	breast	nutlin-3a	p53 stabilization	8h	SRP007261	Nikulenkov et al. (PMID 22790872)	6.05	13537	10896	2641
nutlin_LCL	LCL	lymphoma	B lymphocyte (blood)	nutlin-3a	p53 stabilization	24h	GSE46991	Zeron-Medina et al. (PMID 24120139)	25	25612	24642	970
doxo_LCL	LCL	lymphoma	B lymphocyte (blood)	doxorubicin	DNA damage	24h	GSE46991	Zeron-Medina et al. (PMID 24120139)	14.5	49652	45091	4561

Supplemental Table S8 – Genes associated to strongly bound predicted TP53 sites. Sites were assigned the closest gene within 20kb using GREAT. The binding site, gene name and distance to the gene are provided.

Supplemental Table S9 – Gene ontology results of strongly bound predicted binding sites. 1148 strongly bound predicted TP53 binding sites were assigned their closest gene within 20kb and analyzed using GREAT for their gene ontology.

Term Name	Binom Raw P-Value	Binom FDR Q-Val	Binom Fold Enrichment	Binom Observed Region Hits	Binom Region Set Coverage	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment	Hyper Observed Gene Hits	Hyper Total Genes	Hyper Gene Set Coverage
GO Biological Process											
cellular response to stress	4.79E-12	5.00E-08	3.3271	46	0.0401	1	0.0083093	2.2357	42	1,145	0.1419
cellular response to DNA damage stim	1.73E-11	9.04E-08	4.4301	31	0.027	3	0.01147	2.6889	27	612	0.0912
response to stress	1.13E-10	1.97E-07	2.1782	82	0.0714	5	0.032777	1.5844	77	2,962	0.2601
negative regulation of cell cycle proces	4.05E-10	5.29E-07	6.8689	18	0.0157	2	0.015486	4.2326	15	216	0.0507
negative regulation of biological proces	4.74E-10	5.50E-07	2.0373	89	0.0775	4	0.040634	1.5421	84	3,320	0.2838
signal transduction by p53 class media	3.39E-09	3.21E-06	9.2668	13	0.0113	6	0.050022	5.0791	10	120	0.0338
cell death	1.92E-08	1.25E-05	2.6474	43	0.0375	7	0.043605	1.9725	40	1,236	0.1351
death	2.09E-08	1.21E-05	2.6396	43	0.0375	8	0.040293	1.9677	40	1,239	0.1351
Panther Pathway											
p53 pathway	9.48E-11	1.44E-08	12.5164	13	0.0113	1	8.6789E-05	7.814	10	78	0.0338
p53 pathway feedback loops 2	1.60E-08	1.22E-06	14.9718	9	0.0078	2	0.0067016	8.1266	6	45	0.0203
MsigDB Pathway											
Direct p53 effectors	4.07E-23	5.37E-20	15.522	27	0.0235	1	7.0518E-14	10.2324	23	137	0.0777
p53 signaling pathway	1.11E-16	7.35E-14	19.4608	17	0.0148	2	4.4335E-08	11.6521	13	68	0.0439
Validated transcriptional targets of TA	3.27E-10	1.44E-07	11.2899	13	0.0113	3	7.9018E-05	10.3499	9	53	0.0304
p73 transcription factor network	8.74E-10	2.89E-07	11.8938	12	0.0105	4	0.014989	6.1721	8	79	0.027
p53 Signaling Pathway	2.77E-08	7.32E-06	35.487	6	0.0052	5	0.029198	15.2373	4	16	0.0135
MsigDB Perturbations											
Genes up-regulated in B lymphocytes	4.72E-37	1.59E-33	19.2165	40	0.0348	1	9.331E-26	13.1681	35	162	0.1182
Genes displaying an ionizing radiation	7.69E-23	1.29E-19	23.4077	22	0.0192	2	1.818E-13	15.0286	18	73	0.0608
Genes significantly (FDR < 10%) up-reg	1.30E-21	1.46E-18	17.0104	24	0.0209	3	5.3096E-13	11.6094	20	105	0.0676
Genes up-regulated in B lymphocytes	8.14E-20	6.84E-17	26.2436	18	0.0157	4	3.5E-11	17.0658	14	50	0.0473
Genes up-regulated in the human skin	1.28E-19	8.58E-17	29.4571	17	0.0148	6	1.2282E-10	17.6076	13	45	0.0439
Genes whose up-regulated in NT2/D1 c	1.10E-18	6.17E-16	30.041	16	0.0139	5	5.3836E-11	18.8653	13	42	0.0439
Genes whose promoters contain TP53	6.92E-14	3.33E-11	44.6216	10	0.0087	7	2.4212E-06	25.0968	7	17	0.0236
Primary up-regulated targets of TP53	7.13E-13	3.00E-10	18.6783	13	0.0113	15	9.0075E-05	9.4577	9	58	0.0304
Genes up-regulated in SH-SY5Y cells (n	1.71E-11	6.38E-09	6.8079	21	0.0183	13	7.3762E-05	4.4854	17	231	0.0574
Genes up-regulated by TP53 [GeneID=	2.45E-11	8.24E-09	44.026	8	0.007	9	4.8343E-05	22.856	6	16	0.0203