Appendix 1

- **WES library preparation**

  Library preparation, exome capture, sequencing and data analysis have been done by IntegraGen SA (Evry, France). Genomic DNA is captured using Twist Mouse Exome Panel, Twist Bioscience. For detailed explanations of the process, see Gnirke publication in Nature Methods [1]. Sequence capture, enrichment and elution are performed according to manufacturer's instruction and protocols (Twist Bioscience) without modification except for library preparation performed with NEBNext® Ultra II kit (New England Biolabs®). For library preparation 150 ng of each genomic DNA are fragmented by sonication and purified to yield fragments of 150-200 bp. Paired- end adaptor oligonucleotides from the NEB kit are ligated on repaired, a-tailed fragments then purified and enriched by 7 PCR cycles. 500ng of these purified Libraries are then hybridized to the Twist oligo probe capture library for 16 hr in a singleplex reaction. After hybridization, washing, and elution, the eluted fraction is PCR- amplified with 8 cycles, purified and quantified by QPCR to obtain sufficient DNA template for downstream applications. Each eluted-enriched DNA sample is then sequenced on an Illumina NovaSeq as Paired End 100 reads. Image analysis and base calling is performed using Illumina Real Time Analysis with default parameters. [1] : Gnirke, A. et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol 27, 182-189 (2009).

- **RNA-seq library preparation**

  RNA sequencing: RNA-Seq libraries are performed with NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina according to supplier recommendations (NEB). The capture is then performed on cDNA libraries with the Twist Mouse Exome Enrichment System according to supplier recommendations (Twist Bioscience). First of all, an RNA quality control is performed on Fragment Analyzer (Agilent) with the RNA kit (DNF-489) to check the integrity Subdirección General de Evaluación y Fomento de la Investigación Health Research Project application form, page 14 2021 UNIÓN EUROPEA of the RNA profile and to assess the RNA concentration. The protocol permits to convert total RNA into a library of template molecules of known strand origin. Then a capture of the coding regions of the transcriptome is performed and the resulting library is suitable for subsequent cluster generation and sequencing. Briefly, the RNA is fragmented into small pieces using divalent cations under elevated temperature. cDNA is generated from the cleaved RNA fragments using random priming during first and second strand synthesis and sequencing adapters are ligated to the resulting double-stranded cDNA fragments and enriched by 12 PCR cycles. The coding regions of the transcriptome are then captured from this library using sequence- specific probes to create the final library. For that purpose, 190ng of each purified Libraries are hybridized to the Twist oligo probe capture library for 16 hours in a multiplex reaction. After hybridization, washing, and elution, the eluted fraction is PCR-amplified with 8 cycles, purified and quantified by QPCR to obtain sufficient DNA template for downstream applications. Each eluted enriched DNA sample is then sequenced on an Illumina NovaSeq as paired-end 100b reads. Image analysis and base calling are performed using the last version of Illumina Real Time Analysis with default parameters (V3.4.4). RNA sequencing will be performed at IntegraGen SA (Evry, France).

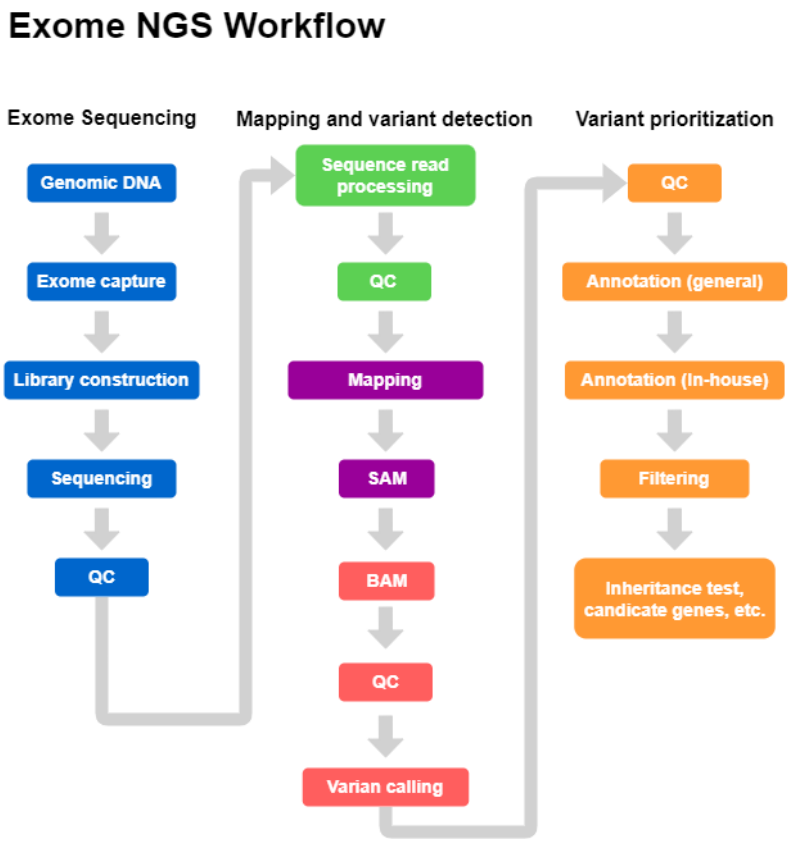

Fig 1. Bioinformatics pipeline for WES processing. Exome NGS workflow including exome sequencing, mapping and variant detection, and variant prioritization.

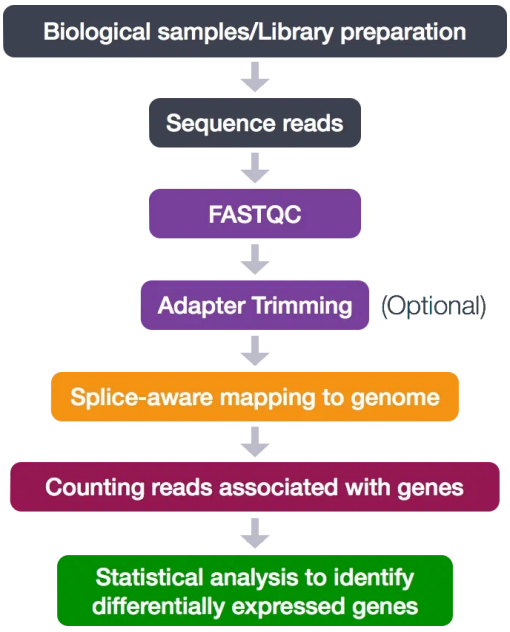

Fig 2. Bioinformatics pipeline for RNAseq processing. RNA sequencing NGS workflow for processing sequences, including a quality control.

	Sample Name	% Aligned	M Aligned	% rRNA
	.	74.7%	66.7	-
1	**RnD1-104_R1_val_1**	72.4%	129.3	29.5%
2	**RnD1-132_R1_val_1**	81.8%	172.3	15.7%
3	**RnD1-134_R1_val_1**	72.3%	91.9	29.7%
4	**RnD1-162_R1_val_1**	81.4%	161.2	15.0%
5	**RnD1-195_R1_val_1**	74.9%	100.6	24.6%
6	**RnD1-207_R1_val_1**	75.2%	122.7	23.6%
7	**RnD1-214_R1_val_1**	72.9%	93.7	22.1%
8	**RnD1-228_R1_val_1**	73.3%	103.7	23.9%
9	**RnD1-36_R1_val_1**	76.7%	135.6	22.6%
10	**RnD1-81_R1_val_1**	71.5%	101.8	25.8%
11	**RnD12-135_R1_val_1**	81.7%	146.7	14.7%
12	**RnD12-145_R1_val_1**	84.4%	138.6	11.0%
13	**RnD12-169_R1_val_1**	83.6%	155.1	13.1%
14	**RnD12-178_R1_val_1**	79.2%	122.9	17.2%
15	**RnD12-184_R1_val_1**	82.3%	163.2	14.1%
16	**RnD12-185-1_R1_val_1**	77.2%	98.6	21.7%
17	**RnD12-185-2_R1_val_1**	71.1%	76.9	25.9%
18	**RnD12-205_R1_val_1**	82.2%	319.6	-
19	**RnD12-212_R1_val_1**	81.4%	146.3	14.5%
20	**RnD12-229_R1_val_1**	79.1%	100.4	20.8%
21	**RnD12-231_R1_val_1**	73.7%	84.2	25.6%
22	**RnD12-24_R1_val_1**	79.8%	140.0	16.7%
23	**RnD12-49_R1_val_1**	67.5%	130.5	29.4%
24	**RnD12-51_R1_val_1**	45.8%	49.0	68.3%
25	**RnD3-11_R1_val_1**	82.5%	130.1	15.0%
26	**RnD3-168_R1_val_1**	82.3%	190.3	-
27	**RnD3-203_R1_val_1**	62.4%	75.7	41.0%
28	**RnD3-212_R1_val_1**	78.5%	170.8	18.8%
29	**RnD3-232_R1_val_1**	79.7%	109.7	18.5%
30	**RnD3-233_R1_val_1**	80.2%	143.7	17.5%
31	**RnD3-66_R1_val_1**	60.2%	106.9	35.7%
32	**RnD3-92_R1_val_1**	81.3%	171.1	15.9%
33	**RnD6-132_R1_val_1**	80.5%	142.1	16.9%
34	**RnD6-152_R1_val_1**	71.1%	83.9	32.3%
35	**RnD6-179_R1_val_1**	75.5%	128.5	27.9%

36	**RnD6-204_R1_val_1**	83.5%	105.5	13.4%
37	**RnPr109_R1_val_1**	80.2%	124.9	16.0%
38	**RnPr140_R1_val_1**	80.3%	211.6	15.9%
39	**RnPr146_R1_val_1**	74.4%	189.0	22.7%
40	**RnPr214_R1_val_1**	79.9%	171.9	9.5%
	STARpass1	75.0%	67.0	-

Table 1. Quality control report. This table shows the mapping quality of the samples. %Aligned: % uniquely mapped reads. M aligned: uniquely mapped reads per million. % RNA: percentage of reads matched to a SortMeRNA database.

	Sample	Histology subtype	Group
1	**RnD1.36**	Squamous	D1
2	**RnD1.81**	Squamous	D1
3	**RnD1.104**	Adenocarcinoma	D1
4	**RnD1.132**	Adenocarcinoma	D1
5	**RnD1.134**	Squamous	D1
6	**RnD1.162**	Adenocarcinoma	D1
7	**RnD1.195**	Adenocarcinoma	D1
8	**RnD1.207**	Adenocarcinoma	D1
9	**RnD1.214**	Adenocarcinoma	D1
10	**RnD1.228**	Other	D1
11	**RnD3.11**	Other	D3
12	**RnD3.66**	Other	D3
13	**RnD3.92**	Adenocarcinoma	D3
14	**RnD3.168**	Adenocarcinoma	D3
15	**RnD3.203**	Adenocarcinoma	D3
16	**RnD3.212**	Adenocarcinoma	D3
17	**RnD3.232**	Adenocarcinoma	D3
18	**RnD3.233**	Adenocarcinoma	D3
19	**RnD6.16**	Other	D6
20	**RnD6.132**	Adenocarcinoma	D6

21	**RnD6.152**	Adenocarcinoma	D6
22	**RnD6.179**	Adenocarcinoma	D6
23	**RnD6.204**	Adenocarcinoma	D6
24	**RnD12.24**	Other	D12
25	**RnD12.49**	Squamous	D12
26	**RnD12.135**	Other	D12
27	**RnD12.145**	Adenocarcinoma	D12
28	**RnD12.169**	Other	D12
29	**RnD12.178**	Adenocarcinoma	D12
30	**RnD12.184**	Adenocarcinoma	D12
31	**RnD12.185.1**	Squamous	D12
32	**RnD12.185.2**	Adenocarcinoma	D12
33	**RnD12.205**	Other	D12
34	**RnD12.212**	Other	D12
35	**RnD12.229**	Adenocarcinoma	D12
36	**RnD12.231**	Adenocarcinoma	D12
37	**RnPr109**	Other	Pr
38	**RnPr140**	Adenocarcinoma	Pr
39	**RnPr146**	Adenocarcinoma	Pr
40	**RnPr214**	Adenocarcinoma	Pr

Table 2. Histology subtype and dose exposure of the overall cohort. The table contains information about the cancer types developed by each sample and which group they pertain. There are two samples from the same rat.

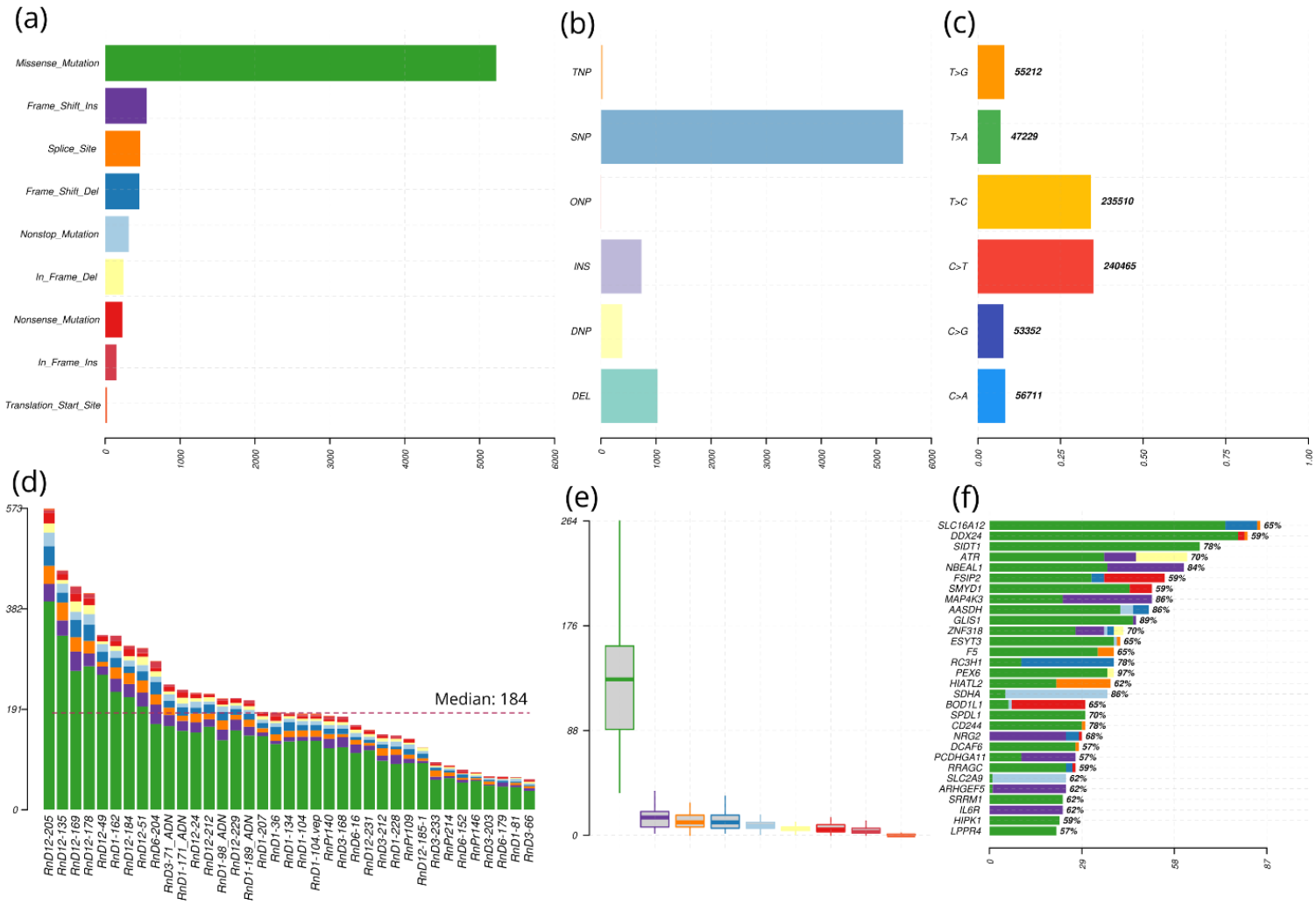

Fig 3. Types of variants. (a): Variant classifications. (b): varian types. (c): SNV classes of our cohort. (d): Variants per sample. (e): Variant classification summary. (f): Top 30 mutated genes.

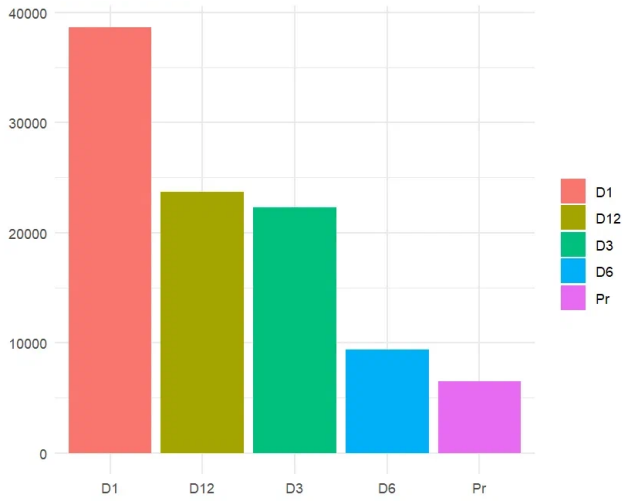

Fig 4. Total gene fusions per exposure group. Number of gene fusions sorted by exposure group.

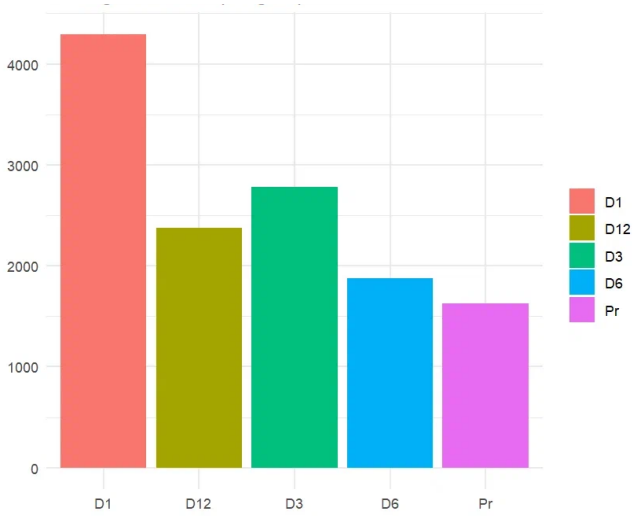

Fig 5. Mean gene fusions per exposure group. Mean distribution of gene fusions sorted by exposure group.

	.y.	group1	group2	n1	n2	p	p.signif	p.adj	p.adj.signif
*	<chr>	<chr>	<chr>	<int>	<int>	<dbl>	<chr>	<dbl>	<chr>
1	Total_gene_fusions	D1	D12	9	10	0.0125	*	0.125	ns
2	Total_gene_fusions	D1	D3	9	8	0.058	ns	0.58	ns
3	Total_gene_fusions	D12	D3	10	8	0.586	ns	1	ns
4	Total_gene_fusions	D1	D6	9	5	0.01	*	0.1	ns
5	Total_gene_fusions	D12	D6	10	5	0.572	ns	1	ns
6	Total_gene_fusions	D3	D6	8	5	0.322	ns	1	ns
7	Total_gene_fusions	D1	Pr	9	4	0.00848	**	0.0848	ns
8	Total_gene_fusions	D12	Pr	10	4	0.432	ns	1	ns
9	Total_gene_fusions	D3	Pr	8	4	0.241	ns	1	ns
10	Total_gene_fusions	D6	Pr	5	4	0.814	ns	1	ns

Table 3. Total gene fusions t-test. T-test comparing total gene fusions among exposure groups. Significant differences among groups are marked with *.

	.y.	group1	group2	n1	n2	p	p.signif	p.adj	p.adj.signif
*	<chr>	<chr>	<chr>	<int>	<int>	<dbl>	<chr>	<dbl>	<chr>
1	Inter.Chromosomal	D1	D12	9	10	0.247	ns	1	ns
2	Inter.Chromosomal	D1	D3	9	8	0.514	ns	1	ns
3	Inter.Chromosomal	D12	D3	10	8	0.0786	ns	0.786	ns
4	Inter.Chromosomal	D1	D6	9	5	0.977	ns	1	ns
5	Inter.Chromosomal	D12	D6	10	5	0.316	ns	1	ns
6	Inter.Chromosomal	D3	D6	8	5	0.597	ns	1	ns
7	Inter.Chromosomal	D1	Pr	9	4	0.0212	*	0.212	ns
8	Inter.Chromosomal	D12	Pr	10	4	0.131	ns	1	ns
9	Inter.Chromosomal	D3	Pr	8	4	0.00669	**	0.0669	ns
10	Inter.Chromosomal	D6	Pr	5	4	0.0355	*	0.355	ns

Table 4. Inter chromosomal gene fusions t-test. T-test comparing inter chromosomal gene fusions among exposure groups. Significant differences among groups are marked with *.

	.y.	group1	group2	n1	n2	p	p.signif	p.adj	p.adj.signif
*	<chr>	<chr>	<chr>	<int>	<int>	<dbl>	<chr>	<dbl>	<chr>
1	Intra.Chromosomal	D1	D12	9	10	0.00571	**	0.0571	ns
2	Intra.Chromosomal	D1	D3	9	8	0.0654	ns	0.654	ns
3	Intra.Chromosomal	D12	D3	10	8	0.365	ns	1	ns
4	Intra.Chromosomal	D1	D6	9	5	0.00802	**	0.0802	ns
5	Intra.Chromosomal	D12	D6	10	5	0.696	ns	1	ns
6	Intra.Chromosomal	D3	D6	8	5	0.261	ns	1	ns
7	Intra.Chromosomal	D1	Pr	9	4	0.002	**	0.02	*
8	Intra.Chromosomal	D12	Pr	10	4	0.27	ns	1	ns
9	Intra.Chromosomal	D3	Pr	8	4	0.0821	ns	0.821	ns
10	Intra.Chromosomal	D6	Pr	5	4	0.509	ns	1	ns

Table 5. Intra chromosomal gene fusions t-test. T-test comparing intra chromosomal gene fusions among exposure groups. Significant differences among groups are marked with *.

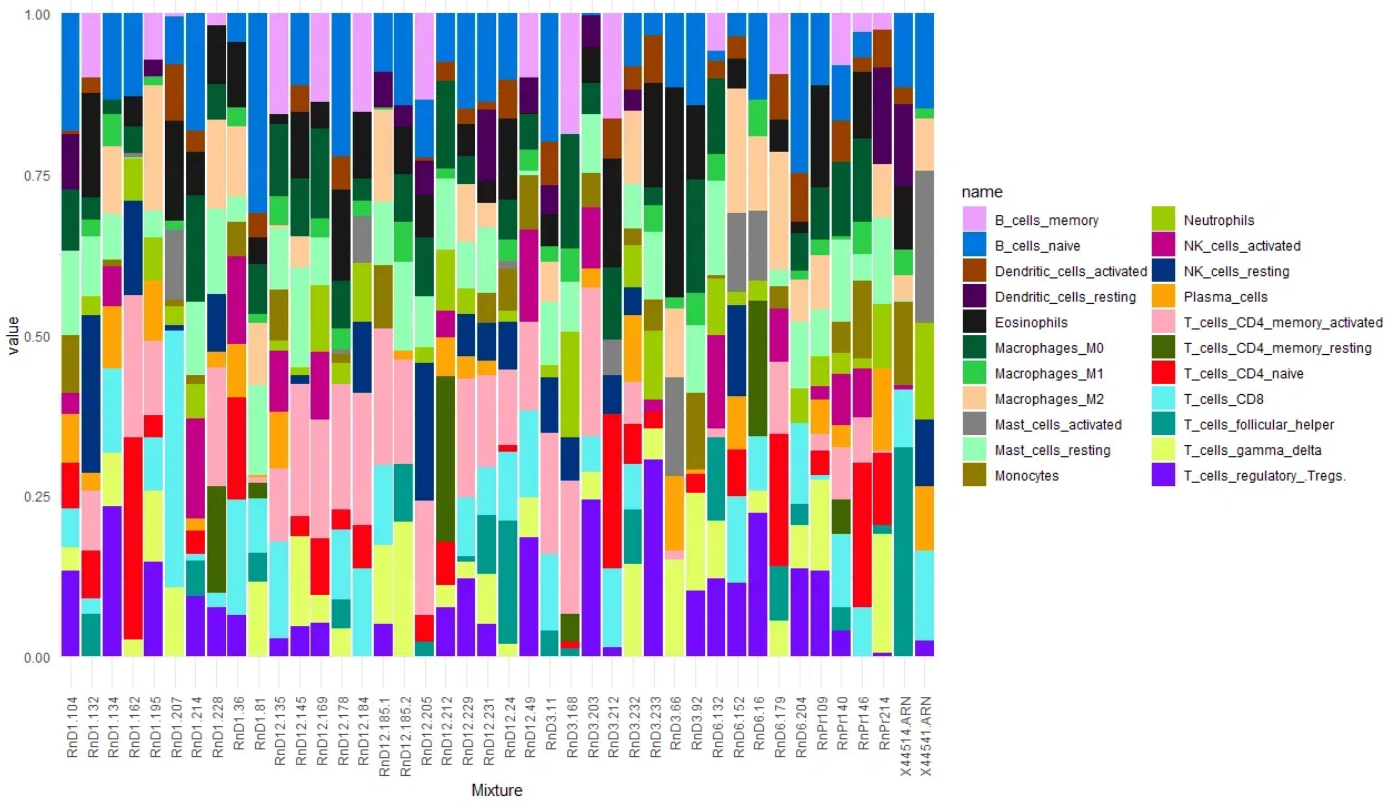

Fig 6. Immune cell proportions. Deconvolution showing the detailed proportions of immune cells. Cell proportions are estimated with cibesortx from expression data.

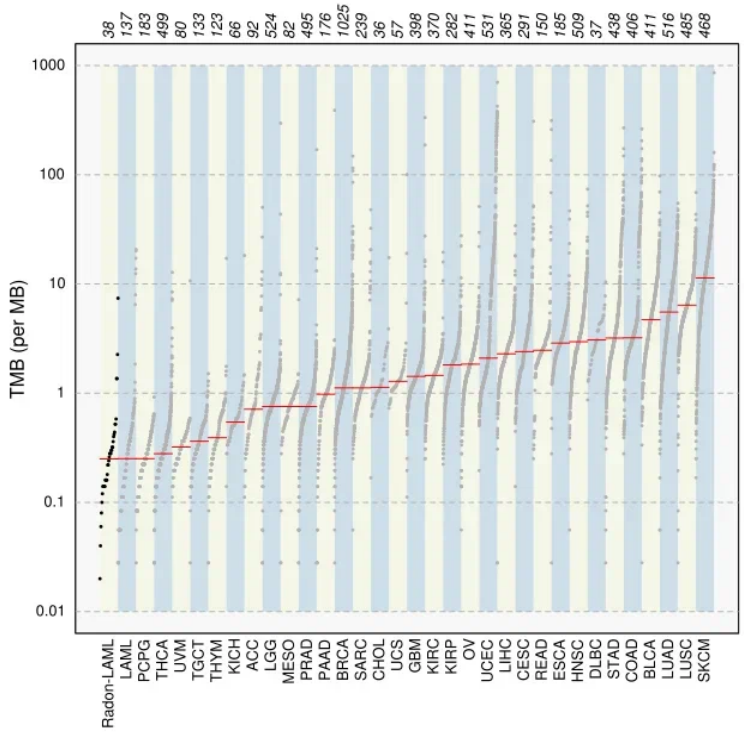

Fig 7. Radon TMB against TCGA cohorts. Tumor Mutational Burden distribution of our rat cohort (Radon-LAML) and 33 TCGA cancer cohorts. Red lines indicate the cohorts median TMB.

	.y.	group1	group2	n1	n2	p	p.signif	p.adj	p.adj.signif
*	<chr>	<chr>	<chr>	<int>	<int>	<dbl>	<chr>	<dbl>	<chr>
1	total_perMB	RnD1	RnD12	10	12	0.0567	ns	0.567	ns
2	total_perMB	RnD1	RnD3	10	6	0.639	ns	1	ns
3	total_perMB	RnD12	RnD3	12	6	0.0366	*	0.366	ns
4	total_perMB	RnD1	RnD6	10	4	0.639	ns	1	ns
5	total_perMB	RnD12	RnD6	12	4	0.0598	ns	0.598	ns
6	total_perMB	RnD3	RnD6	6	4	0.957	ns	1	ns
7	total_perMB	RnD1	RnPr	10	4	0.491	ns	1	ns
8	total_perMB	RnD12	RnPr	12	4	0.0368	*	0.368	ns
9	total_perMB	RnD3	RnPr	6	4	0.797	ns	1	ns
10	total_perMB	RnD6	RnPr	4	4	0.853	ns	1	ns

Table 6. Tumor mutational burden t-test. T-test comparing the TMB of all radon exposure groups in our cohort. Significant differences among groups are marked with *.