

The effect of interferon-beta and vitamin D on immune cells

Gunjan Janak Sethia
Pedro Herrero-Vidal

Multiple sclerosis

- **Multiple sclerosis (MS)** is a demyelinating disease → loss of nerves' isolation covers
- Demyelination results in a **slower synaptic transmission** and possible cell death
- **Symptoms:**
 - Spasticity, Walking (Gait) Difficulties
 - Vision Problems, Bladder Problems, Dizziness and Vertigo
 - Fatigue, Numbness, Weakness
 - Cognitive Changes, Emotional Changes, Depression
- More than 2.3 million people are affected by MS worldwide
- Unknown causes, but highly associated with **auto-immune** processes, and viral infections
- Has no cure yet, but it is commonly treated with **interferon- β** → modulate immune system
- Low serum **vitamin D** is associated with increased MS risk → immune system

Effect of IFN- β and vitamin D in monocytes

What transcriptional changes are induced by IFN- β and vitamin D in human immune cells, monocytes?

Data sets

- NCBI project: PRJNA258216
- Paired-end **RNAseq data**
- **Triplicates**
- Monocyte samples from **MS patients** before and 24 hours after the first treatment with **IFN- β**
- NCBI project: PRJNA285092
- Unpaired-end **RNAseq data**
- **Triplicates**
- Human cultured monocyte samples from 24 hours after treatment with **vitamin D** or ethanol

Analysis pipeline

Data pre-processing

Raw reads

FASTQC

Reads QC

Trimomatic

Trim reads

FASTQC

Trimmed reads QC

Alignment

Filtered reads

hiSAT2

Align

HTseq

Table of counts

DEG analysis

Table of counts

edgeR

Data summary

edgeR

DEG

DAVID

DEG interpretation

Data analysis

➤ **Download fastq** files from NCBI website:

➤ Project PRJNA258216

➤ Project PRJNA285092

```
CAGGACTCTGGATACCTTTATTTTTGCTTCAGGTGGCTTTTAATCAGATT
+
CCCCFFFFHHHHJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ
@SRR1550993.2 CHAUSSABELL -ILL:183:D16YPACXX:5:1101:8696:1977/1
GCTTATGTTTAAGCATTTAAAGTTGGCAAACATGTTATCAATGTATTAT
+
BCCFFFBFBHHHHJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJ
@SRR1550993.3 CHAUSSABELL -ILL:183:D16YPACXX:5:1101:13893:1995/1
CAGGGGGTCACGAATGCGGCCGCACATATTCACCCTCAGCGTGCCTTTCC
```

```
$ module load sratoolkit/2.8.0
$ wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR155/008/SRR1550998/SRR1550998_2.fastq.gz
$ gzip -d SRR1550998_2.fastq.gz
```

Analysis pipeline



Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

Analysis pipeline

Download

➤ **Quality control** of the fastq files using **FASTQC**:

FastQC Report

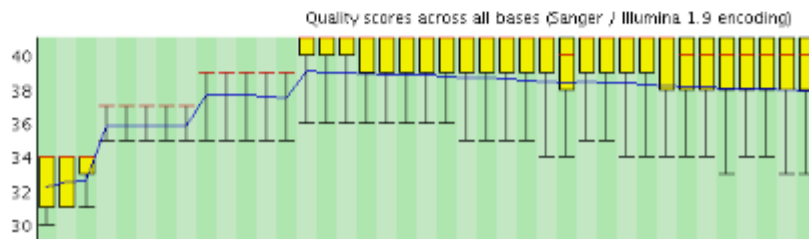
Summary

- ✓ [Basic Statistics](#)
- ✓ [Per base sequence quality](#)
- ! [Per tile sequence quality](#)
- ✓ [Per sequence quality scores](#)
- ✗ [Per base sequence content](#)
- ✓ [Per sequence GC content](#)
- ✓ [Per base N content](#)
- ✓ [Sequence Length Distribution](#)
- ! [Sequence Duplication Levels](#)
- ✓ [Overrepresented sequences](#)
- ✓ [Adapter Content](#)
- ✗ [Kmer Content](#)

✓ Basic Statistics

Measure	Value
Filename	SRR1550993_2.fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	16466842
Sequences flagged as poor quality	0
Sequence length	50
%GC	51

✓ Per base sequence quality



```
$ module load fastqc/0.11.2
$ fastqc MonoC_control_01.fastq
```

➤ Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

➤ Trim the reads based on the quality control analysis using TRIMMOMATIC:

- Trim first five bases of all reads
- Minimum length of the read 40

```
@SRR1551102.1 STE-DEDICATED:691:C1EY8ACXX:1:1101:1127:2419/1
GACGAGTAGGAGGGCCGCTGCGGTGAGCCTTGAAGCCTAGGGCGC
+
AAACA<11::<*)0=AAAA8AAAA@@@7>;5(;5(;=?=?####
@SRR1551102.2 STE-DEDICATED:691:C1EY8ACXX:1:1101:1084:2433/1
AAAGTGGATATTGTTGCCATCAATGACCCCTTCATTGACCTCAAC
+
EFFHCAFFIGB<FIGIBHIGGCEGIIBHEEFDHGHIDGHIIIIII
@SRR1551102.3 STE-DEDICATED:691:C1EY8ACXX:1:1101:1163:2469/1
GTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGCAGCTGGGATTA
```

```
$ module load trimmomatic/0.32
$ java -jar $TRIMMOMATIC_ROOT/trimmomatic.jar SE -threads 1 -phred33
MonoC_control_01.fastq MS_52_post_1.fastq HEADCROP:5 TRAILING:0 MINLEN:40
```

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

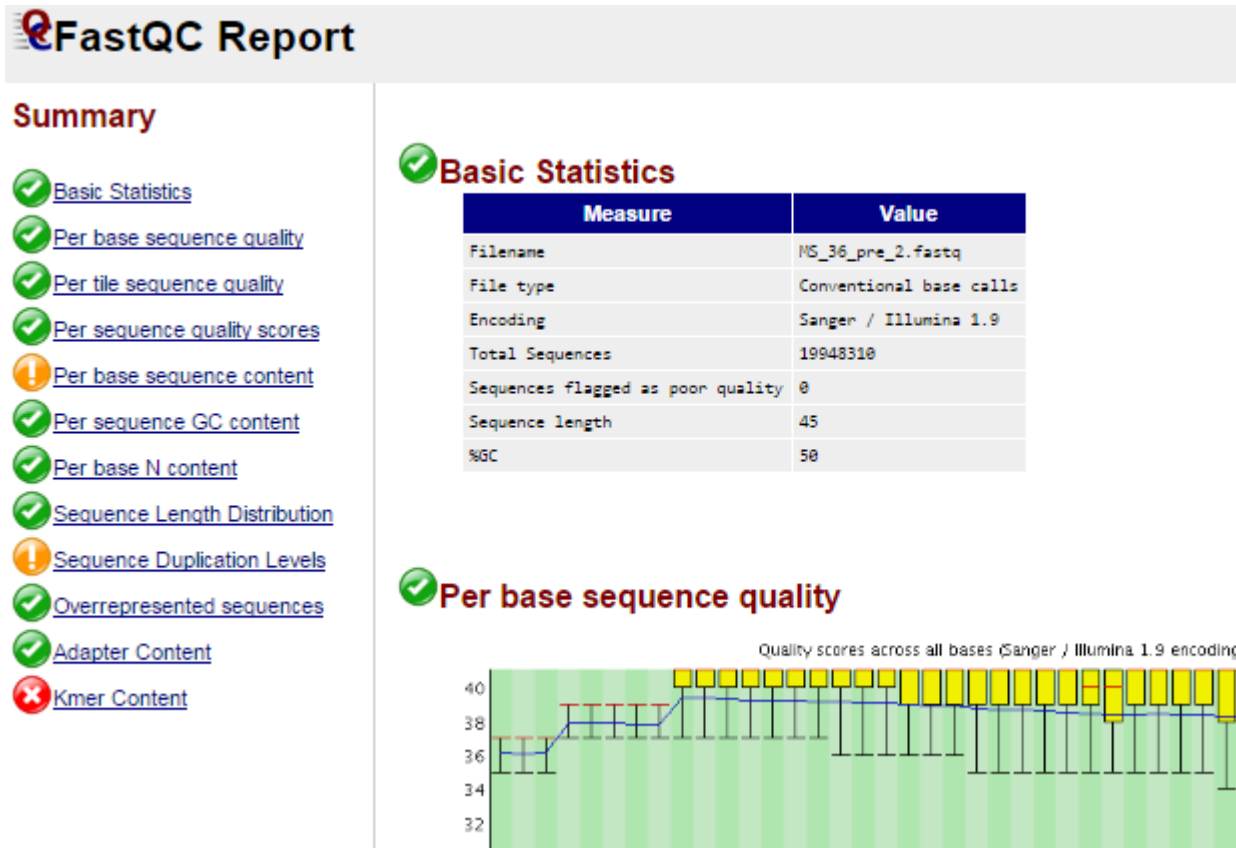
Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

- **Quality control** of the fastq files using **FASTQC**:



```
$ module load fastqc/0.11.2
$ fastqc MonoC_trimmed_control_01.fastq
```

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

➤ **Align** the filtered genes to the reference genome using **hiSAT2**:

- Reference genome: **UCSC hg38**
- **Overall alignment rate:**
 - MS patients: 92,99% ± 1,54%
 - Cultured monocytes: 92,83% ± 0,57%

```
26407049 reads; of these:
 26407049 (100.00%) were unpaired; of these:
 1981353 (7.50%) aligned 0 times
 19058623 (72.17%) aligned exactly 1 time
 5367073 (20.32%) aligned >1 times
92.50% overall alignment rate
```

```
$ module load samtools
$ module load hisat2
$ filename=MonoC_control_trimmed_0${PBS_ARRAYID}
$ hisat2 --threads 16 -x ../hisat2_ref/hg38/genome -U
${filename}.fastq | samtools view -bS -@ 16 > ${filename}.bam
$ samtools sort -@ 16 ${filename}.bam -o
${filename}.sorted.bam
$ samtools index ${filename}.sorted.bam
```

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)



Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

- Create **tables of counts** using **HTseq**:

A1BG	91	108	62	64	94	266	
A1BG-AS1		75	67	48	75	76	254
A1CF	0	0	1	0	0	0	
A2M	27	31	31	22	14	100	
A2M-AS1	21	14	10	16	8	49	
A2ML1	3	2	4	0	3	16	
A2MP1	0	0	0	0	0	0	
A3GALT2	0	0	0	0	0	1	
A4GALT	179	174	155	98	98	336	
A4GNT	7	3	5	2	7	13	

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

➡ Table of counts (HTseq)

Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

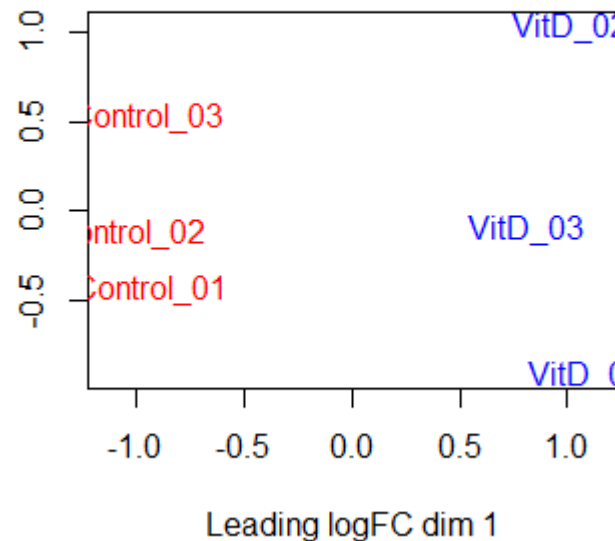
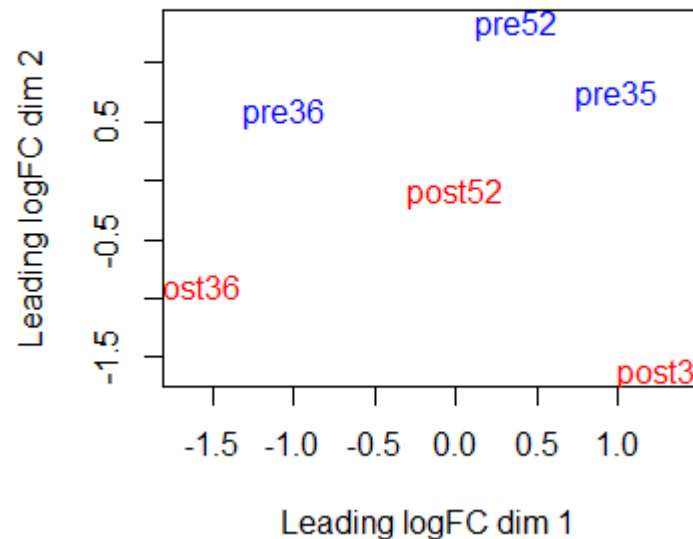
```
$ module load pysam/intel/0.8.2.1
$ module load htseq/intel/0.6.1
$ htseq-count -f bam MonoC_control_trimmed_01.sorted.bam /scratch/ggg256-
share/reference_genomes/Homo_sapiens/UCSC/hg38/Annotation/Genes/genes.gtf >
MonoC_control_01_counts.txt
```

Data analysis

- **Multidimensional scaling plot of distances (MDS)** between gene expression profiles:

IFN- β treatment

Vitamin D treatment



paired analysis

unpaired analysis

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

➡ Data summary (edgeR)

DEG (edgeR)

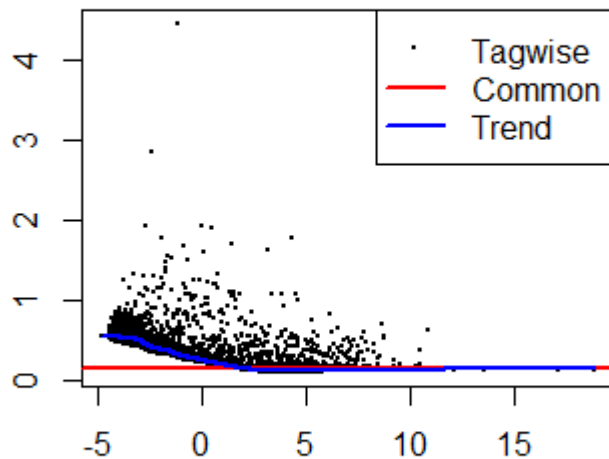
DEG interpretation (DAVID)

Data analysis

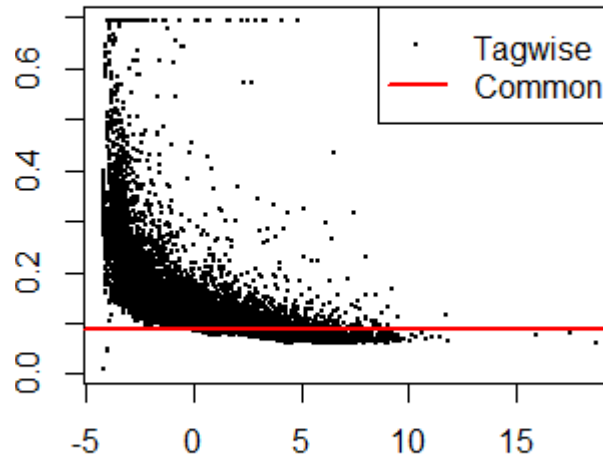
➤ Genewise biological coefficient of variation (BCV):

IFN- β treatment

Vitamin D treatment



BCV of 16%



BCV of 9%

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

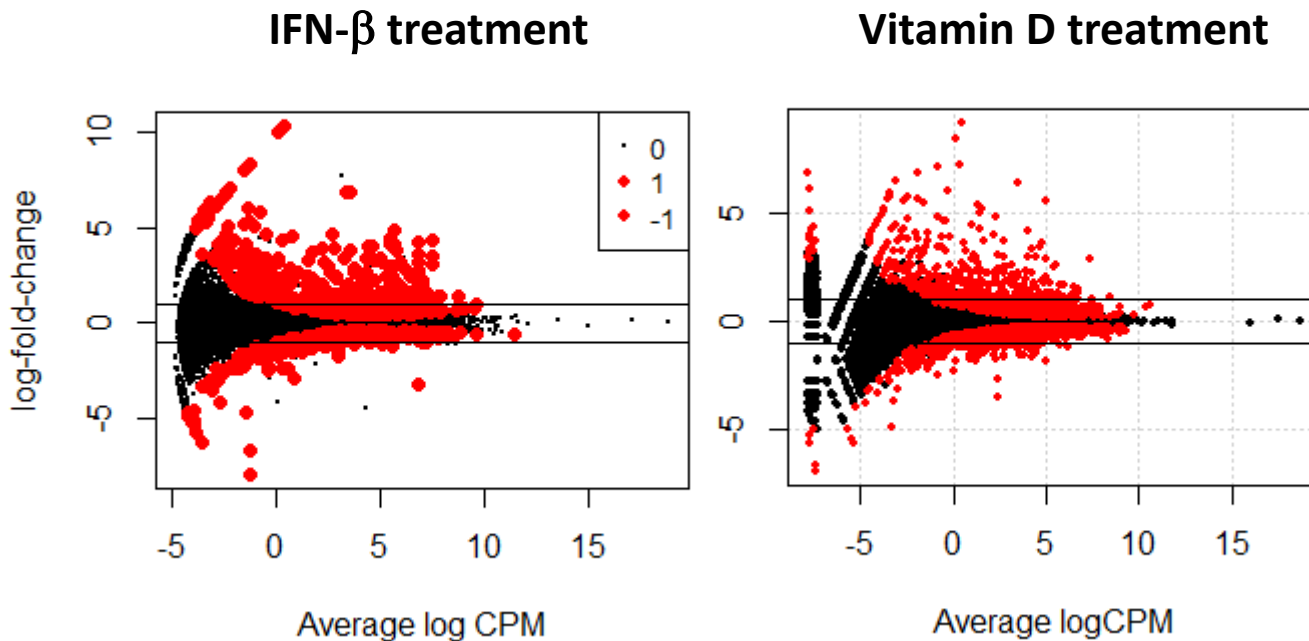
➡ Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

➤ Differentially expressed genes (DEG):



IFN-β treatment Vitamin D treatment

Upregulated DEG	1177	642
Downregulated DEG	1455	406
Total DEG	2632	1048

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

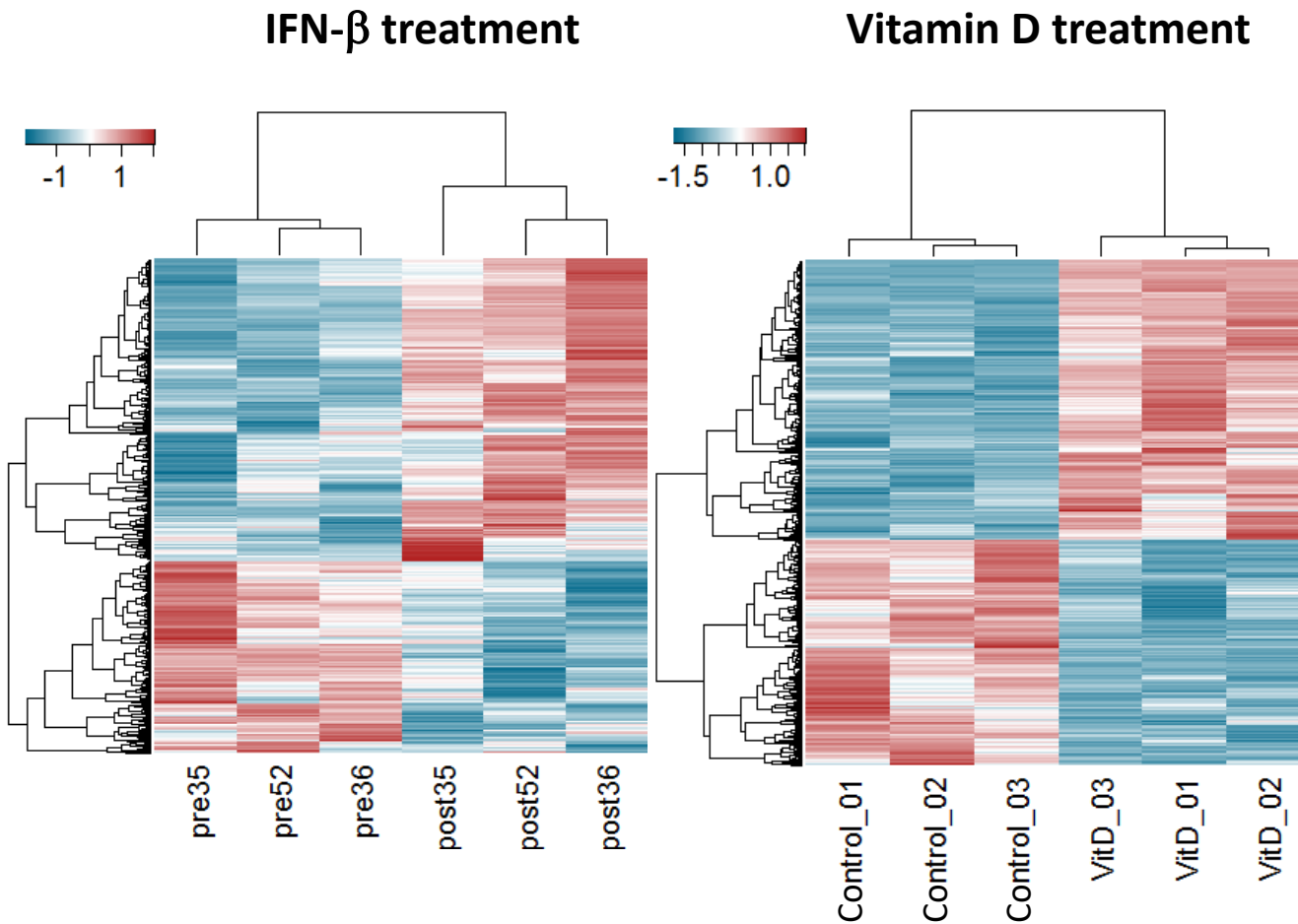
Data summary (edgeR)

➡ DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

➤ Differentially expressed genes (DEG):



Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

➡ DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

➤ Differentially expressed genes (DEG):

➤ IFN- β treatment

Comparison of groups: IFN_b-control

	logFC	logCPM	PValue	FDR
IFITM1	4.844856	5.677926	4.115588e-163	7.496543e-159
OTOF	6.899962	3.604281	4.252187e-155	3.872679e-151
IFIT2	4.164084	5.742876	2.390159e-125	1.451225e-121

Interferon response

➤ Vitamin D treatment

Comparison of groups: vitD-control

	logFC	logCPM	PValue	FDR
CD14	6.401384	5.671870	0.000000e+00	0.000000e+00
FBP1	5.575315	6.838013	0.000000e+00	0.000000e+00
SLC37A2	3.608211	5.758871	6.239596e-249	4.088183e-245

Immune and metabolic functions

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

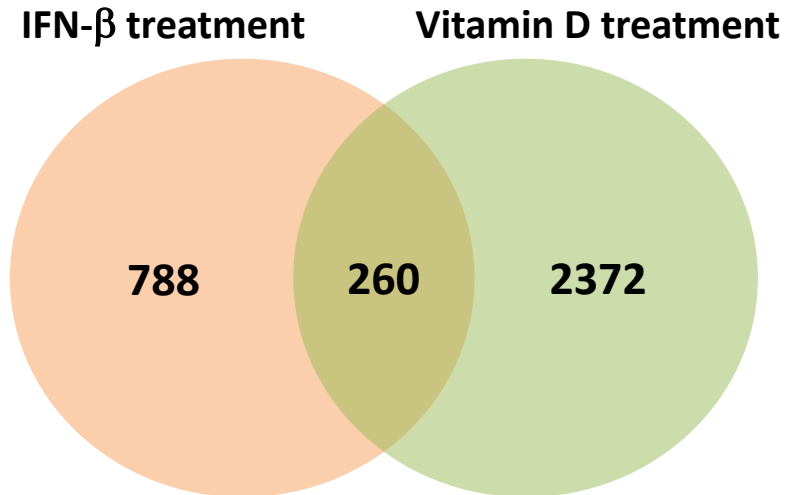
Data summary (edgeR)

 DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

➤ Differentially expressed genes (DEG):



➤ Common DEG:

		Vitamin D	
		Up-regulated	Down-regulated
IFN - β	Up-regulated	104	60
	Down-regulated	41	55

Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

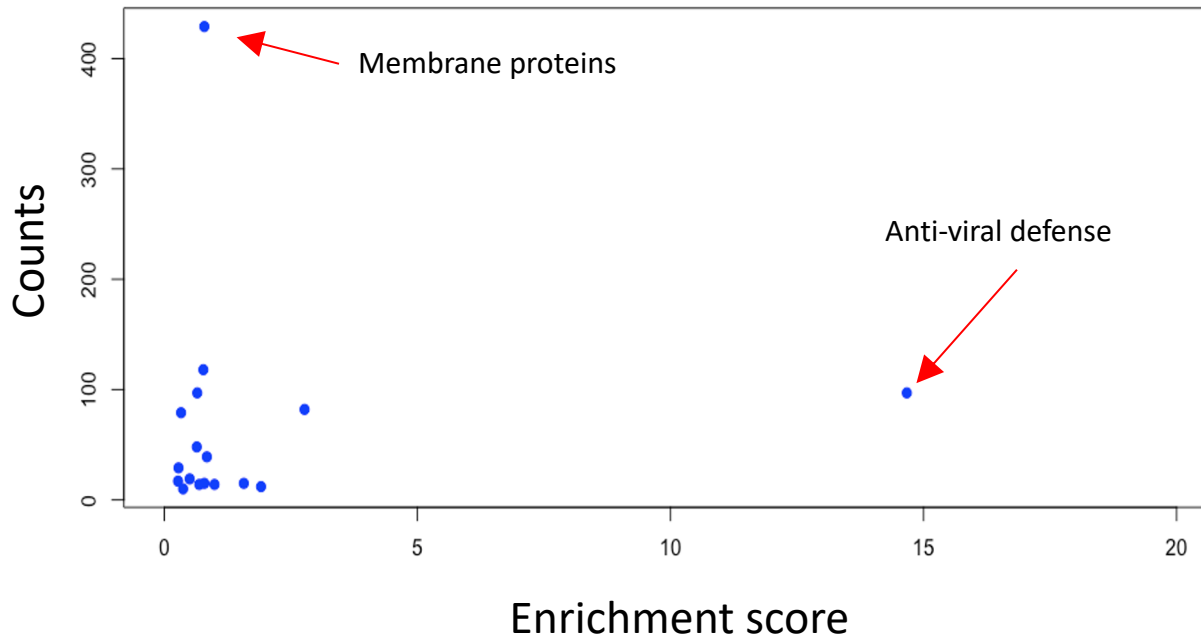
DEG (edgeR)

DEG interpretation (DAVID)



Data analysis

- **Common up-regulated DEG after IFN- β and vitamin D treatment:**



Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

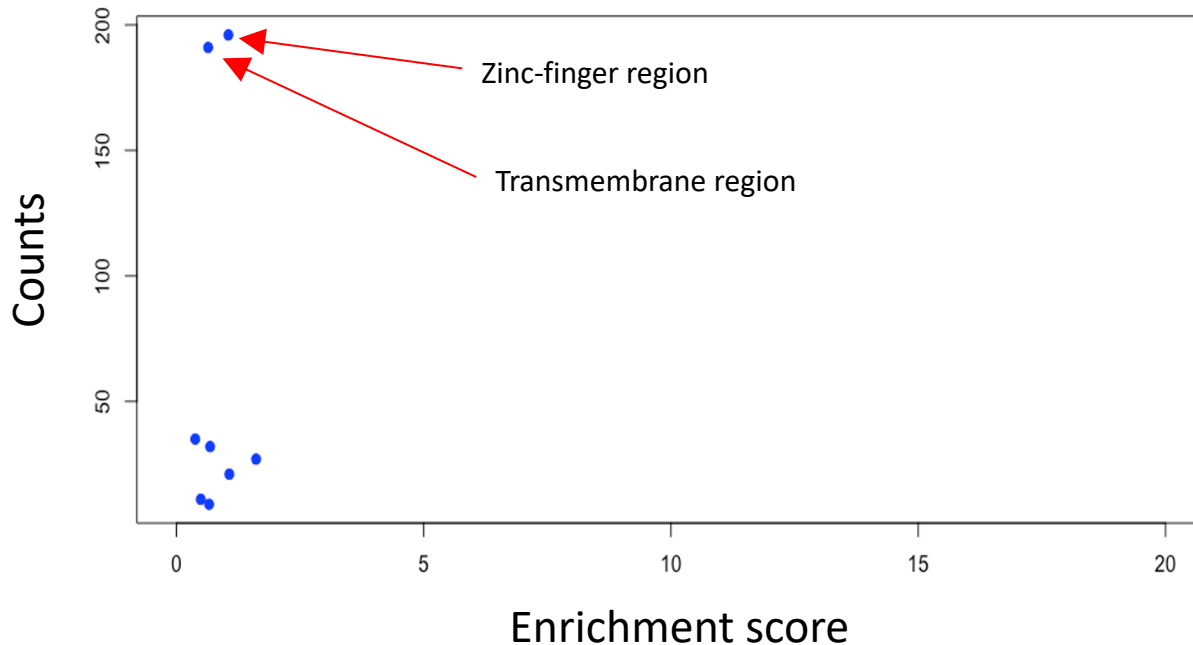
Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Data analysis

- **Common down-regulated DEG after IFN- β and vitamin D treatment:**



Analysis pipeline

Download

Quality control (FASTQC)

Trim reads (TRIMMOMATIC)

Quality control (FASTQC)

Align (hiSAT2)

Table of counts (HTseq)

Data summary (edgeR)

DEG (edgeR)

DEG interpretation (DAVID)

Conclusions

- We found **1048 DEG** in after **IFN- β** treatment in MS patients:
 - **642** of them were **up-regulated**
 - **406** of them were **down-regulated**
- Our analysis showed **2632 DEG** after **vitamin D** treatment in cultured monocytes:
 - **1177** of them were **up-regulated**
 - **1455** of them were **down-regulated**
- Interestingly, **260** genes were differentially expressed after **both treatments**:
 - **104** of them were **commonly up-regulated**
 - **55** of them were **commonly down-regulated**
- Among the **commonly up-regulated DEG**, we found functional clusters associated with **immune viral response**
- Among the **commonly down-regulated DEG**, we found functional clusters as **DNA binding**
- We successfully identified some **common functional pathways** after both treatments that may be relevant for the **treatment of MS patients**

Future work

- **A larger sample size** could strength the robustness of our analysis
- We plan on exploit our analysis pipeline to question the effect of both treatments in **other immune cell types**
- **A deeper analysis and understanding of the different DEG** is necessary to assess the relevance of these findings an possible medical interpretations
- Given the variability of the different tools to analyze DEG, it would be interesting to cross-validate our results using **alternative methods for DEG analysis**

References

- Anders, S., Pyl, P. T. & Huber, W. **HTSeq--a Python framework to work with high-throughput sequencing data**. Bioinformatics 31, 166–169 (2015).
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data**. Bioinformatics 26, 139–140 (2010).
- Kim, D., Langmead, B. & Salzberg, S. L. **HISAT: a fast spliced aligner with low memory requirements**. Nat. Methods 12, 357–360 (2015).
- Linsley, P. S., Speake, C., Whalen, E. & Chaussabel, D. **Copy number loss of the interferon gene cluster in melanomas is linked to reduced T cell infiltrate and poor patient prognosis**. PLoS One 9, e109760 (2014).
- Seuter, S., Neme, A. & Carlberg, C. **Epigenome-wide effects of vitamin D and their impact on the transcriptome of human monocytes involve CTCF**. Nucleic Acids Res. 44, 4090–4104 (2016).