

Ranjini Ghosh Dastidar

**A Dynamic Regulatory Model of *In Vivo* Vitamin D Responsiveness:
From Transcriptomic Discovery to Systems Identification
in Human Immune Cells**

Nutrigenomics Laboratory

InLife Institute of Animal Reproduction and Food Research,

Polish Academy of Sciences, Olsztyn, Poland

under the supervision of Prof. Carsten Carlberg

Doctoral thesis carried out within the WELCOME2 ERA Chair project (No. 952601)
funded by the European Union's Horizon 2020 research and innovation programme

“The art of being wise is the art of knowing what to overlook.”

— William James

Table of Contents

Abstract	10
Streszczenie (Abstract in Polish)	12
Abbreviations	14
Reader's Guide to the Studies in This Thesis	19
Chapter 1	21
General Introduction	21
<i>1.1 Vitamin D: From Sunlight to Endocrine Signal</i>	21
<i>1.2 VDR and the Genomic Programme It Controls</i>	22
<i>1.3 Physiological Functions: Beyond Calcium and Bone</i>	24
<i>1.4 Why Large-Scale Trials Have Failed: The Heterogeneity Problem</i>	25
<i>1.5 The Studies in This Thesis: Names, Designs, and the Conceptual Progression</i>	26
<i>1.6 From Targeted Assays to Genome-wide RNA-seq: A Necessary Technical and Conceptual Advance</i>	27
<i>1.7 Why PBMCs? The Rationale for Peripheral Blood Immune Cells</i>	28
<i>1.8 Bolus versus Daily Vitamin D Supplementation: A Pharmacological Rationale</i>	29
<i>1.9 In Vivo Human Studies versus In Vitro Cell Culture: Why They Don't Always Agree</i>	30
<i>1.10 The Evolution of the Vitamin D Response Index</i>	31
<i>1.11 Mechanistic Modelling: Making Sense of Biological Complexity</i>	32
Chapter 2	35
Objectives and Hypothesis	35
<i>2.1 Objectives</i>	35
<i>2.2 Research Hypotheses</i>	36
Chapter 3	37
<i>In Vivo Vitamin D Targets Reveal the Upregulation of Focal Adhesion-Related Genes in Primary Immune Cells of Healthy Individuals</i>	37
Publication details	38
Chapter 4	39

Evaluation of the Vitamin D Response Index in a Saudi Cohort	39
Publication details	40
Chapter 5.....	41
Transcriptomic Profiling of Immune Modulation Induced by Vitamin D₃ in the VitDPAS and VitDHiD Cohort Studies	41
Publication details	42
Chapter 6.....	43
Dynamic System Identification Defines a Minimal Transcription Factor Control Architecture of Human Endocrine Gene Regulation	43
Publication details	45
Chapter 7.....	46
<i>Combined Observations.....</i>	<i>46</i>
7.1 <i>The Problem of Interindividual Baseline Differences</i>	<i>46</i>
7.2 <i>Neither Vitamin D Levels Nor Transcriptomic Change Alone Explain the Response.....</i>	<i>49</i>
7.3 <i>Why a Mechanistic Model Is Needed.....</i>	<i>50</i>
Chapter 8.....	52
General Discussion.....	52
8.1 <i>From Observation to Explanation: Integrated Findings Across the Four Studies.....</i>	<i>52</i>
8.2 <i>What the Transcriptional Findings Suggest About Vitamin D's Role in the Immune System</i>	<i>53</i>
8.3 <i>HLA-C and What Baseline Differences Tell Us.....</i>	<i>54</i>
8.4 <i>Epigenetic Memory, Baseline Fluctuation, and the Implications for Trial Design.....</i>	<i>55</i>
8.5 <i>Toward Predictive Personalised Supplementation.....</i>	<i>55</i>
8.6 <i>Baseline Transcriptional Heterogeneity: Baseline Transcriptomic Architecture Before Supplementation Begins.....</i>	<i>56</i>
8.7 <i>The Evolution of the Response Index and Baseline Transcriptional Heterogeneity</i>	<i>57</i>
Chapter 9.....	60
Conclusions.....	60
Chapter 10	62
Implications and Recommendations	62

10.1 Limitations	62
10.2 Public Health: Moving Beyond One-Size-Fits-All Guidelines	63
10.3 Clinical Trial Design: What Needs to Change	63
10.4 Translational Medicine: The Path to a Predictive Assay	64
10.5 Future Research Directions	64
Chapter 11	66
List of Figures	66
General Introduction	66
Chapter 3 – VitDHiD (Scientific Reports 2024)	67
Chapter 4 – Saudi Cohort (Saudi Pharmaceutical Journal 2024)	67
Chapter 5 – VitDPAS (Scientific Reports 2025)	68
Chapter 6 – Manuscript (submitted 2026)	68
Chapter 12	69
List of Tables	69
Chapter 3 – VitDHiD	69
Chapter 4 – Saudi Cohort	69
Chapter 5 – VitDPAS	69
Chapter 6 – Manuscript, under revision	70
Chapter 13	71
List of Publicly Available Data	71
Chapter 3 – VitDHiD	71
Chapter 5 – VitDPAS	71
Chapter 6 – Manuscript	71
Chapter 14	72
References	72
Chapter 15	79
Acknowledgements	79
Chapter 16	84
Author’s Statements	84

<i>Candidate's Contribution to Each Publication</i>	<i>84</i>
Chapter 3 – Scientific Reports (2024)	84
Chapter 4 – Saudi Pharmaceutical Journal (2024).....	84
Chapter 5 – Scientific Reports (2025)	84
Chapter 6 – Under review (2026)	85
Chapter 17	86
Official Documents	86
<i>Declaration of Authorship and Originality.....</i>	<i>86</i>
Chapter 18	105
Appended Publications	105

Abstract

Vitamin D₃ is a pleiotropic fat-soluble hormone precursor whose biologically active metabolite, 1 α ,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃), regulates gene expression through the nuclear receptor vitamin D receptor (VDR). Despite decades of clinical supplementation trials and widespread public health interest, the molecular basis of its health effects has remained poorly resolved. A central and underappreciated reason for this failure is that individuals respond to vitamin D₃ supplementation with striking and reproducible interindividual variability. Some people activate hundreds of target genes robustly, while others show minimal transcriptomic change even when their serum vitamin D levels rise equivalently. This thesis investigates the origins, cross-population distribution, molecular composition, and mechanistic basis of that variability.

Four studies define the experimental arc. The VitDHiD intervention trial (NCT03537027; 25 healthy Finnish participants) used RNA-seq (RNA sequencing) to characterise the genome-wide transcriptional response to a single high-dose vitamin D₃ bolus in peripheral blood mononuclear cells (PBMCs), identifying 452 *in vivo* target genes and enabling the first transcriptomics-based classification of individuals into high, mid, and low vitamin D responders [27]. A Saudi Arabian cohort of 100 participants independently confirmed [2] that low responders exist at comparable frequencies in a genetically distinct, vitamin D-insufficient, population (though living in a sun-rich country), establishing the broader view of the response index across geographies. The VitDPAS study (NCT06104111; 45 Polish participants; three monthly bolus cycles) replicated and extended these findings, identifying a conserved 26-gene core whose dominant signature [30], is the downregulation of NF κ B (nuclear factor kappa B) pathway inhibitors, suggesting enhanced immune readiness. Finally, 287 longitudinal RNA-seq libraries from VitDPAS were used to build and validate an ordinary differential equation (ODE)-based regulatory kernel of 11 transcription factors (TFs) and 117 interactions that mechanistically explains interindividual variability, with the TP53–AP1–MYC regulatory axis (comprising the tumour suppressor tumor protein P53 (TP53), the activator protein 1 (AP1) TF complex, and the proto-oncogene MYC, a bHLH TF) emerging as the principal discriminator between high and low responders.

This thesis establishes that the vitamin D response index is a mechanistically tractable, widely distributed biological property, encoded in each individual's pre-existing transcriptional regulatory architecture and mechanistically predictable from baseline immune cell gene expression before any supplementation begins. The findings provide both a biological reframing of vitamin D's immune role and a quantitative foundation for personalised supplementation strategies.

Keywords: vitamin D · response index · transcriptomics · RNA-seq · PBMCs · focal adhesion · ODE modeling · TP53-AP1-MYC · personalised supplementation · human intervention trial · epigenetic memory

Streszczenie (Abstract in Polish)

Witamina D₃ jest plejotropowym, rozpuszczalnym w tłuszczach prekursorem hormonu, którego biologicznie aktywny metabolit, 1 α ,25-dihydroksywitamina D₃ (1,25(OH)₂D₃), reguluje ekspresję genów docelowych za pośrednictwem jądrowego receptora witaminy D (VDR). Pomimo dziesięcioleci badań klinicznych nad suplementacją oraz szerokiego zainteresowania zdrowiem publicznym, molekularne podstawy jej efektów zdrowotnych pozostają słabo poznane. Kluczowym, a często niedocenianym powodem tego stanu rzeczy jest fakt, że odpowiedź na suplementację witaminą D₃ wykazuje u ludzi wyraźną i powtarzalną zmienność. U niektórych osób dochodzi do silnej aktywacji setek genów docelowych, podczas gdy u innych obserwuje się minimalne zmiany transkryptomiczne, mimo podobnego wzrostu stężenia witaminy D w surowicy. Niniejsza praca bada źródła tej zmienności, jej rozkład w różnych populacjach, skład molekularny oraz mechanistyczne podstawy.

Cztery badania wyznaczają eksperymentalny zakres pracy. Badanie interwencyjne VitDHiD (NCT03537027; 25 zdrowych uczestników z Finlandii) wykorzystało sekwencjonowanie RNA (RNA-seq) do scharakteryzowania odpowiedzi transkrypcyjnej w całym genomie na pojedynczą wysoką dawkę witaminy D₃ w komórkach jednojądrzastych krwi obwodowej (PBMC), identyfikując 452 geny docelowe *in vivo* i umożliwiając pierwszą klasyfikację opartą na transkryptomice – podział na osoby o wysokiej, średniej i niskiej responsywności na witaminę D [27]. Kohorta 100 uczestników z Arabii Saudyjskiej niezależnie potwierdziła [2], że osoby słabo reagujące występują z porównywalną częstością w genetycznie odmiennej populacji z niedoborem witaminy D (mimo życia w kraju o dużym nasłonecznieniu), co dowodzi uniwersalności wskaźnika odpowiedzi w różnych regionach geograficznych. Badanie VitDPAS (NCT06104111; 45 uczestników z Polski; trzy miesięczne cykle bolusowe) powtórzyło i rozszerzyło te wyniki, identyfikując konserwowany zestaw 26 genów, którego dominująca cecha [30] – regulacja w dół inhibitorów szlaku NF κ B (jądrowego czynnika kappa B) – sugeruje zwiększoną gotowość immunologiczną. Ostatecznie 287 bibliotek RNA-seq z pomiarów podłużnych przeprowadzonych w ramach badania VitDPAS wykorzystano do stworzenia i walidacji modelu regulacyjnego opartego na równaniach różniczkowych zwyczajnych (ODE), obejmującego 11 czynników transkrypcyjnych (TF) i 117 interakcji, który mechanistycznie

wyjaśnia zmienność międzyosobniczą. Oś regulacyjna TP53–AP1–MYC (obejmująca supresor nowotworowy TP53, kompleks czynnika transkrypcyjnego AP1 oraz proto-onkogen MYC) okazała się głównym czynnikiem różnicującym osoby silnie i słabo reagujące.

Niniejsza praca dowodzi, że wskaźnik odpowiedzi na witaminę D jest uniwersalną, możliwą do mechanistycznego opisanie cechą biologiczną, zakodowaną w istniejącej wcześniej architekturze regulacyjnej transkrypcji każdego człowieka, i możliwą do przewidzenia na podstawie profilu ekspresji genów w komórkach układu odpornościowego jeszcze przed rozpoczęciem suplementacji. Wyniki te dostarczają zarówno nowego spojrzenia na rolę witaminy D w odporności, jak i ilościowej podstawy dla spersonalizowanych strategii suplementacji.

Słowa kluczowe: witamina D · wskaźnik odpowiedzi · transkryptomika · RNA-seq · PBMC · VitDHiD · VitDPAS · kohorta saudyjska · adhezja ogniskowa · NFκB · HLA-C · modelowanie ODE · TP53-AP1-MYC · pamięć epigenetyczna · spersonalizowana suplementacja

Abbreviations

1,25(OH)₂D₃	1 α ,25-Dihydroxyvitamin D ₃ (biologically active form of vitamin D)
25(OH)D₃	25-Hydroxyvitamin D ₃ (biomarker of vitamin D status)
AP1	Activator protein 1 (TF complex comprising FOS and JUN family subunits; responds to cytokines, growth factors, and stress signals)
ATAC-seq	Assay for transposase-accessible chromatin with high-throughput sequencing
BMI	Body mass index
<i>CAMP</i>	Cathelicidin antimicrobial peptide (innate immunity; vitamin D target gene)
<i>CD14</i>	CD14 molecule (monocyte marker; vitamin D target gene)
ChIP-seq	Chromatin immunoprecipitation sequencing
CPM	Counts per million (RNA-seq normalization unit)
<i>CSRNP1</i>	Cysteine and serine rich nuclear protein 1 (TF; vitamin D target gene)
CV	Coefficient of variation
<i>CYP24A1</i>	Cytochrome P450 24A1 (vitamin D catabolizing enzyme; vitamin D target gene)
<i>CYP27B1</i>	Cytochrome P450 27B1 (renal 1 α -hydroxylase; vitamin D target gene)
<i>CYP2R1</i>	Cytochrome P450 2R1 (hepatic 25-hydroxylase)
D-HEALTH	D-HEALTH Trial (Australian randomised controlled trial; 3,340 IU vitamin D ₃ /day for 5 years; all-cause mortality endpoint)
DEFB4	Defensin beta 4A (β -defensin; innate antimicrobial peptide; vitamin D target gene)
DEG	Differentially expressed gene
DESeq2	Differential expression analysis using a negative binomial model (Bioconductor R package)

DHCR7	7-Dehydrocholesterol reductase
DoRothEA	Discriminant Regulon Expression Analysis (Curated TF–target gene regulatory interaction database)
<i>EGR1</i>	Early growth response 1 (immediate-early stress TF)
ELISA	Enzyme-linked immunosorbent assay
ETS1 / ETS2	ETS proto-oncogene 1 / 2 (TFs)
FAIRE-qPCR	Formaldehyde-assisted identification of regulatory elements by quantitative polymerase chain reaction (PCR)
FC / logFC	Fold change / log ₂ fold change
FDR	False discovery rate (Benjamini-Hochberg correction)
<i>FOS</i>	Fos proto-oncogene, AP1 TF subunit
<i>FOSB</i>	FosB proto-oncogene, AP1 TF subunit
GC / VDBP	Vitamin D-binding protein (serum transport for vitamin D metabolites)
GEO	Gene Expression Omnibus (NCBI public repository for functional genomics data)
GRN	Gene regulatory network
H3K27ac	Histone H3 lysine 27 acetylation (active enhancer mark)
H3K4me3	Histone H3 lysine 4 trimethylation (active promoter mark)
HBB	Hemoglobin subunit beta
HIF1A	Hypoxia-inducible factor 1 subunit alpha (TF)
HLA-C	Human leukocyte antigen C (major histocompatibility complex (MHC) class I; vitamin D target gene)
<i>IFI44L</i>	Interferon induced protein 44 like (interferon-stimulated gene; baseline variability marker)
<i>IFIT1</i>	Interferon induced protein with tetratricopeptide repeats 1 (antiviral innate immune gene)
IL6 / IL8	Interleukin 6 / 8 (proinflammatory cytokines; response index biomarkers)
<i>ITGA2B</i>	Integrin subunit alpha 2B (focal adhesion; vitamin D target gene)

<i>ITGB3</i>	Integrin subunit beta 3 (focal adhesion; vitamin D target gene)
<i>ITGB5</i>	Integrin subunit beta 5 (focal adhesion; vitamin D target gene)
IU	International unit (standard unit for vitamin D dose; 1 IU = 0.025 µg vitamin D ₃)
IκBα	Inhibitor of kappa B alpha (encoded by NFKBIA; canonical NFκ- κB inhibitory protein)
<i>JCHAIN</i>	Immunoglobulin J chain (B cell/plasma cell marker; baseline variability gene)
JUN / JUNB	Jun / JunB proto-oncogenes, AP1 TF subunits
JUP	Junction plakoglobin (cell–cell adhesion protein; desmosomal/adherens junction component)
MDS	Multidimensional scaling (transcriptome visualization)
MHC	Major histocompatibility complex
<i>MYC</i>	MYC proto-oncogene, bHLH TF (master regulator of proliferation and metabolism)
<i>MYL9</i>	Myosin light chain 9 (focal adhesion effector; vitamin D target)
<i>MYLK</i>	Myosin light chain kinase (focal adhesion effector; vitamin D target)
NFKBIA / NFKBIZ	NFκB inhibitor alpha / zeta (conserved vitamin D response core genes)
NFκB / NFKB1	Nuclear factor kappa B / subunit 1 (master inflammatory TF)
NK	Natural killer (cell; innate immune lymphocyte)
ODE	Ordinary differential equation
PBMC	Peripheral blood mononuclear cell
PTH	Parathyroid hormone
RNA-seq	RNA sequencing (genome-wide transcriptome profiling)
RXR	Retinoid X receptor (VDR heterodimerization partner)
Saudi Cohort	Study of 100 Saudi participants supplemented with vitamin D ₃ bolus (Chapter 4)

<i>SIGLEC1</i>	Sialic acid binding immunoglobulin-type lectin 1 (monocyte/macrophage marker; baseline variability gene)
STAR	Spliced Transcripts Alignment to a Reference (ultrafast RNA-seq aligner)
STAT1 / STAT3	Signal transducer and activator of transcription 1 / 3
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins (protein interaction network database used for gene set enrichment visualisation)
T_H1	T helper 1 (lymphocyte subset; drives cell-mediated immunity; suppressed by vitamin D)
T_H17	T helper 17 (lymphocyte subset; drives mucosal inflammation; suppressed by vitamin D)
<i>THBD</i>	Thrombomodulin (anticoagulant glycoprotein; primary vitamin D target gene)
TMM	Trimmed mean of M-values (edgeR normalization)
TNF	Tumor necrosis factor (proinflammatory cytokine; response index biomarker)
<i>TNFAIP3</i>	TNF alpha-induced protein 3 (vitamin D target gene)
<i>TP53</i>	Tumor protein P53 (master stress-sensing TF)
Tregs	Regulatory T cells (suppressive lymphocytes; vitamin D target population)
<i>TRPV6</i>	Transient receptor potential vanilloid 6 (epithelial calcium channel; vitamin D target gene)
TSS	Transcription start site
UVB	Ultraviolet-B radiation (290–315 nm; required for skin vitamin D ₃ synthesis)
<i>VDR</i>	Vitamin D receptor (nuclear receptor TF)
VDRE	Vitamin D response element (VDR genomic binding site)
ViDA	Vitamin D Assessment study (New Zealand Randomised Controlled Trial; 100,000 IU monthly)

VITAL	VITamin D and OmegA-3 TriaL (US RCT; 2,000 IU/day; n=25,871)
VitDbol	Vitamin D bolus trial (NCT02063334; Kuopio, Finland; FAIRE-qPCR chromatin study; Chapter context only)
VitDHiD	Vitamin D Human <i>in vivo</i> Dose trial (NCT03537027; Kuopio, Finland; Chapter 3)
VitDmet	Vitamin D metabolism trial (NCT01479933; Kuopio, Finland; first response index study)
VitDPAS	Vitamin D Polish Academy of Sciences trial (NCT06104111; Olsztyn, Poland; Chapters 5–6)

Reader's Guide to the Studies in This Thesis

This thesis is built around three focal experimental studies (VitDHiD, the Saudi Cohort, and VitDPAS) conducted between 2018 and 2025, and draws on two earlier precursor studies (VitDmet and VitDbol) that established the key concepts and classification method. Each study has a short codename used consistently throughout the text. A reader coming to this work without prior familiarity with the Nutrigenomics team will find it helpful to understand these studies before reading further. The two precursor studies are described first as background while the three focal studies are the basis of the appended papers.

VitDmet (Vitamin D Metabolism Trial, NCT01479933): Conducted in Kuopio, Finland over five months in 71 pre-diabetic individuals. VitDmet measured the expression of vitamin D target genes, particularly *CD14* (CD14 molecule) and *THBD* (thrombomodulin), in PBMCs and adipose tissue by quantitative Polymerase Chain Reaction (qPCR) [18]. Crucially, it found that only the top half of participants ranked by their serum 25-hydroxyvitamin D₃ (25(OH)D₃) response showed a positive correlation between gene expression changes and vitamin D status, revealing for the first time that individuals differ reproducibly in their molecular responsiveness. VitDmet introduced the concept later formalised as the vitamin D response index. It is referenced in this thesis as historical context and is not one of the four appended papers.

VitDbol (Vitamin D Bolus Trial, NCT02063334): Conducted in Kuopio, Finland in 40 healthy adults receiving a single oral vitamin D₃ dose (2,000 micrograms), with blood collected before supplementation and at days 1, 2, and 30. Within 24 hours, serum 25(OH)D₃ and 1,25(OH)₂D₃ rose approximately 20%. Rather than genome-wide transcriptomics, VitDbol assessed immune cell chromatin accessibility at selected genomic loci by FAIRE-qPCR (Formaldehyde-assisted identification of regulatory elements by quantitative polymerase chain reaction), alongside PTH (parathyroid hormone) measurements, to classify participants into 14 high, 11 mid, and 10 low vitamin D responders [33]. This established both the bolus design and the individual responsiveness classification concept that all subsequent studies built upon [33]. VitDbol is referenced for context and its data are not appended.

VitDHiD (Vitamin D Human *in vivo* Dose Trial, NCT03537027): Conducted in Kuopio, Finland in 2018; 40 healthy adult participants. This is the first appended paper (Chapter 3) and the foundation of the thesis. Participants received a single high-dose vitamin D₃ bolus. Blood was taken before the dose (day 0) and 24 hours after (day 1). RNA-seq was performed on PBMCs. Key findings include 452 vitamin D target genes highlighting dominant upregulation of focal adhesion genes and first transcriptomic classification of individuals as high, mid, or low vitamin D responders.

Saudi Cohort: Conducted at King Abdulaziz University, Jeddah, Saudi Arabia; 100 adult participants with documented vitamin D deficiency. This is the second appended paper (Chapter 4). Participants received a bolus of vitamin D₃ and blood cytokines (IL6, IL8, and TNF (interleukin 6, interleukin 8, and tumor necrosis factor)) were measured by ELISA (Enzyme-linked immunosorbent assay). A subset of 37 participants also had seven vitamin D target genes quantified by qPCR. Full RNA-seq was not performed due to budget constraints. Key finding is that low vitamin D responders exist at comparable rates in this sun-rich, genetically distinct population, establishing geographic universality of the response index.

VitDPAS (Vitamin D Polish Academy of Sciences Trial, NCT06104111): Conducted in Olsztyn, Poland from 2023-25; 45 healthy adult participants. This study provides the data for both Chapter 5 and Chapter 6. Participants received three monthly bolus doses of vitamin D₃ (weight-adjusted: 1,000 IU (international units) per kg body weight) and blood was taken at seven time points across twelve weeks. RNA-seq was performed at every time point, generating 287 total sequencing libraries. Chapter 5 analyses the first acute response and cross-cohort comparison with VitDHiD. Chapter 6 uses the full longitudinal dataset for mechanistic mathematical modelling.

A note on the mathematical model (Chapter 6): The ODE model built in Chapter 6 is a mathematical description of how 11 key regulatory proteins interact over time to produce the observed gene expression patterns. No prior knowledge of differential equations is required to follow the biological argument: the model is a tool for asking “which molecular interactions are necessary and sufficient to explain why high responders respond more than low responders?”

Chapter 1

General Introduction

*“In questions of science, the authority of a thousand is not worth
the humble reasoning of a single individual.”*

— Galileo Galilei

1.1 Vitamin D: From Sunlight to Endocrine Signal

Among the molecules that influence human physiology, vitamin D occupies an exceptional position. It is not primarily a dietary nutrient but its metabolite $1,25(\text{OH})_2\text{D}_3$ acts as a hormone manufactured by the body itself, provided the skin receives sufficient ultraviolet-B (UVB) radiation in the 290–315 nm wavelength range. When UVB photons strike skin cells, they drive the photolysis of 7-dehydrocholesterol into previtamin D_3 , which isomerises thermally to vitamin D_3 within hours [\[11\]](#). This is the starting point of a two-organ metabolic activation cascade. In the liver, the enzyme CYP2R1 (Cytochrome P450 2R1) adds a hydroxyl group at carbon 25, generating $25(\text{OH})\text{D}_3$, the major circulating storage form whose serum concentration is the standard clinical measure of an individual’s vitamin D status. In the kidney, a second enzyme, CYP27B1, adds a further hydroxyl group at carbon 1, converting $25(\text{OH})\text{D}_3$ to $1,25(\text{OH})_2\text{D}_3$, the biologically active hormone, that enters cells and activates the VDR [\[11\]](#). A third enzyme, CYP24A1, which is itself switched on by the VDR, degrades both circulating forms and thereby limits the duration and intensity of the signal, acting as an autoregulatory brake.

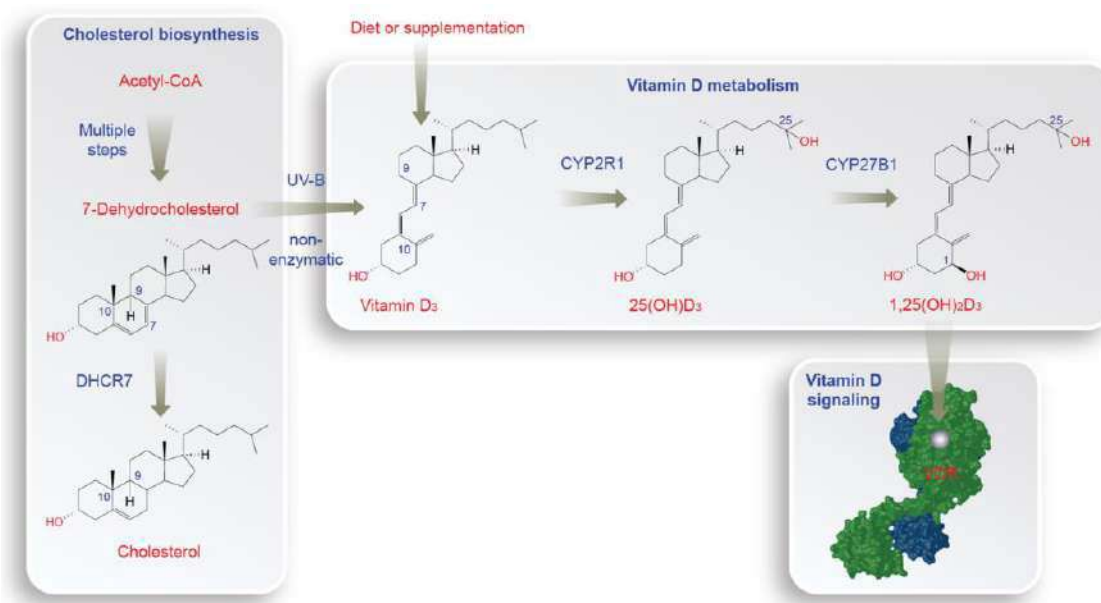


Figure 1. The vitamin D metabolic pathway. Vitamin D₃ is synthesised in the skin from 7-dehydrocholesterol by UVB light, then hydroxylated in the liver (CYP2R1 → 25(OH)D₃) and kidney (CYP27B1 → 1,25(OH)₂D₃). The active hormone binds VDR in immune and other cells. CYP24A1 mediates feedback inactivation. Taken from Carlberg C & Velleuer E (2022) *Biochem Pharmacol* 196:114735 [15].

Despite the sun's availability to virtually the entire world, vitamin D deficiency affects an estimated one billion people [35, 37]. The reason is behavioural rather than geographical. Indoor work and leisure, widespread sunscreen use, clothing that covers the skin, and at high latitudes the winter solar angle that prevents UVB from penetrating the atmosphere have together made cutaneous synthesis inadequate for much of the modern population [38, 59]. The clinical consequences extend far beyond the bones, because the VDR is broadly expressed across immune cells, intestine, kidney, and skin, though expression levels vary considerably by cell type.

1.2 VDR and the Genomic Programme It Controls

The biological actions of 1,25(OH)₂D₃ are mediated through the VDR, a 427-amino acid protein belonging to the nuclear receptor superfamily of ligand-activated TFs. When 1,25(OH)₂D₃ binds VDR with sub-nanomolar affinity, VDR changes shape and pairs with a partner protein called the retinoid X receptor (RXR). The VDR-RXR heterodimer travels into the cell nucleus (or is already present there), finds specific DNA sequences called vitamin D response elements (VDREs), which typically consist of two hexameric motifs

separated by three nucleotides [47], and recruits a machinery of helper proteins that remodels chromatin and drives or suppresses the transcription of nearby genes. The result is a wave of coordinated gene expression changes spanning hundreds of loci simultaneously.

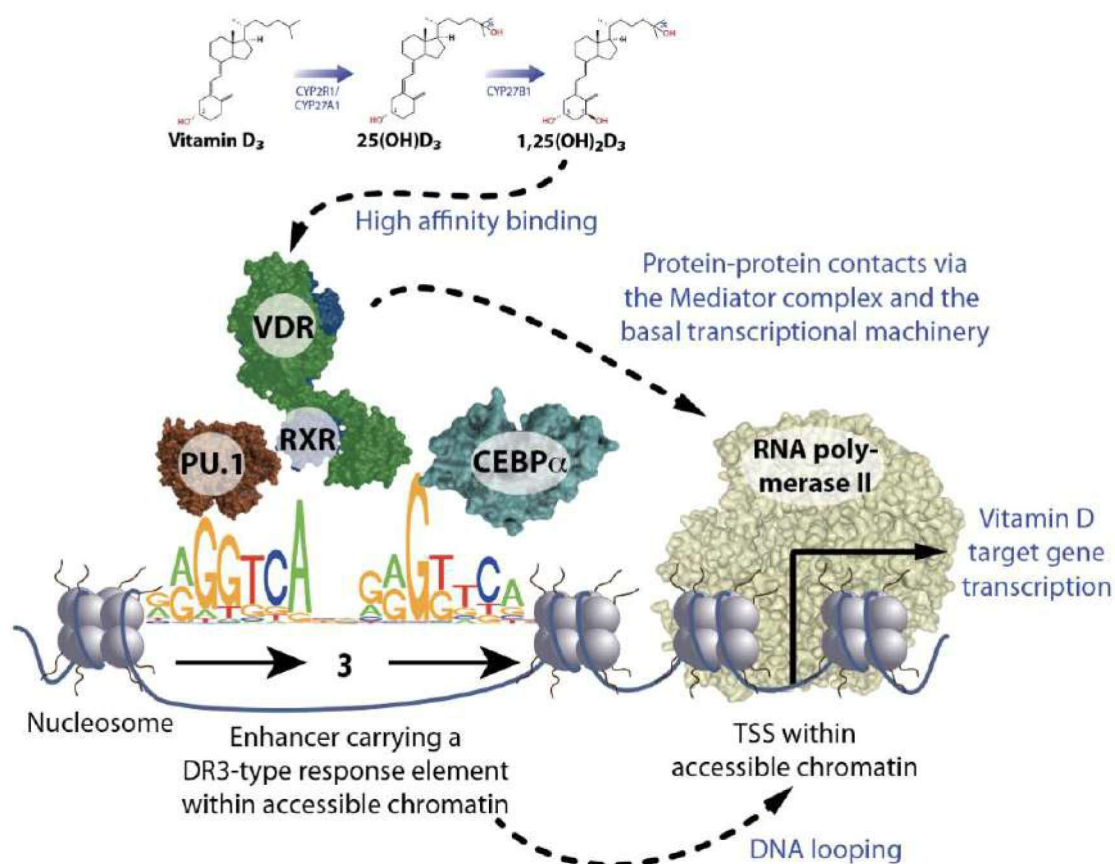


Figure 2. VDR-mediated transcription. 1,25(OH)₂D₃ binds VDR, triggering heterodimerisation with RXR, VDRE binding, and co-activator recruitment, resulting in target gene activation or repression. Taken from Carlberg C & Velleuer E (2024) *Nutrients* 16:398 [16].

VDR's presence in immune cells, intestinal epithelium, bone, pancreas, and adipose tissue gives vitamin D regulatory influence across an extraordinary breadth of physiology like calcium absorption, immune defence, cell proliferation, and metabolic regulation. Genome-wide mapping of VDR binding sites has identified thousands of potential response elements [47, 49], suggesting that vitamin D signalling is woven into a large fraction of the human gene regulatory network (GRN). This breadth is both the biological appeal of vitamin D and the source of the central problem that motivates this thesis.

1.3 Physiological Functions: Beyond Calcium and Bone

The canonical role of vitamin D in calcium homeostasis is well established i.e. VDR activation in intestinal cells induces the calcium transport proteins TRPV6 (transient receptor potential vanilloid 6) and calbindin-D9K, increasing dietary calcium absorption. Through modulation of PTH secretion, vitamin D also regulates calcium release from bone and reabsorption in the kidney. Severe deficiency in children causes rickets while in adults it leads to osteomalacia. These consequences motivated twentieth-century fortification of dairy products and remain the primary justification for clinical supplementation guidelines [\[37, 41\]](#).

The immune roles of vitamin D are equally important but far more complex. Macrophages and dendritic cells express CYP27B1 and can locally activate vitamin D [\[6\]](#), directly linking pathogen sensing to VDR-mediated transcription. VDR activation in these cells drives expression of cathelicidin (CAMP) and β -defensin (DEFB4), broad-spectrum antimicrobial peptides that kill bacteria and viruses directly [\[42\]](#). In adaptive immunity, $1,25(\text{OH})_2\text{D}_3$ suppresses the differentiation of pro-inflammatory T helper 1 ($\text{T}_{\text{H}}1$) and T helper 17 ($\text{T}_{\text{H}}17$) lymphocytes while promoting regulatory T cells (Tregs), thereby shifting the balance toward immune tolerance [\[3\]](#). Beyond these established domains, epidemiological associations implicate vitamin D status in cancer risk, cardiovascular disease, type 2 diabetes, and respiratory infection susceptibility [\[58\]](#). Converting these associations into clinical benefit through supplementation has, however, proven extraordinarily difficult, for reasons that go to the heart of this thesis.

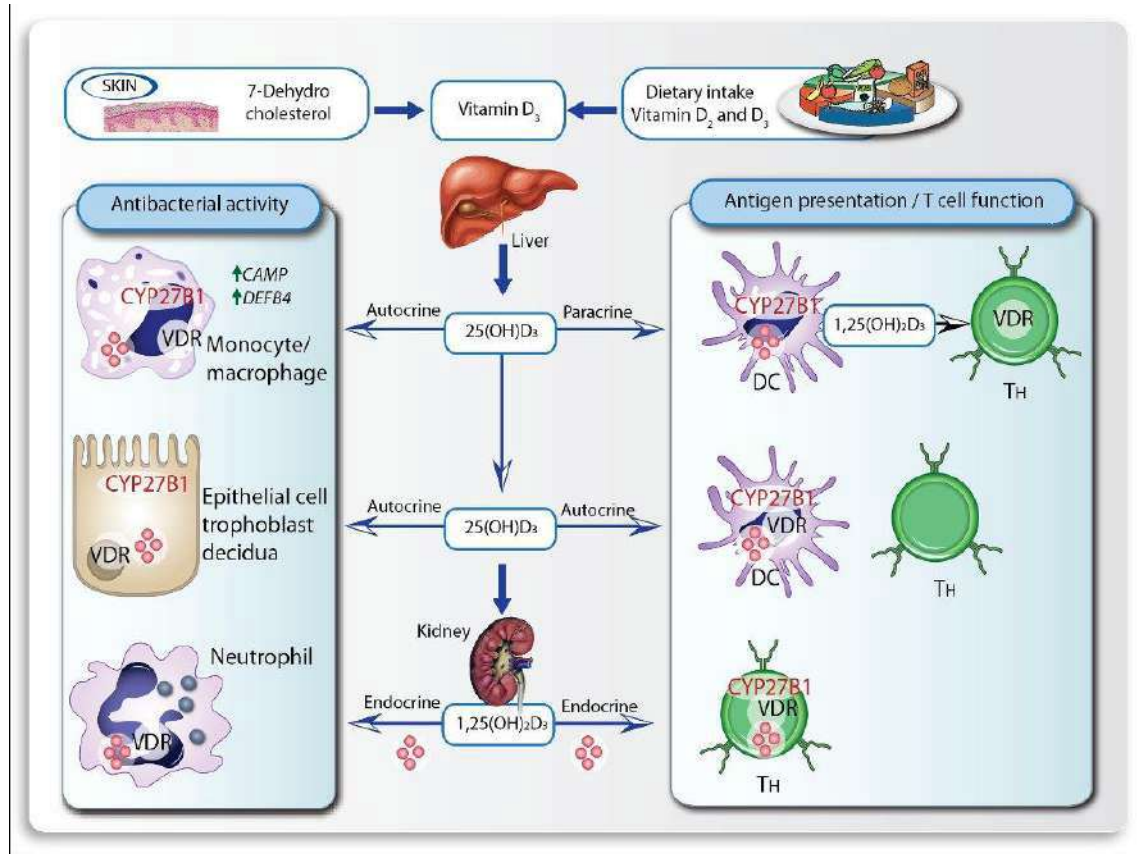


Figure 3. Vitamin D and immune cell activation. The immune roles of vitamin D: monocytes/macrophages and epithelial cells express CYP27B1, locally converting 25(OH)D₃ to 1,25(OH)₂D₃ (autocrine/paracrine), boosting antibacterial activity (CAMP, DEFB4). Kidney-derived endocrine 1,25(OH)₂D₃ modulates antigen presentation by dendritic cells (DC) and T helper cell (T_H) function. Taken from Carlberg C & Molnár F (2025) *Nutrigenomics*. Springer [13].

1.4 Why Large-Scale Trials Have Failed: The Heterogeneity Problem

The clinical importance of vitamin D deficiency motivated some of the largest nutritional intervention trials ever conducted. The VITAL (VITamin D and Omega-3 Trial) trial enrolled 25,871 American adults [44] and followed them for 5.3 years with 2,000 IU of vitamin D₃ per day, finding no significant reduction in cancer incidence or cardiovascular events [44]. The D-HEALTH trial administered 3,340 IU per day for five years in Australia and found no effect on all-cause mortality [45]. ViDA (Vitamin D Assessment study; 100,000 IU monthly, New Zealand) similarly failed to reduce cardiovascular disease or respiratory infections [53]. Across these trials, a consistent and frustrating pattern emerges: no significant average effect, yet post-hoc subgroup analyses frequently identify apparent responders.

This pattern has a straightforward biological explanation that the trials were not designed to detect. If approximately one third of participants are constitutively low responders (individuals whose genome-wide transcriptional response to vitamin D is minimal, formally classified by the vitamin D response index; see Section 3.3) whose transcriptional machinery cannot be meaningfully activated by vitamin D regardless of how much their serum 25(OH)D₃ rises, then including them in a trial designed to detect an average effect will dilute the true signal in the responsive subgroup to statistical insignificance. The fundamental problem with the large trials is not statistical power, but scientific framing: rather than asking what the average population effect of vitamin D supplementation is, the more productive question is which individuals respond, to what degree, and through what mechanism.

This reframing is the intellectual foundation of the entire body of work described in this thesis.

1.5 The Studies in This Thesis: Names, Designs, and the Conceptual Progression

Before entering the scientific detail, it is important to introduce the four studies whose findings constitute this thesis, because their codenames appear throughout. The studies are introduced more fully in the Reader's Guide preceding this chapter; here, the scientific logic connecting them is explained.

The research programme began with two predecessor studies that are not appended to this thesis but provide essential context. VitDmet [18] was the first study to track the gene expression response to vitamin D₃ supplementation in individuals over time and discovered that they differed reproducibly in how many genes they activated, regardless of dose. This reproducibility was extraordinary and suggested that responsiveness is a stable biological property of the individual, not a dose-dependent threshold. VitDbol refined the approach by administering a single large oral dose (a bolus) and sampling blood 24 hours later, creating a sharp, pharmacokinetically tractable signal that isolated direct vitamin D target gene activation from later regulatory feedbacks. From these two studies, the concept of the vitamin D response index was formally developed. Each individual receives a score based

on the number and magnitude of their transcriptional changes, and this score is reproducible [27, 30].

VitDHiD [27] deployed genome-wide RNA-seq for the first time in this series, replacing the limited qPCR readouts used in earlier work. This choice was transformative in revealing that the dominant biological programme activated by vitamin D *in vivo* was focal adhesion and cytoskeletal reorganisation, a programme absent from the prior *in vitro* literature and could never have been found with a targeted approach. The Saudi cohort (Chapter 4) tested whether the low-responder phenotype was specific to Northern Europeans or universal. 100 Saudi participants were supplemented and assessed, with cytokine measurements in all and qPCR of seven target genes in 37 participants. Full RNA-seq was not possible due to budget constraints, but the available data confirmed universality. VitDPAS (Chapters 5 and 6) combined larger participant numbers, weight-adjusted dosing, and three monthly supplementation cycles to generate both a rigorous cross-cohort replication of VitDHiD findings and the rich longitudinal dataset needed for mechanistic modelling.

The progression from VitDmet through VitDHiD, Saudi, and VitDPAS is therefore not arbitrary but deliberate. Each study was designed to answer the question that the previous study's findings had raised. VitDmet asked whether individual differences exist; VitDHiD asked what genes underlie them and whether RNA-seq reveals unexpected biology; Saudi asked whether the finding is geographically universal; VitDPAS asked what the conserved molecular core looks like and what mechanisms determine the response phenotype.

1.6 From Targeted Assays to Genome-wide RNA-seq: A Necessary Technical and Conceptual Advance

The progression from targeted molecular assays in VitDmet and VitDbol to genome-wide RNA-seq in VitDHiD represents both a technical and a conceptual leap forward. VitDmet quantified expression of a handful of pre-selected vitamin D target genes by qPCR, and VitDbol assessed chromatin accessibility at selected loci by FAIRE-qPCR. These are powerful and precise methods for the questions they were designed to answer, but they share a fundamental limitation: they measure only what the experimenter already knows to look for. Every gene panel, every set of FAIRE primers, reflects prior hypotheses formed largely from *in vitro* experiments.

RNA-seq overcame this limitation in two distinct ways. The first is conceptual i.e, by measuring all expressed transcripts simultaneously without pre-selection, RNA-seq is genuinely hypothesis-free and can reveal biological programmes that no prior knowledge would have predicted. The second is technical: modern RNA-seq platforms quantify gene expression across the entire transcriptome with dynamic ranges exceeding five orders of magnitude, at single-nucleotide resolution, with sensitivity sufficient to detect low-abundance regulatory transcripts that would be invisible to any targeted panel. Alignment to reference genomes and normalisation procedures such as trimmed-mean-of-M-values (TMM) allow rigorous cross-sample comparisons at the level of raw transcript counts [50, 60]. These technical properties collectively enable the kind of discovery-driven analysis that yielded the focal adhesion signature in VitDHiD.

The application of RNA-seq to VitDHiD immediately revealed a biological programme that VitDmet and VitDbol could not have detected: the dominant *in vivo* vitamin D transcriptional response in human PBMCs involves the coordinated upregulation of focal adhesion pathway genes, including *ITGA2B* (integrin alpha IIb), *ITGB3* (integrin beta 3), *ITGB5* (integrin beta 5), *MYL9* (myosin light chain 9), and *MYLK* (myosin light chain kinase), none of which appeared in any prior vitamin D gene panel. This discovery required hypothesis-free genome-wide measurement [27].

The Saudi cohort illustrates what is lost when budget constraints require a return to targeted methods: measuring only seven genes by qPCR and three cytokines by ELISA can confirm the presence of the response index phenotype and validate its cross-population universality, but cannot characterise the full transcriptional programme or generate the mechanistic resolution that genome-wide data enables. The findings of Chapter 4 are therefore complementary to, rather than fully replicative of, Chapters 3 and 5 [2].

1.7 Why PBMCs? The Rationale for Peripheral Blood Immune Cells

All four studies in this thesis measure gene expression in PBMCs which is the mixed population of lymphocytes (T cells, B cells, natural killer (NK) cells) and monocytes that circulate in the bloodstream and can be isolated from a standard venepuncture by density gradient centrifugation. The choice is deliberate and merits explanation.

The most fundamental reason is access. Vitamin D exerts important biological effects in muscle, liver, intestine, and other metabolically active tissues, but obtaining repeated biopsies from these tissues in healthy participants receiving a nutritional supplement is ethically unjustifiable and practically impossible at the participant numbers required for transcriptomic power. A venepuncture, by contrast, is routine, low-risk, and can be performed repeatedly in the same individual across multiple time points. This is the only design that can generate the longitudinal datasets required to understand how the transcriptional response to vitamin D evolves across cycles of supplementation.

The biological justification is equally strong. PBMCs are among the tissues most profoundly regulated by vitamin D [\[3, 6\]](#), VDR expression is high in monocytes and lymphocytes, and their transcriptional responses to VDR activation are extensive. Critically, circulating immune cells are a genuine systemic readout of the organism's immune regulatory state, not an approximation of it. When vitamin D signalling is altered, the gene expression changes in circulating PBMCs reflect both the direct transcriptional effects of VDR activation in those cells and the broader hormonal and cytokine landscape in which they circulate.

PBMCs are also not immunologically homogeneous: they comprise monocytes, CD4⁺ T cells, CD8⁺ T cells, B cells, and NK cells, each with distinct VDR expression levels. The bulk RNA-seq approach used here measures a weighted average across these populations, which captures the integrated multicellular immune response. Finally, the use of standardised PBMC isolation protocols enables rigorous cross-cohort comparison between Finnish (VitDHID) and Polish (VitDPAS) datasets, which is central to the cross-cohort validation strategy of this thesis [\[20, 27, 30, 34\]](#).

1.8 Bolus versus Daily Vitamin D Supplementation: A Pharmacological Rationale

All studies in this thesis employ a bolus design: a single large oral dose of vitamin D₃ (typically 80,000 to 100,000 IU) rather than a daily small dose. This choice reflects both pharmacokinetic reasoning and experimental logic.

Daily vitamin D₃ supplementation at conventional doses (600–4,000 IU per day) raises serum 25(OH)D₃ [\[35, 36\]](#) slowly over weeks, reaching a new steady state without

generating a clear temporal signal. At any single transcriptomic measurement time point during gradual supplementation, the observed gene expression reflects a mixture of direct VDR target induction, secondary regulatory effects, feedback adaptations, and ongoing baseline transcription, all superimposed without temporal separation. It is essentially impossible to identify direct VDR target genes from this signal.

A bolus dose creates a pharmacokinetic pulse: serum 25(OH)D₃ rises sharply within 24–48 hours and then declines over subsequent weeks as hepatic hydroxylation temporarily saturates and CYP24A1-driven catabolism accelerates. By sampling blood at day 0 (immediately before the bolus) and day 1 (24 hours after), one captures the acute transcriptional response before secondary regulators have had time to operate. Genes that are differentially expressed at day 1 are therefore strong candidates for early, direct VDR genomic targets. This temporal precision is impossible with gradual supplementation [\[18, 19\]](#).

The multi-cycle VitDPAS design took this logic further: three monthly bolus doses with blood sampling at seven time points across twelve weeks enabled tracking of how the transcriptional response changes with repeated VDR activation. As described in Chapter 6, this revealed progressive temporal amplification of the response, with the number of differentially expressed genes growing from approximately 758 after the first cycle to roughly 2,600 after the third, suggesting epigenetic priming of VDR target gene loci across repeated cycles of activation. This deepening of the response over time is entirely invisible in short daily supplementation trials and provides a compelling biological explanation for why brief intervention studies consistently underestimate vitamin D's effects [\[28, 30\]](#).

1.9 *In Vivo* Human Studies versus *In Vitro* Cell Culture: Why They Don't Always Agree

A reader familiar with the vitamin D literature will notice that many findings in this thesis differ from those of *in vitro* cell culture studies. This discordance is not a methodological failure but a predictable consequence of the fundamental differences between the two experimental contexts.

In vitro studies typically treat cultured cell lines or freshly isolated cells with supraphysiological concentrations of 1,25(OH)₂D₃ (commonly 10-100 nM), bypassing the

metabolic activation cascade and the serum binding proteins that govern how much active hormone actually reaches immune cells *in vivo* [10]. In the bloodstream, more than 99% of circulating 1,25(OH)₂D₃ is bound to the vitamin D binding protein (GC/VDBP), making the free fraction available to cells orders of magnitude lower than concentrations used in culture experiments. This means that *in vitro* studies identify genes that are VDR-responsive under forced conditions, but many of those genes may not receive sufficient hormonal signal in a living person to be meaningfully regulated.

Cell lines also carry profoundly remodelled epigenomes after dozens of passages in *in vitro* culture: their chromatin accessibility landscapes, histone modification patterns, and TF occupancy profiles differ fundamentally from those of freshly isolated immune cells in a human donor. A gene that is transcriptionally accessible and VDR-responsive in a cell line like THP-1 (human monocytic leukaemia) may be inaccessible in a freshly isolated human monocyte, and vice versa. This is precisely why the focal adhesion signature was invisible in prior *in vitro* work: integrin genes are not prominently regulated by supraphysiological 1,25(OH)₂D₃ in standard cell lines, but they constitute a major part of the physiological VDR transcriptional programme in human immune cells.

In vivo human intervention studies with genome-wide transcriptomics are therefore not merely confirmatory of cell culture findings but represent a qualitatively distinct and more physiologically meaningful level of understanding. The findings of this thesis should be read with this distinction in mind [9, 27, 30, 48].

1.10 The Evolution of the Vitamin D Response Index

The concept of the vitamin D response index has undergone a progressive transformation across the course of this research programme, from an empirical observational score to a mechanistically grounded, quantitatively predictable property of each individual's gene regulatory architecture.

In its earliest form (VitDmet), the response index was a simple count: participants were ranked by the number of genes that changed expression after supplementation, and this rank order was found to be stable across repeated measurements. Useful for demonstrating that individual differences are real and reproducible, but offering no mechanistic explanation [12, 18, 55].

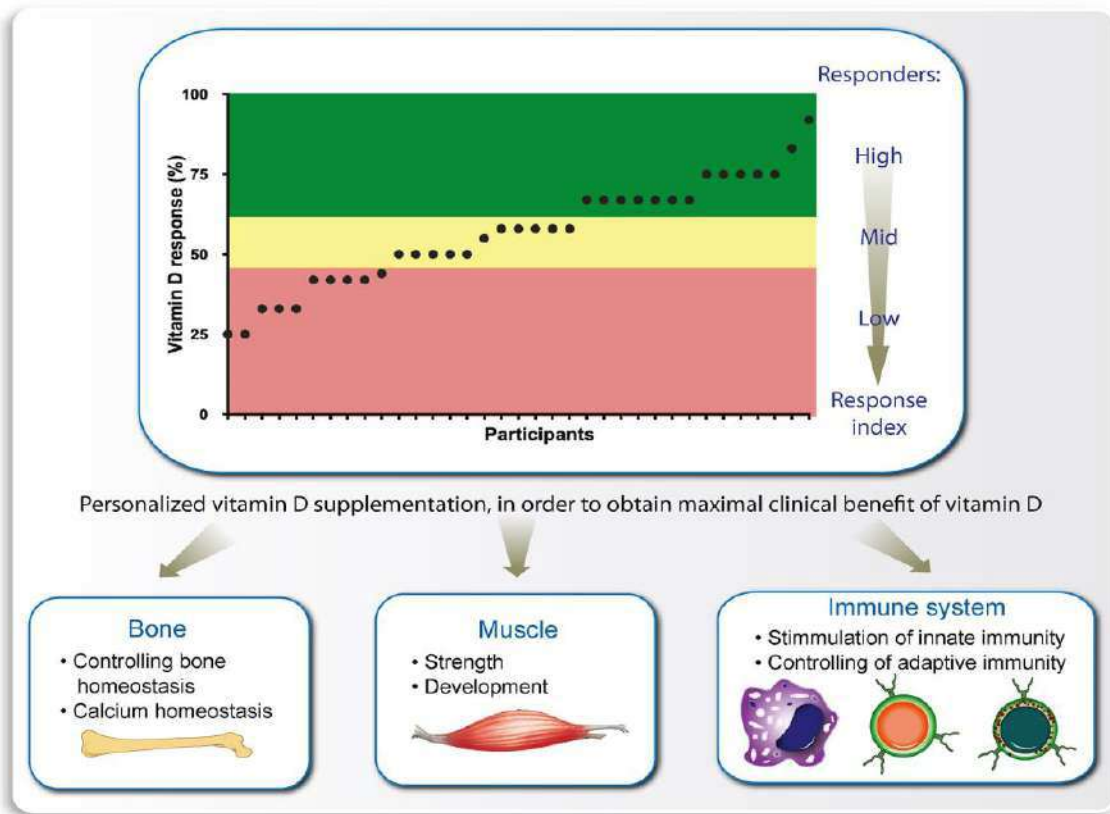


Figure 4. Personalised vitamin D supplementation framework. Interindividual variation in the vitamin D response index across participants (high, mid, low responders) guides personalised vitamin D₃ dosing, with downstream clinical benefits for bone, muscle, and immune health. Taken from Carlberg C, Raczyk M & Zawrotna N (2023) *Redox Biol* 62:102695 [17].

The subsequent refinement and mechanistic explanation of the response index across VitDHiD, VitDPAS, and the ODE modelling study is discussed in Section 8.6 [27, 28, 30].

1.11 Mechanistic Modelling: Making Sense of Biological Complexity

Even the most comprehensive transcriptomic dataset, covering hundreds of individuals, thousands of genes, and several time points, cannot by itself reveal the causal regulatory mechanism that determines why some people respond more than others. Knowing that *FOS* (*Fos proto-oncogene, AP1 TF subunit*) or *JUN* (*Jun proto-oncogene, AP-1 TF subunit*) expression is highly variable at baseline, or that *MYC* expression tracks response amplitude, tells us that these genes are important but not how they are connected, which direction the causality runs, or how their dynamics evolve over time. This is where mechanistic mathematical modelling becomes indispensable.

GRN modelling formalises the regulatory relationships between TFs and their target genes as a directed graph: each node is a gene or protein while each edge is an activation or repression interaction with a defined strength. In principle, a complete GRN model of VDR-regulated immunity would contain thousands of nodes. In practice, such a model is computationally intractable and biologically uninformative because with everything connected to everything, no individual interaction can be identified as load-bearing.

The key conceptual advance applied in Chapter 6 comes from the work of Kim et al. [40], who showed that the most biologically important information in a complex regulatory network is typically carried by a small “kernel” of highly connected, functionally critical nodes. Iterative reduction proceeds by collapsing low-degree, peripheral nodes while preserving regulatory logic through propagation: each removed indirect path $A \rightarrow B \rightarrow C$ is replaced by a direct $A \rightarrow C$ edge whose sign and magnitude are the product of the corresponding upstream values, so that the regulatory consequence of B is preserved even after B itself is dropped. The procedure terminates when no further node falls below the chosen connectivity threshold, leaving a kernel that is structurally minimal but dynamically equivalent to the parent network for the purposes of mechanistic modelling.

The starting point for kernel reduction is itself a curated regulatory network. In this thesis the network used in Chapter 6 is built on top of DoRothEA (Discriminant Regulon Expression Analysis) [24], a publicly available human regulon resource that catalogues directed transcription-factor – target gene interactions assembled from four orthogonal evidence streams: literature-curated interactions, ChIP-seq (Chromatin Immunoprecipitation Sequencing) peak calls mapped to nearby promoters, *in silico* TF binding-site prediction, and inferred associations from large gene-expression compendia. Each interaction in DoRothEA carries a confidence level (A–E, with A indicating the highest level of multi-evidence support) and a signed mode of regulation that distinguishes activating from repressive edges. By restricting the input network to high-confidence, signed interactions, DoRothEA provides a biologically defensible prior over which TFs regulate which targets and in which direction, without committing the model to any single regulatory hypothesis [24]. In Chapter 6, this prior is combined with the longitudinal VitDPAS data to identify a cohort-relevant sub-network and is then condensed by iterative kernel reduction into the minimal regulatory module that is fitted with ordinary differential equations. The size of that minimal module, the parameter values that emerge from fitting, and the cross-cohort generalisation performance are reported in Chapter 6.

The mathematical framework for this modelling was developed in collaboration with Dr. Elisa Domínguez-Hüttinger (Instituto de Investigaciones Biomédicas, UNAM, Mexico City), whose work on ODE-based models of immune regulation in skin disease provided the methodological blueprint. In their work on atopic dermatitis and skin-barrier dynamics, Domínguez-Hüttinger and colleagues demonstrated that Hill-type ODEs can capture the sigmoidal, switch-like behaviour of TF activation in immune cells with quantitative precision. Moreover, they showed that differential evolution, a global optimisation algorithm, can efficiently navigate the high-dimensional parameter spaces characteristic of regulatory network models [10,11]. The same methodological principles were adapted and applied to the vitamin D response context in Chapter 6: Hill-type kinetics model the cooperative activation and repression relationships between the kernel TFs. Then differential evolution optimisation in the Julia programming language (a high-performance open-source language for numerical computing, suited to differential equation systems) fits the model parameters to the longitudinal VitDPAS time series, and pharmacokinetic modelling of individual serum 25(OH)D₃ trajectories provides the hormone input function. The result is intended not as curve-fitting to known data but as a mechanistically interpretable model: it describes the biological interactions that generate the observed dynamics and produces hypotheses about which specific regulatory interactions distinguish high from low responders. The size of the resulting kernel, its parameter values and its cross-cohort behaviour are reported in Chapter 6.

Chapter 2

Objectives and Hypothesis

2.1 Objectives

The central objective of this thesis is to determine why individuals differ so profoundly in their molecular responses to vitamin D₃ supplementation, and to move from a descriptive observation to a mechanistic, quantitatively predictive understanding of the regulatory architecture that governs this variability. The thesis pursues this objective through four sequential studies whose findings are mutually reinforcing.

Stage 1 (Chapter 3; VitDHiD). Transcriptomic discovery addressing two questions simultaneously. First: what genes does vitamin D actually regulate in immune cells of living disease-free humans, and what biological programmes do they constitute? This aimed to produce the first genome-wide, *in vivo* map of VDR target genes in PBMCs, going beyond classical targets identified in cell culture. Second: do individuals differ reproducibly in the extent of their transcriptional response to vitamin D₃ supplementation? This aimed to develop and validate the first transcriptomics-based vitamin D response index from RNA-seq data, enabling quantitative classification of individuals as high, mid, or low responders.

Stage 2 (Chapter 4; Saudi Cohort). Cross-population universality: To determine whether the low-responder phenotype is specific to Northern European populations or universal, by applying the response index framework to a genetically and geographically distinct cohort using cytokine and qPCR readouts.

Stage 3 (Chapter 5; VitDPAS). Cross-cohort integration: To replicate VitDHiD findings in an independent Central European cohort, identify the high-confidence conserved core of regulated genes, and characterise the chromatin and regulatory logic of the conserved response mechanism.

Stage 4 (Chapter 6; VitDPAS Modelling). Mechanistic systems identification: To apply ODE-based regulatory network modelling to the complete longitudinal VitDPAS transcriptomic dataset, identify the minimal TF kernel that mechanistically governs the

response index, predict individual response phenotypes from baseline molecular profiles, and validate the model in the independent Finnish cohort.

2.2 Research Hypotheses

The general hypothesis of this thesis is that interindividual variability in transcriptional responses to vitamin D₃ supplementation is a reproducible, biologically meaningful, and mechanistically explicable property of each individual's pre-existing immune regulatory architecture. This general hypothesis is decomposed into five specific, testable sub-hypotheses.

Hypothesis 1: Vitamin D₃ supplementation activates a specific and reproducible transcriptional programme in human immune cells *in vivo* that is distinct from what *in vitro* models predicted, including biological pathways not previously associated with vitamin D signalling. Tested in Chapter 3.

Hypothesis 2: A reproducible vitamin D response index exists in human PBMCs. Individuals can be classified as high, mid, or low responders from RNA-seq data, and this classification will replicate across measurements and may correlate with independent immune biomarkers. Tested in Chapter 3.

Hypothesis 3: The low-responder phenotype is universal and not population-specific. Low vitamin D responders will appear at comparable frequencies in a genetically and geographically distinct population, ruling out pharmacokinetic explanations. Tested in Chapter 4.

Hypothesis 4: A conserved core of vitamin D target genes exists that is consistently regulated across independent cohorts, geographies, and repeated supplementation cycles. Identifying this core will provide the robust foundation for the response index. Tested in Chapter 5.

Hypothesis 5: Interindividual variability in the vitamin D response index can be mechanistically determined by the pre-existing regulatory interactions among a small kernel of TFs, capturable by an ODE model that distinguishes responder phenotypes and generalises across independent cohorts. Tested in Chapter 6.

Chapter 3

***In Vivo* Vitamin D Targets Reveal the Upregulation of Focal Adhesion-Related Genes in Primary Immune Cells of Healthy Individuals**

*“The most exciting phrase to hear in science, not ‘Eureka!’ but
‘That’s funny...’”*

— Isaac Asimov

Background

VitDHiD enrolled 40 healthy Finnish adults during winter (baseline 25(OH)D₃ 15.6 to 49.8 ng/ml (mean: 29.5 ng/ml), of which 25 were selected for detailed inspection of their transcriptomes. A single high-dose oral vitamin D₃ bolus was administered; PBMCs isolated at day 0 and day 1 for genome-wide RNA-seq, the first transcriptome-wide measurement in this programme. All bioinformatic analysis was performed by the candidate as lead bioinformatician [27].

Key findings

Differential expression analysis identified 452 significantly regulated genes at day 1 (FDR<0.05). The dominant enriched programme involved the focal adhesion pathway, including integrins (*ITGA2B*, *ITGB3*, *ITGB5*) and cytoskeletal effectors (*MYL9*, *MYLK*), invisible to prior targeted panels and not predicted by any *in vitro* study [27]. Classical VDR targets (*CAMP*, *CYP24A1*, *CD14*, *THBD*) and immune regulators (*NFKB1A* [NFκB inhibitor alpha, encoding IκBα]), *CSRNP1* [cysteine and serine rich nuclear protein 1]) were also present.

k-means clustering of transcriptome-wide response scores (per-individual statistics quantifying genome-wide transcriptional change after supplementation) produced the first RNA-seq-based vitamin D response index: high (n=4), mid (n=12), and low (n=9) responders. *HLA-C* (human leukocyte antigen C, a major histocompatibility class I (MHC) class I gene involved in immune cell recognition) was four-fold higher in high versus low responders at day 0, before supplementation,

implying the index reflects a pre-existing regulatory state. Within the 452 shared targets, individual response directions diverged sharply [27, 51].

Significance for the thesis

VitDHiD simultaneously answers what vitamin D transcriptionally regulates in living human immune cells and whether individuals differ reproducibly. The 452-gene target set and three-tier response index are the analytical foundation for Chapters 4, 5, and 6.

Author contribution

The candidate led all bioinformatic analysis: RNA-seq QC, alignment, normalisation, differential expression, pathway enrichment, clustering, and most figure generation. Study organization and wet-laboratory work was performed by co-authors.

Publication details

Also referred to as: VitDHiD

Ranjini Ghosh Dastidar¹, Julia Jaroslawska¹, Marjo Malinen², Tomi-Pekka Tuomainen³, Jyrki K. Virtanen³, Igor Bendik⁴ & Carsten Carlberg¹

¹ InLife Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn ² South-Eastern Finland University of Applied Sciences ³ University of Eastern Finland, Kuopio ⁴ DSM Nutritional Products, Basel

** The full text of this publication is appended at the end of this thesis as Appended Paper I.*

Chapter 4

Evaluation of the Vitamin D Response Index in a Saudi Cohort

“It is not the strongest of the species that survives, nor the most intelligent, but the one most responsive to change.”

— Charles Darwin

Background

Saudi Arabia presents a paradox: intense year-round sun exposure yet high vitamin D deficiency. One hundred Saudi participants received an 80,000 IU oral bolus of vitamin D₃, mirroring VitDHiD. Resource constraints required cytokine measurement (IL6, IL8, TNF) and qPCR of seven vitamin D target genes: *ASAP2* (ArfGAP with SH3 domain, ankyrin repeat and PH domain 2), *CAMP*, *CD14*, *FBP1* (fructose-bisphosphatase 1), *ITGAM* (integrin subunit alpha M), *LRRC25* (leucine rich repeat containing 25) and *NINJI* (ninjurin 1). All statistical analysis was performed by the candidate; study organization and wet-laboratory work by King Abdulaziz University collaborators [2].

Key findings

Cytokine-based classification found 39% of Saudi participants were low responders, comparable with VitDHiD's frequency despite different baseline status, genetics, and geography [2]. qPCR-based classification in 37 participants showed 88.6% concordance with cytokine classification, validating targeted assays as practical alternatives.

Significance for the thesis

Chapter 4 establishes the geographical and genetic universality of the response index. Although higher than average (~40%), low-responder prevalence is conserved across populations, essential for Chapter 6's mechanistic model to represent general human biology [2].

Author contribution

The candidate led all statistical analysis, scoring algorithm development, concordance analysis, and figure generation. Study organization and all wet-laboratory work was performed by King Abdulaziz University collaborators.

Publication details

Also referred to as: Saudi Cohort

Shareefa A. AlGhamdi^{2*}, Ranjini Ghosh Dastidar^{1*}, Maciej Rybiński¹, Hadeil M. Alsufiani², Sawsan O. Khoja², Nusaibah N. Enaibsi², Safa F. Saif² & Carsten Carlberg¹

¹ InLife Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn ² King Abdulaziz University, Jeddah, Saudi Arabia * **equal contribution**

***** *The full text of this publication is appended at the end of this thesis as Appended Paper II.*

Chapter 5

Transcriptomic Profiling of Immune Modulation Induced by Vitamin D₃ in the VitDPAS and VitDHiD Cohort Studies

“In nature we never see anything isolated, but everything in connection with something else.”

— Johann Wolfgang von Goethe

Background

VitDPAS enrolled 45 healthy Polish participants from Olsztyn, who received a single weight-adjusted vitamin D₃ bolus of 1,000 IU/kg body mass. PBMCs were collected immediately before (d0) and 24 h after (d1) supplementation, yielding 90 RNA-seq libraries for transcriptome-wide differential expression analysis. The Polish cohort was geographically and demographically distinct from the previously published Finnish VitDHiD cohort (25 participants, Kuopio, 63°N), enabling a direct cross-cohort assessment of which in vivo vitamin D target genes are reproducibly identified across two independent European populations. The principal goal of Paper III was therefore to (i) define a high-confidence, cross-cohort conserved set of in vivo vitamin D target genes and (ii) validate and extend the vitamin D response index concept in a second European cohort [30].

Key findings

Differential expression analysis of d1 versus d0 PBMC transcriptomes identified 758 significantly responsive vitamin D target genes ($p < 0.05$) in the VitDPAS cohort. Stratification of the 45 participants by the average absolute log₂ fold change of these target genes against the d1/d0 ratio of serum 25(OH)D₃ reproduced the high/mid/low responder structure of the vitamin D response index (17 high, 19 mid, 9 low responders), thereby replicating the index in an independent European cohort. Cross-cohort comparison of the VitDPAS gene list with the corresponding VitDHiD list (1,654 targets at $p < 0.05$) yielded 232 shared in vivo vitamin D target genes; applying the more stringent threshold of FDR < 0.05 to both cohorts narrowed this to a high-confidence conserved core of 26 genes consistently regulated in healthy individuals across both populations. Of these 26 genes, 22 are downregulated and only 4 upregulated by vitamin D₃ supplementation. The dominant

biological theme of the conserved core is modulation of inflammatory signalling, primarily through the TNF and NFκB pathways: prominent members include the NFκB inhibitors *NFKBIA* and *NFKBIZ*, the upstream inhibitor *TNFAIP3*, and additional regulators spanning immune-cell migration, transcriptional control, stress response and metabolism [30].

Significance for the thesis

Paper III is the second pillar of the response-index validation programme of this thesis: it shows that the high/mid/low responder stratification observed in the Finnish VitDHiD cohort (Chapter 3) and in the Saudi cohort (Chapter 4) is reproducible in a third, geographically and demographically distinct population, and it delivers the first cross-cohort high-confidence list of 26 conserved in vivo vitamin D target genes whose dominant biological function is modulation of inflammatory signalling. The conserved core therefore provides the most robust descriptive map of the in vivo vitamin D response in healthy human PBMCs available to date, and it is the biological backdrop against which the dynamic, mechanistic modelling of Chapter 6 is interpreted. Chapter 6 itself addresses a logically separate question (regulatory architecture from longitudinal data) using a different participant subset and a different gene set, as detailed there [30].

Author contribution

The candidate was the lead bioinformatician on Paper III: alignment and quantification of the d0 and d1 RNA-seq libraries from all 45 VitDPAS participants, differential expression analysis with edgeR, harmonisation and joint analysis with the previously published VitDHiD transcriptome dataset, derivation of the 232 shared and 26 high-confidence cross-cohort target genes, response-index classification of all 45 individuals, and generation of 12 out of 14 figures. Study organisation, recruitment, blood draws, sample processing and serum 25(OH)D₃ measurements were performed by co-authors.

Publication details

Also referred to as: VitDPAS

Emilia Gospodarska^{1*}, Ranjini Ghosh Dastidar^{1*}, Julia Jaroslawska¹, Maciej Rybiński¹, Marianna Raczyk¹, Kornelia Tokarczyk-Malesa², Jerzy Romaszko² & Carsten Carlberg¹

¹ InLife Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn ² Dept of Family Medicine, University of Warmia and Mazury, Olsztyn * **equal contribution**

***** *The full text of this publication is appended at the end of this thesis as Appended Paper III.*

Chapter 6

Dynamic System Identification Defines a Minimal Transcription Factor Control Architecture of Human Endocrine Gene Regulation

“The ability to simplify means to eliminate the unnecessary so that the necessary may speak.”

— Hans Hofmann

Background and rationale

Chapters 3–5 establish that the vitamin D response index is a real, reproducible and cross-population property of healthy individuals, and that its dominant transcriptional signature in PBMCs is the modulation of inflammatory signalling. Chapter 6 moves from this descriptive picture to a mechanistic one and asks: what regulatory architecture causes one individual to respond robustly and another minimally to the same VDR activation? To address this question, the VitDPAS cohort was extended into a longitudinal design. Forty-one of the original VitDPAS participants completed three monthly weight-adjusted vitamin D₃ bolus cycles (administered on days 0, 28 and 56), with PBMC sampling at seven time points per participant: days 0, 1, 28, 29, 56, 57 and 84. The resulting $7 \times 41 = 287$ RNA-seq libraries form the longitudinal substrate on which all subsequent modelling is built and constitute the principal experimental contribution of Paper IV beyond Paper III [28].

Model construction

Differential expression was first computed for the four contrasts d1 vs. d0, d28 vs. d0, d56 vs. d0 and d84 vs. d0. Seven TF genes were significantly regulated ($\text{FDR} < 0.05$) in all four contrasts and were therefore selected as persistently regulated primary regulators: *CSRNP1*, *ETS2*, *FOS*, *HIF1A*, *JUNB*, *KLF10* and *NR4A2*. Around this seed set, a directed regulatory network was constructed from the human DoRothEA TF–target regulon database [24], restricted to high-confidence curated interactions (1,633 regulatory edges) with activating and repressing modes of regulation. Iterative

kernel reduction, performed by collapsing low-degree nodes while preserving regulatory logic through the product of the corresponding mode-of-regulation values, yielded a minimal regulatory kernel of 11 TF genes connected by 117 directed interactions: *TP53* (tumour protein p53), *MYC*, *STAT1* (signal transducer and activator of transcription 1), *STAT3*, *FOS*, *JUN*, *HIF1A* (hypoxia inducible factor 1 subunit alpha), *EGR1* (early growth response 1), *NFKB1* (NFκB subunit 1), *ETS1* and *ETS2* (ETS proto-oncogene TFs). For each kernel gene, Hill-type ordinary differential equations describing activation and repression by the other kernel members were generated automatically; pharmacokinetic modelling of individual serum 25(OH)D₃ trajectories provided the hormonal forcing function. Differential evolution optimisation in Julia [56] fitted the 373 free ODE parameters per cycle simultaneously across all seven time points and the three supplementation cycles [28].

Key findings

The fitted ODE model reproduces individual transcriptional dynamics across all 11 kernel TFs and all seven time points, demonstrating that the reduced kernel is dynamically sufficient to explain the observed longitudinal response. A first major mechanistic finding is temporal amplification: the transcriptome-wide response is substantially larger after the third cycle than after the first despite identical doses, consistent with progressive accumulation of VDR-driven regulatory input across successive supplementations. A second finding concerns the structure of the kernel itself: the TP53 protein, the AP1 components (encoded by *FOS*, *FOSB*, *JUN* and *JUNB*), *MYC*, *EGR1* and *STAT1* form a tightly interconnected control circuit. The TP53–AP1–MYC axis acts as the primary discriminator between response classes: interaction strengths within this axis differ systematically between high and low responders, whereas aggregate expression levels of the same genes do not, indicating that responder status is encoded in regulatory wiring rather than in static expression magnitude [28]. Out-of-sample evaluation in the independent VitDHiD cohort (with median $R^2 > 0.94$) confirmed that the inferred kernel generalises beyond the VitDPAS training data without refitting.

Clinical and translational significance

The model has a direct translational implication. Because it is fitted to a single individual's longitudinal data rather than requiring cohort-level comparisons, it can in principle be applied to one patient's blood transcriptome measured across a short supplementation window. The output is a predicted response class before a full-length trial is completed. This is clinically critical: a practitioner cannot acquire a matched population cohort in routine practice, and existing group-level statistics offer no individual-level prediction. The kernel also identifies which regulatory

interactions are aberrant in a given low responder, pointing toward mechanistic targets for future intervention [28].

Author contribution

The candidate performed all bioinformatic analysis: differential expression across seven time points, TF network construction, iterative kernel reduction, ODE model formulation, and cross-cohort validation. ODE parameter optimisation was conducted in Julia under the guidance of Dr. Elisa Domínguez-Hüttinger. The author also generated all the figures. Study organisation and all wet-laboratory work was performed by co-authors.

Publication details

Also referred to as: VitDPAS Modelling

Ranjini Ghosh Dastidar¹, Emilia Gospodarska¹, Kornelia Tokarczyk-Malesa², Jerzy Romaszko², Elisa Domínguez-Hüttinger³ & Carsten Carlberg¹

¹ InLife Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn; ² Dept of Family Medicine, University of Warmia and Mazury; ³ Instituto de Investigaciones Biomédicas, UNAM, Mexico City

***** *The full text of this publication is appended at the end of this thesis as Appended Paper IV.*

Chapter 7

Combined Observations

7.1 The Problem of Interindividual Baseline Differences

Before asking what vitamin D supplementation does, it is worth asking what the immune cell transcriptome looks like across different people before any vitamin D is given. Even at baseline, before any supplement has been administered, individuals differ from one another in a profound and biologically meaningful way.

Analysis of the day 0 expression profiles of all 45 VitDPAS participants across the full transcriptome of 14,788 expressed genes reveals a median CV of 44% across individuals at baseline meaning the typical gene is expressed at levels that vary by nearly half across participants before any intervention. More striking still, 573 genes show greater than 2-fold variation across the 45 individuals at baseline. These are not obscure, lowly expressed genes but include key regulatory factors.

Among the most variable genes at baseline are the TFs that later emerge as the mechanistic kernel of the vitamin D response. *FOS* spans a 5-fold range across individuals (455 to 2,288 CPM), *JUN* a 13.5-fold range (35 to 476 CPM), *MYC* a 5.1-fold range, and *STAT1* a 4.4-fold range. Even when taken collectively, the aggregate expression of all 11 kernel TFs per individual at day 0 spans a 1.82-fold range (2,803 to 5,093 CPM total, CV = 14%). [19, 21].

Figure 5 shows the day 0 CPM expression of the eight most variable genes across all 25 VitDHiD participants and all 45 VitDPAS participants. Each subplot represents one gene and each bar is one person, sorted from highest to lowest expression. The staircase shape is the signature of strong interindividual variability. The gene *HBB* (hemoglobin subunit beta) spans from over 2,500 CPM in one person to nearly zero in another. *JCHAIN*, an immunoglobulin assembly gene (immunoglobulin J chain), shows a 10-fold difference across VitDPAS participants. These are not outliers. This is the normal range of human immune cell gene expression.

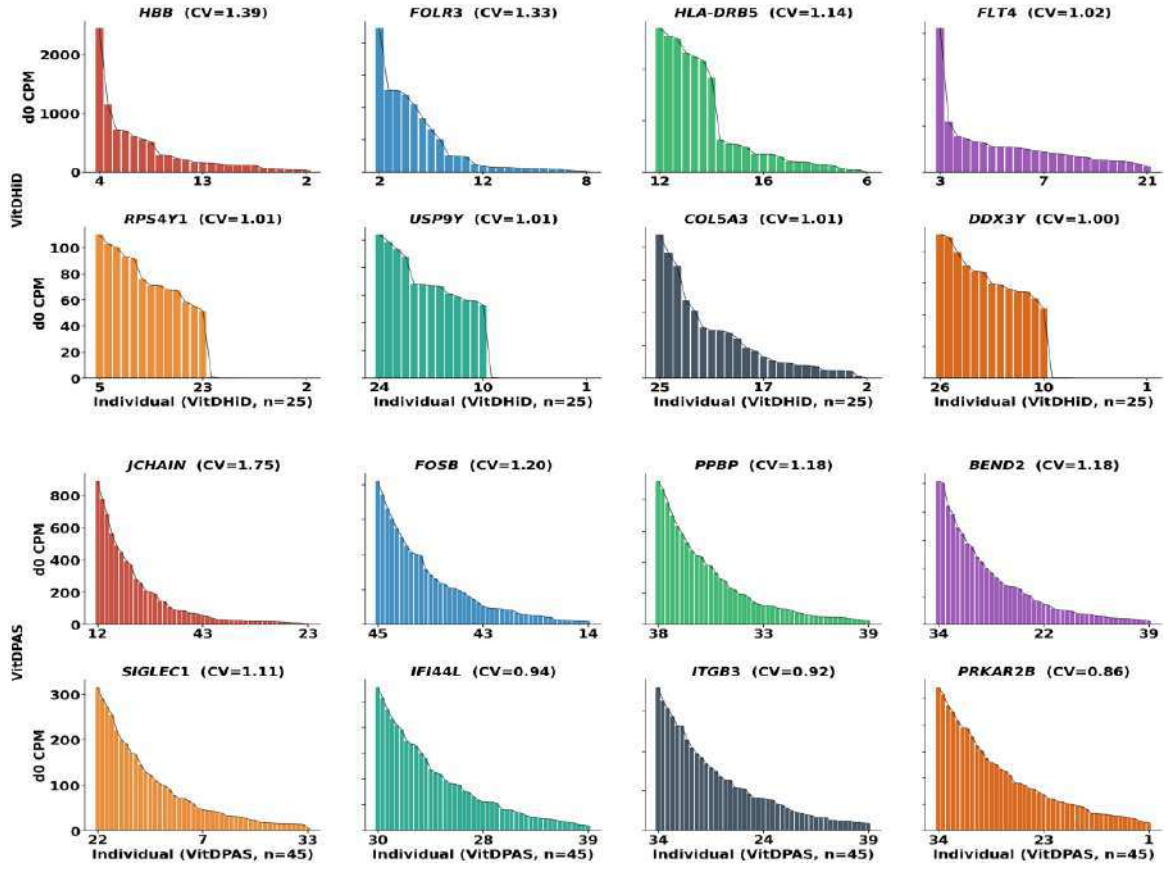


Figure 5. Top 8 most variable genes at baseline (day 0, CPM) across VitDHiD ($n=25$, top) and VitDPAS ($n=45$, bottom) individuals. Each panel sorted independently by that gene's expression level Coefficient of variation (CV) values shown in parentheses. [16, 18].

Figure 6 extends this view to the top 20 most variable genes as a heatmap. Each row is a gene, each column is an individual, and the colour runs from deep blue (low z-scored expression) to deep red (high). If the variability were driven by a single global factor, some columns would be uniformly red and others uniformly blue. Instead, the pattern is gene-specific. One individual is red for *HBB* and blue for *JUP* (junction plakoglobin, a cell–cell adhesion protein). Another shows the opposite. The baseline transcriptome is shaped differently, gene by gene, in each person.

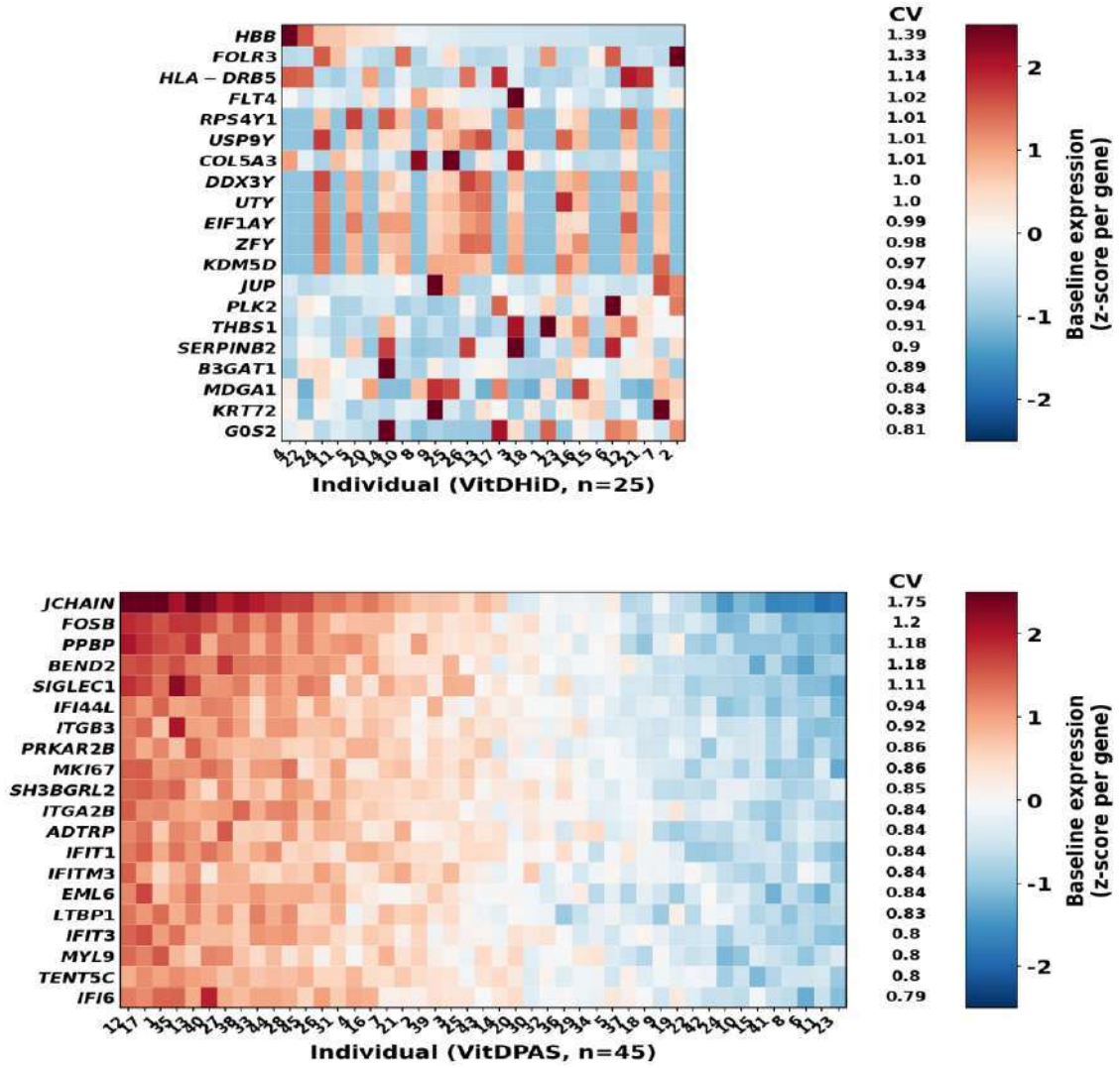


Figure 6. Heatmap of the 20 most variable genes at baseline across *VitDHiD* ($n=25$, top) and *VitDPAS* ($n=45$, bottom) individuals. Individuals sorted by the expression of the most variable gene. Z-scored expression (red = high, blue = low). CV values shown at right. [27, 30].

This gene-specific heterogeneity is not noise. It is the molecular signature of each person's accumulated immunological history like past infections, dietary patterns, prior vitamin D exposures, and genetic variants that alter how TFs bind their target DNA. The epigenetic memory hypothesis holds that this pre-existing chromatin configuration determines how strongly an individual can respond when VDR is activated by $1,25(\text{OH})_2\text{D}_3$ [27, 30].

7.2 Neither Vitamin D Levels Nor Transcriptomic Change Alone Explain the Response

A natural next step is to ask whether these baseline differences predict the response. The data show that two simple proxies both fail.

Serum 25(OH)D₃ level tells you how much vitamin D arrived in the bloodstream. It does not tell you what happened inside the cell. In VitDHiD and VitDPAS, participants whose serum levels rose equivalently showed very different numbers of differentially expressed genes and very different response scores. The molecule was metabolised in every study participant. What differed was whether the cell machinery downstream of VDR translated that signal into a coordinated transcriptional programme.

Figure 7 makes the second point clearly. It shows the distribution of differentially expressed gene (DEG) expression at day 0 (blue boxes) and day 1 (red boxes) for each individual in both cohorts, sorted from the highest to the lowest response score. If baseline expression level predicted response, the blue boxes should be systematically lower for high responders. That is not what is observed. The blue and red boxes sit at similar heights across the full range of response scores. Neither how much the target genes are expressed at baseline, nor how much they are expressed after supplementation, separates high from low responders.

What does differ is the degree of coordinated directional change. The response score measures the sum of absolute per-individual log-fold changes across all differentially expressed genes. A high responder is someone whose immune transcriptome shifts in a coherent programme across many genes simultaneously. A low responder makes smaller and less coordinated adjustments. This distinction between expression magnitude and coherence of change is the conceptual foundation of the response index.

This also explains the failure of large-scale population trials. When a trial averages changes across a mixed population of high and low responders, the coordinated upward shifts of high responders and the minimal uncoordinated changes of low responders cancel each other out. The biological signal disappears in the average.

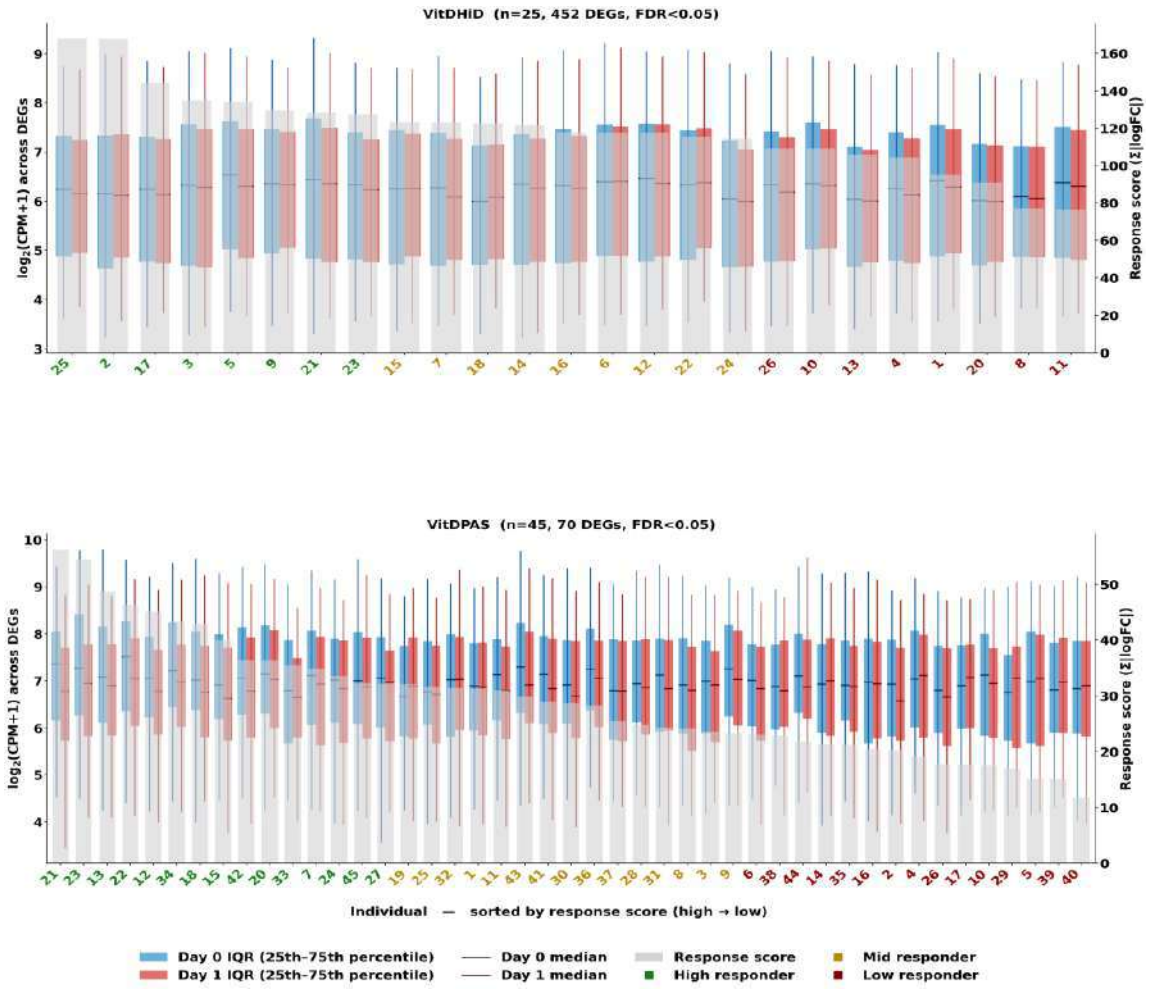


Figure 7. Expression distribution of DEGs (FDR<0.05) per individual at day 0 (blue) and day 1 (red), sorted by response score (grey bars, right axis). Individual labels coloured by response class: dark green = High, gold = Mid, maroon = Low. Panels: VitDHiD ($n=25$, 452 DEGs, top), VitDPAS ($n=45$, 70 DEGs, bottom). Neither baseline nor post-supplementation expression level separates responder classes, demonstrating that the response index measures regulatory responsiveness rather than expression magnitude. [27, 30].

7.3 Why a Mechanistic Model Is Needed

The findings described in Figures 5-7 lead to a clear conclusion: the response index cannot be reliably calculated from any single measurement. It cannot be read from a serum vitamin D level. It cannot be inferred from the absolute expression of any individual gene at baseline or after supplementation. It requires knowing something about the regulatory architecture of the cell, specifically how the TFs governing the vitamin D response are connected and how strongly those connections operate in a given individual.

This is precisely why Chapter 6 builds an ODE-based regulatory model. By fitting this model to the longitudinal transcriptomic data from VitDPAS across three supplementation

cycles, it becomes possible to identify not just which genes change but what regulatory interactions produce those changes. The TP53-AP1-MYC axis emerged as the mechanistic discriminator between high and low responders. High responders have systematically stronger interactions within this axis at baseline. This is a property of the regulatory network that is invisible to any single measurement and can only be estimated by modelling the full system dynamics.

Chapter 8

General Discussion

8.1 From Observation to Explanation: Integrated Findings Across the Four Studies

Looking across the four studies, a pattern emerges that is more than the sum of individual findings. VitDHiD established that the vitamin D transcriptional response in human immune cells is dominated by a programme centred on focal adhesion, not by the classical immune regulatory targets previously characterised in cell culture. This was a discovery that required genome-wide measurement to make and that immediately raised a follow-up question: is this finding specific to Finnish winter adults, or is it a general property of human immune biology [27]?

The Saudi cohort answered that question at the level of the response index: low responders appear at similar frequencies regardless of geography, genetic background, or baseline vitamin D status [2]. But the Saudi study, limited to cytokine and qPCR readouts, could not tell us whether the molecular programme characterised in VitDHiD generalises to other populations. That question was addressed by VitDPAS, which replicated the focal adhesion signature in an independent Polish cohort and identified the 26-gene conserved core that persists across studies, cohorts, and repeated supplementation cycles [30].

The conserved core is scientifically significant not only because it validates the VitDHiD findings but because it provides the specific set of consistently regulated genes that are robust enough to be informative for reproducible classification of the individuals into High, Mid and Low responder type. Without knowing which genes respond reliably across individuals and cohorts, the credibility of the response index comes into question and without the reliable classification a model cannot be trained efficiently. The cross-cohort comparison in Chapter 5 was therefore not simply a replication exercise: it was a necessary data reduction step that defined the biological substrate for Chapter 6.

Chapter 6 then closes the explanatory loop. Given a set of consistently regulated TFs, the ODE modelling framework asks: what is the minimal TF network whose dynamics,

constrained by the longitudinal VitDPAS time series, can reproduce the observed transcriptional trajectories? The emergence of the TP53-API-MYC regulatory axis as the discriminating interaction was not imposed: it arose from the model fitting as the relationship whose sign and strength best distinguished high from low responders across all time points and cohorts [28]. Each study was therefore logically necessary for the next. The architecture of the thesis reflects a deliberate progression from discovery to universality to conserved core to mechanism, with each stage generating the specific information the next required.

8.2 What the Transcriptional Findings Suggest About Vitamin D's Role in the Immune System

The transcriptional findings of this thesis, taken together, raise a genuine interpretive question about vitamin D's immune role that goes beyond the scope of any individual study. The combination of focal adhesion upregulation and downregulation of specific regulatory genes in the conserved core does not map neatly onto the conventional categories of pro-inflammatory or anti-inflammatory. One plausible interpretation, developed here as a working framework, is that vitamin D shifts the regulatory tone of circulating immune cells in ways that alter their readiness to respond to subsequent stimuli, without constitutively activating or suppressing inflammation. Whether this interpretation is correct requires experimental validation beyond the scope of the current studies. What the current data do establish unambiguously is that vitamin D produces a specific, reproducible, and conserved pattern of transcriptional change in human immune cells that is distinct from what *in vitro* models predicted [27, 30].

The transcriptional data from both VitDHiD and VitDPAS present an interesting interpretive challenge. The dominant *in vivo* response in human PBMCs involves upregulation of focal adhesion and cell-environment sensing genes, alongside downregulation of a conserved core that includes NFκB pathway regulators. Rather than fitting neatly into an anti-inflammatory or pro-inflammatory category, this pattern suggests a more nuanced regulatory shift whose biological consequences are context-dependent. One interpretation is that vitamin D calibrates immune cells toward a state of altered regulatory tone, with the specific functional outcome depending on what signals those cells subsequently encounter. This is a working hypothesis rather than an established conclusion,

but it provides a useful conceptual framework for interpreting the combination of findings and for generating testable predictions about when and in whom vitamin D₃ supplementation will produce clinically meaningful immune effects.

The temporal amplification documented in Chapter 6 adds a dimension that is not captured by any single-measurement study. The transcriptional response to vitamin D grows substantially across three monthly supplementation cycles, not because the hormonal stimulus changes but presumably because the chromatin landscape at VDR-occupied loci is progressively modified with each cycle. This finding has a direct and practical implication: the biological response measured at any single time point in a supplementation study represents a snapshot of a dynamic process, and the time point chosen will substantially affect what is observed. Short studies may systematically underestimate the transcriptional effects of consistent supplementation [28].

8.3 *HLA-C* and What Baseline Differences Tell Us

Among the findings of VitDHiD, the constitutively four-fold higher expression of *HLA-C* in high compared with low responders at baseline deserves particular attention because it directly demonstrates that the response index reflects a pre-existing regulatory state [27]. *HLA-C* is itself a direct VDR target gene, so its higher expression in individuals who will subsequently show the strongest vitamin D response cannot be coincidental. It suggests that the same regulatory landscape that enables strong VDR-mediated transcription is already active at baseline, before any supplementation is administered. The response index is not created by vitamin D, rather revealed by it.

The biological basis of this *HLA-C* difference is not fully established, but the chromatin data from Rybiński et al. (2026) [21] provide relevant context: chromatin accessibility at vitamin D-regulated loci differs systematically between high and low responders before supplementation, consistent with the idea that epigenetic differences underlie the differential baseline expression [51]. The mechanistic link between chromatin accessibility at VDR-bound enhancers near the MHC locus and the *HLA-C* baseline expression difference is a specific and testable prediction that future chromatin conformation studies could address.

8.4 Epigenetic Memory, Baseline Fluctuation, and the Implications for Trial Design

The striking interindividual variation in baseline gene expression documented in Chapter 7.1, covering a 44% median CV across 14,788 expressed genes, greater than 2-fold variation in 573 genes, and a 1.82-fold range in aggregate kernel TF expression, quantifies the scale of pre-existing biological heterogeneity that vitamin D trials ignore when they treat participants as exchangeable [28]. This is not a statistical nuisance: it is the biological signal that encodes each individual’s immune regulatory history, a conclusion reinforced by the chromatin-level evidence in Rybiński et al. (2026), which shows that this pre-supplementation variability has a corresponding epigenetic signature at VDR-regulated loci [51].

Combined with the temporal amplification finding, these data build a compelling case that clinical vitamin D trials need to be redesigned around two principles that current trials uniformly ignore. First, pre-trial stratification by response index phenotype is not optional but biologically necessary: including approximately one third of non-responders in a trial powered for an average effect will always dilute the true signal below detection. Second, the measurement time point matters profoundly: a trial measuring outcomes after one month is measuring a different and incomplete biological state compared to one measuring after three months of consistent bolus supplementation, when epigenetic priming has had time to deepen the transcriptional programme.

8.5 Toward Predictive Personalised Supplementation

The ODE regulatory kernel provides a specific and actionable path toward clinical prediction. The model accepts as input the baseline expression levels of the 11 kernel TFs from the baseline expression values of the 11 kernel TFs as initial conditions for the ODE model, and the fitted model reproduces the observed response trajectories of different responder subgroups. Cross-cohort validation demonstrates that this generalises to the independent VitDHiD cohort. The next step toward clinical application is a minimum viable assay: a targeted RNA quantification panel measuring the 11 kernel TFs from a standard venepuncture, interpretable through the ODE model, that classifies individuals as high, mid, or low vitamin D responders before supplementation begins.

This is not merely a theoretical possibility. The biological foundation demonstrated in this thesis, namely that the baseline transcriptional state of immune cells is mechanistically linked to the response phenotype through the fitted ODE model, is the most important prerequisite for such a clinical tool. What remains is prospective validation in a new, independent cohort designed specifically for prediction testing, followed by assay development and regulatory review. These are substantial but straightforwardly defined challenges.

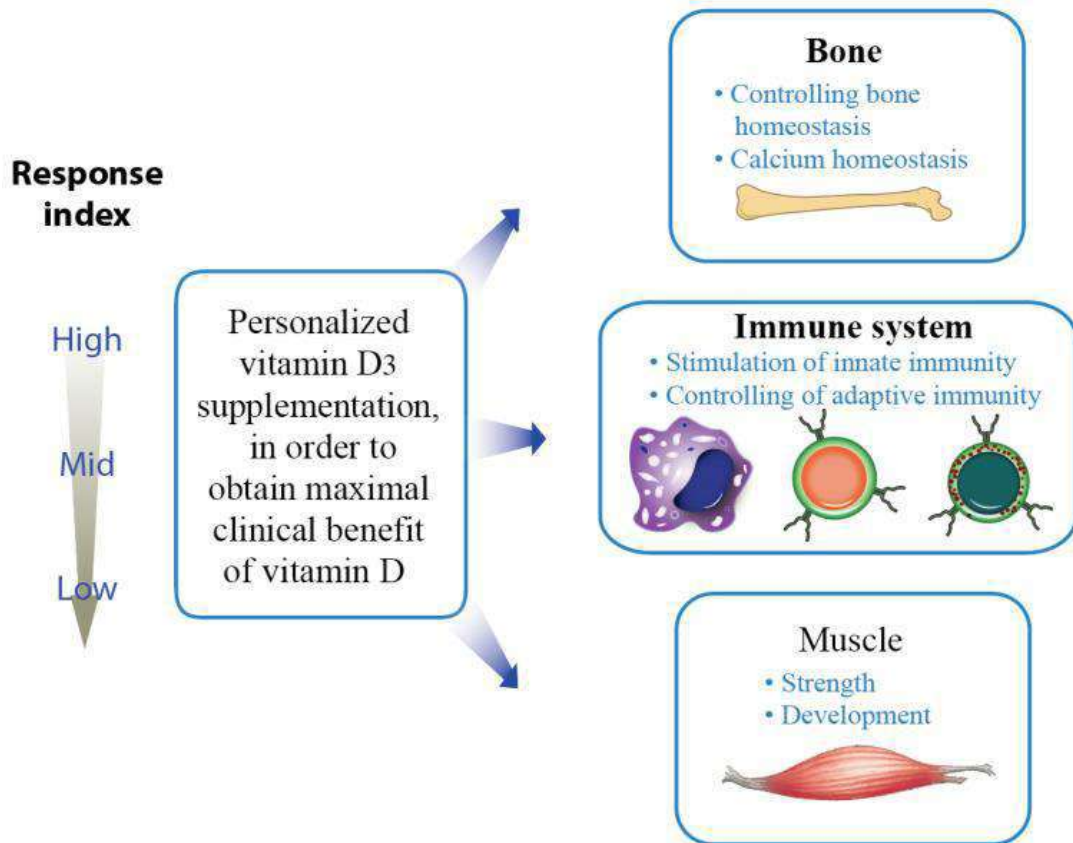


Figure 8. Overview of personalised vitamin D supplementation. Participant stratification by vitamin D response index (high, mid, low) enables targeted clinical benefit across bone homeostasis, muscle strength and development, and immune system function (innate and adaptive immunity). Taken from Carlberg C & Mycko MP (2023) *Cells* 12:2391[14].

8.6 Baseline Transcriptional Heterogeneity: Baseline Transcriptomic Architecture Before Supplementation Begins

One of the most striking and biologically informative patterns in the VitDPAS dataset is the degree of interindividual variation in gene expression at day 0, before any vitamin D

has been administered. This baseline fluctuation is not experimental noise rather it is reproducible across time points and biologically meaningful.

This pre-existing heterogeneity, which is present before any vitamin D is administered and is therefore not caused by it, is the molecular signature of each individual's accumulated immunological history: past infections that transiently remodelled chromatin in haematopoietic progenitors, dietary and metabolic exposures, prior seasonal vitamin D status, and genetic variants that alter TF binding site affinity or chromatin looping architecture. The epigenetic memory hypothesis holds that this pre-existing landscape is precisely what determines the individual's capacity to respond to VDR activation [51]: participants whose regulatory circuitry is already in a particular configuration will amplify the VDR signal into a broad transcriptional programme, while those whose circuitry is in a different state will produce a minimal response. If correct, this implies that a baseline PBMC transcriptomic profile, taken before any supplementation, contains sufficient information to predict the response phenotype. This prediction is tested and supported by the modelling in Chapter 6.

The top fluctuating genes, identified by CV across individuals at day 0, include focal adhesion genes (*ITGB3*, *ITGA2B*, *MYL9*), immune composition markers (*JCHAIN*, *SIGLEC1* [sialic acid binding immunoglobulin-type lectin 1]), an AP1 TF (*FOSB*), and interferon-induced genes (*IFI44L* [interferon induced protein 44 like], *IFIT1* [interferon induced protein with tetratricopeptide repeats 1]). Strikingly, the focal adhesion genes that are the most enriched pathway in the VitDHiD transcriptional response are also among the most variable at baseline, suggesting that the chromatin landscape at these loci is a major determinant of response capacity before any supplementation begins [27, 30].

8.7 The Evolution of the Response Index and Baseline Transcriptional Heterogeneity

The vitamin D response index, introduced in its earliest form in VitDmet as a simple gene-count classification, was progressively refined through each of the four studies in this thesis. This section discusses how those refinements deepened biological understanding and what the baseline transcriptional heterogeneity documented across studies implies for the biology of the vitamin D response.

VitDHiD advanced the concept by quantifying it transcriptome-wide using RNA-seq: each participant received a response score based on the number and magnitude of differentially expressed genes, and k-means clustering cleanly separated high, mid, and low responders. Crucially, the study also found that *HLA-C*, a MHC class I gene, was expressed at four-fold higher levels in high responders at day 0, before any supplementation. This pre-existing baseline difference suggested that the response capacity is not created by vitamin D but is already present in the individual's immune cells before the hormone arrives.

VitDPAS deepened this insight by showing that a 26-gene conserved core, including *NFKBIA*, *NFKBIZ*, and *TNFAIP3*, is consistently downregulated across the whole cohort as part of the shared vitamin D transcriptional response. Chromatin at the transcription start sites (TSSs) of these genes was found to be constitutively accessible in resting PBMCs by ATAC-seq [8], suggesting that these loci are persistently available targets rather than newly opened by VDR activation. The response index began to look less like a variable sensitivity to vitamin D and more like a property of the pre-existing regulatory landscape that determines how immune cells amplify the VDR signal.

The ODE modelling in Chapter 6 completes this conceptual arc. By fitting a dynamic mathematical model to the longitudinal VitDPAS transcriptomic time series, the study identifies the specific regulatory interactions among 11 TFs, particularly the interaction between TP53, AP1, and MYC, whose sign and strength at baseline determine whether an individual's response will be amplified or attenuated. The response index, initially a rank order based on outcome, has become an explanation of mechanism based on input.

Three arguments establish why an ODE-based mechanistic model is necessary for individual response prediction, rather than expression profiling alone. First, neither baseline nor post-supplementation expression levels of differentially expressed genes predict the response score. Response capacity is encoded in the regulatory architecture, not in transcript abundance. Second, the regulatory kernel accepts a single individual's longitudinal expression trajectory as input and does not require cohort-level data. This is clinically essential as a practitioner assessing one patient cannot acquire a matched population cohort for comparison. The ODE model works on one person's data. Third, the TP53-AP1-MYC interaction strengths distinguishing high from low responders are invisible at the level of individual gene expression but emerge from the fitted regulatory dynamics [28, 30]. These three points explain why the progression from transcriptomic

discovery to systems identification is a conceptual necessity, not merely a technical refinement.

Critically, neither the day 0 nor the day 1 expression level predicts the response score. Individuals with low absolute expression of the DEGs can have high response scores and vice versa. This demonstrates that the response index captures a regulatory property, the capacity for coherent directional change, that is invisible to any simple expression-level measurement [27, 30].

Chapter 9

Conclusions

The four studies constituting this doctoral thesis converge on a coherent, experimentally grounded, and internally consistent account of why individuals differ in their transcriptional responses to vitamin D₃ supplementation and how that variability is mechanistically encoded.

The vitamin D response index is a universal biological phenomenon. High, mid, and low responders exist in comparable proportions in Finnish and Saudi cohorts separated by thousands of kilometres of geography and profoundly different sun exposure histories. This universality rules out pharmacokinetic explanations and implies that the response phenotype is encoded in the intrinsic regulatory machinery of the individual's immune cells rather than in their relationship with the vitamin D metabolic pathway.

The *in vivo* transcriptional programme of vitamin D in human immune cells involves the coordinated upregulation of focal adhesion pathway genes (*ITGA2B*, *ITGB3*, *ITGB5*, *MYL9*, *MYLK*) alongside the downregulation of a conserved core that includes regulators of the NFκB pathway (*NFKBIA*, *NFKBIZ*, *TNFAIP3*). One interpretation of this pattern, proposed here as a working conceptual framework, is that vitamin D modulates immune cell readiness rather than simply suppressing inflammation. This remains an interpretive framing requiring further experimental testing, but it offers a coherent explanation for the combination of findings across VitDHiD and VitDPAS [27, 30].

A 26-gene high-confidence conserved core, identified through rigorous cross-cohort comparison of VitDHiD and VitDPAS datasets, constitutes the most replicable molecular phenotype of vitamin D responsiveness yet described and provides the foundation for biomarker development and minimal response assay design.

HLA-C expression is four-fold higher in high responders than low responders at baseline before any supplementation, pointing toward the MHC locus chromatin architecture as a potential genetic determinant of response capacity and inviting chromatin conformation analysis as a next experimental step.

Progressive temporal amplification of the transcriptional response across three monthly supplementation cycles provides the first direct evidence in humans of epigenetic priming

of VDR target gene loci over repeated cycles of activation, fundamentally changing how the therapeutic timeline of vitamin D supplementation should be conceptualised.

An ODE-based regulatory kernel of 11 TFs and 117 interactions, fitted to the VitDPAS longitudinal dataset and validated in the independent VitDHiD cohort, reproduces observed temporal dynamics with $R^2 > 0.94$ and identifies the TP53-AP1-MYC regulatory axis as the mechanistic discriminator between high and low response phenotypes. Baseline expression of the 11 kernel TFs at day 0 contains sufficient predictive information to forecast response trajectories before supplementation begins.

Finally, this thesis demonstrates that intervention-driven dynamic systems identification from longitudinal transcriptomic data is a productive and generalisable strategy for resolving interindividual variability in nutrigenomic response, with broad applicability beyond vitamin D to other dietary compounds and pharmacological agents whose molecular effects unfold across extended time scales in human immune cells.

Chapter 10

Implications and Recommendations

10.1 Limitations

This thesis has several limitations that should be acknowledged alongside its contributions.

The cohort sizes are relatively small. VitDHiD comprised 25 participants and VitDPAS 45. While standard for deep transcriptomic studies, and while cross-cohort replication across four studies strengthens confidence in the findings, these sizes limit the detection of low-frequency responder subtypes. Larger cohorts will be needed before clinical translation.

The studies measure PBMCs only. Vitamin D responses occur across many tissue types including gut, bone, and adipose tissue, which are inaccessible in healthy participants. The transcriptomic response characterised here may not fully represent responses in the tissues where vitamin D has its main clinical effects. Furthermore, PBMCs are a heterogeneous mixture of monocytes, T cells, B cells, and NK cells whose proportions vary between individuals. This cellular composition variability is an uncontrolled confounder that may contribute to observed interindividual transcriptional differences.

The VitDHiD and Saudi cohort studies used a single bolus dose [2, 16]. The three-cycle VitDPAS design partially addresses this by showing temporal stability of response classification, but the 28-day window does not capture slower epigenetic changes operating over months.

The ODE model in Chapter 6 is a simplification. The regulatory kernel of 11 TFs and 117 interactions omits many co-regulators and feedback loops. Model parameters should be interpreted as effective interaction strengths within the defined kernel, not absolute biochemical constants. A further important caveat concerns causal inference: although ODE fitting identifies regulatory interactions that are mathematically necessary to reproduce the observed trajectories, this does not establish causality in the experimental sense; direct perturbation of the identified TF interactions would be required to confirm their causal roles.

Finally, all cohorts are drawn from European or Saudi Arabian populations. Whether the TP53-AP1-MYC axis operates similarly as a discriminator in other populations remains to be determined.

10.2 Public Health: Moving Beyond One-Size-Fits-All Guidelines

The most immediate practical implication of this thesis is that current vitamin D supplementation guidelines, which recommend identical doses for all adults within broad age and weight categories, are biologically inadequate. A reproducible low-responder phenotype affecting approximately one quarter to one third of the population means that a substantial fraction of supplemented individuals receive vitamin D without meaningful transcriptional benefit to their immune regulatory programme, regardless of how much their serum 25(OH)D₃ rises. Population-wide uniform dosing optimises a statistical average that may be biologically meaningful for almost no specific individual [\[2, 16, 30\]](#).

The weight-adjusted dosing protocol of VitDPAS (1,000 IU per kilogram of body weight per month) represents a first step toward pharmacokinetically rational personalisation.[\[30\]](#) But the deeper personalisation this thesis argues is biologically necessary involves pre-supplementation classification by response index phenotype. High responders may benefit from moderate, well-timed supplementation that maintains VDR activation without inducing excess CYP24A1-driven catabolism. Low responders may require fundamentally different approaches, potentially including higher or more frequent doses, alternative delivery routes, or combination strategies targeting the TP53-AP1-MYC regulatory bottleneck rather than simply increasing VDR ligand availability [\[28\]](#).

10.3 Clinical Trial Design: What Needs to Change

The failure of large-scale vitamin D trials (VITAL, D-HEALTH, ViDA) to demonstrate significant average effects is explained by this thesis in biological terms: the trials included large fractions of non-responders, measured outcomes at time points before epigenetic priming had fully established the transcriptional benefit, and were powered for average effects that the biological heterogeneity of their populations made statistically inaccessible [\[44, 45, 53\]](#).

Future vitamin D trials need three design elements that current trials uniformly lack. First, pre-trial stratification of participants by vitamin D response index phenotype, ideally using a baseline PBMC transcriptomic measurement, is needed to separate the analysable subgroups and enable subgroup-specific effect estimation. Second, a minimum intervention duration of twelve months with monthly bolus dosing, rather than daily dosing over shorter periods, is needed to allow epigenetic priming to establish the full transcriptional programme before outcomes are assessed. Third, the primary analysis should distinguish effects in high, mid, and low responders rather than pooling all participants into a single average.

10.4 Translational Medicine: The Path to a Predictive Assay

The ODE regulatory kernel, validated across two independent cohorts, provides the computational core of a potential clinical predictive tool [28]. A minimum viable assay measuring the 11 kernel TFs from a baseline venepuncture could, through the ODE model, classify individuals as high, mid, or low responders before any vitamin D is administered [28].

Realising this clinical potential requires prospective validation in a new independent cohort designed specifically for prediction testing, followed by assay miniaturisation, clinical laboratory compatibility testing, health economic evaluation, and regulatory engagement. These are substantial challenges, but the biological foundation established by this thesis, demonstrating that pre-supplementation transcriptomics contains mechanistically grounded predictive information, is the most critical intellectual prerequisite for the entire translation pathway [27, 28, 30].

10.5 Future Research Directions

HLA-C's constitutively higher expression in high responders before supplementation invites genetic analysis of the MHC locus to test whether specific chromatin looping configurations or VDRE-containing enhancer variants underlie both the higher basal *HLA-C* expression and the greater VDR responsiveness. This requires chromatin conformation data from sorted immune cell populations stratified by response phenotype.

The ODE model should be validated in cohorts with different supplementation regimens, including daily dosing and paediatric or elderly populations, and in genetically diverse non-European populations to determine whether the TP53-AP1-MYC axis discriminates response phenotypes universally or specifically in European ancestry groups.

The epigenetic priming hypothesis, which predicts that repeated VDR activation progressively opens chromatin at target gene loci, should be tested directly using longitudinal ATAC-seq across the three VitDPAS supplementation cycles. The prediction is specific: chromatin accessibility at VDR-occupied enhancers should increase progressively in high responders across cycles but remain static in low responders.

Finally, the vitamin D response index framework should be deployed in prospective cohort studies with embedded transcriptomic measurements to test whether baseline PBMC expression predicts clinically meaningful outcomes including respiratory infection susceptibility, vaccine response magnitude, and autoimmune disease incidence.

*“For every answer that opens ten new questions, may we always
have the courage to keep asking.”*

— Anonymous

Chapter 11

List of Figures

General Introduction

Figure 1. The vitamin D metabolic pathway. UVB-driven cutaneous synthesis, hepatic CYP2R1 (the main hepatic 25-hydroxylase) hydroxylation, and renal CYP27B1 activation; CYP24A1 feedback inactivation.

Figure 2. VDR-mediated transcription mechanism. $1,25(\text{OH})_2\text{D}_3$ -induced VDR-RXR heterodimerisation, VDRE binding, and co-activator recruitment.

Figure 3. Vitamin D and immune cell activation. Autocrine, paracrine, and endocrine roles of $25(\text{OH})\text{D}_3$ in monocytes/macrophages, epithelial cells, neutrophils, dendritic cells, and T helper cells.

Figure 4. Personalised vitamin D supplementation framework. Participant stratification by vitamin D response index enabling targeted clinical benefit for bone, muscle, and immune health.

Figure 5. Top 8 most variable genes at baseline (day 0, CPM) across VitDHiD (n=25, top) and VitDPAS (n=45, bottom) individuals. Each panel sorted independently by that gene's expression level, illustrating the dramatic interindividual differences in baseline gene expression. CV values shown in parentheses.

Figure 6. Heatmap of the 20 most variable genes at baseline across VitDHiD (n=25, top) and VitDPAS (n=45, bottom) individuals. Individuals sorted by the expression of the most variable gene. Z-scored expression (red = high, blue = low). CV values shown at right.

Figure 7. Expression distribution of differentially expressed genes (FDR<0.05) per individual at day 0 (blue) and day 1 (red), sorted by response score (grey bars, right axis). Individual labels coloured by response class: dark green = High, gold = Mid, maroon = Low. Panels: VitDHiD (n=25, 452 DEGs, top), VitDPAS (n=45, 70 DEGs, bottom). [27, 30]

Figure 8. Overview of personalised vitamin D supplementation clinical impact. Stratification of participants by response index (high, mid, low) with downstream benefits for bone, muscle, and immune function.

Chapter 3 – VitDHiD (Scientific Reports 2024)

Figure 1.1. Venn diagram: overlap between 452 *in vivo* and 142 *in vitro* VitDHiD vitamin D target genes.

Figure 1.2. Schematic: location and function of 16 dual-confirmed vitamin D target genes across nuclear receptor signalling, focal adhesion, and immune surveillance nodes.

Figure 1.3. Focal adhesion and PI3K-AKT-TOR pathways with VitDHiD target genes highlighted.

Figure 1.4. Segregation of 25 VitDHiD participants into high, mid, and low vitamin D responders by k-means clustering; HLA-C baseline expression shown.

Figure 1.5. Differentially expressed genes in high versus low responders at day 1 (VitDHiD).

Supplementary Figures S1–S5. MDS plots; elbow method clustering; MA plot; VDR-binding ChIP-seq enhancer maps; alternative classification subsets.

Chapter 4 – Saudi Cohort (Saudi Pharmaceutical Journal 2024)

Figure 2.1. Serum cytokine changes (IL6, IL8, TNF) in 100 Saudi participants after vitamin D₃ bolus.

Figure 2.2. Ranking of 100 Saudi participants by cumulative response score; k-means clustering.

Figure 2.3. qPCR-based response scoring of 37 Saudi participants (7 target genes).

Figure 2.4. Concordance and discordance between cytokine-based and gene expression-based response classification.

Chapter 5 – VitDPAS (Scientific Reports 2025)

Figure 3.1. MDS and volcano plots for VitDPAS d1 vs d0 transcriptomic analysis.

Figure 3.2. Classification of 45 VitDPAS participants into high (n=17), mid (n=19), and low (n=9) responders.

Figure 3.3. Venn diagram: overlap between 758 VitDPAS and 1,654 VitDHiD target genes.

Figure 3.4. Venn diagram: overlap between 70 high-confidence VitDPAS targets and 452 VitDHiD targets; identification of the 26-gene conserved core.

Supplementary Figures S1–S8. Study design; MA/volcano plots; correlation analyses; clustering; expression heatmap; reproducibility; ATAC-seq; STRING network.

Chapter 6 – Manuscript (submitted 2026)

Figure 4.1. VitDPAS longitudinal study design: three bolus cycles, seven time points.

Figure 4.2. Temporal amplification: Volcano plots for all six pairwise comparisons showing progressive expansion of the differentially expressed gene set.

Figure 4.3. Classification of 41 VitDPAS participants into stable high (n=8), mid (n=8), and low (n=6) responders across three cycles.

Figure 4.4. Network reduction procedure: from initial DoRothEA (Discriminant Regulon Expression Analysis) network to 11-TF kernel across four stages.

Figure 4.5. ODE model fits: temporal TF expression dynamics for all 11 kernel genes across seven time points.

Figure 4.6. TP53-AP1-MYC regulatory interactions distinguishing high and low responders.

Supplementary Figures S1–S8. MDS; pharmacokinetic fits; Venn diagrams; kernel reduction; ODE equations; sensitivity analysis; subgroup fits; cross-cohort validation.

Chapter 12

List of Tables

Chapter 3 – VitDHiD

Table 1.1. Participant demographics: age, sex, BMI, serum 25(OH)D₃ at d0 and d1.

Table 1.2. Gene enrichment analysis of the 91 most responsive *in vivo* vitamin D target genes (EnrichR/Wikipathways).

Supplementary Table S1.1. VitDHiD datasets. Individual number, type of intervention, treatment, GEO accession and number of uniquely aligned reads are indicated.

Supplementary Table S1.2. The vitamin D₃-triggered transcriptome of PBMCs *in vivo* of 9374 expressed genes (average CPM > 10).

Supplementary Table S1.3. The 1,25(OH)₂D₃-triggered transcriptome of PBMCs *in vitro* of 9628 expressed genes.

Supplementary Table S1.4. Difference of gene expression profile of high and low vitamin D responders at d0.

Supplementary Table S1.5. Difference of gene expression profile of high and low vitamin D responders at d1.

Chapter 4 – Saudi Cohort

Table 2.1. qPCR oligonucleotide primers for the seven vitamin D target genes measured.

Supplementary Table S1. Participant demographics and complete cytokine/25(OH)D₃ data for all 100 Saudi participants.

Chapter 5 – VitDPAS

Table 3.1. Participant demographics: age, sex, BMI, weight, administered dose, and serum 25(OH)D₃.

Supplementary Table S1.1. VitDHiD datasets. Individual number, type of intervention, treatment, GEO accession and number of uniquely aligned reads are indicated.

Supplementary Table S1.2. The vitamin D₃-triggered transcriptome of PBMCs *in vivo* of 8397 expressed genes (average CPM > 10).

Chapter 6 – Manuscript, under revision

Supplementary Table S1. Longitudinal participant demographics and 25(OH)D₃ at all seven time points.

Supplementary Table S2. RNA-seq quality metrics for all 287 sequencing libraries.

Supplementary Table S3. Complete differentially expressed gene lists across all six pairwise comparisons with logFC, CPM, p-value, FDR.

Supplementary Table S4. ODE model parameter values for full cohort and three responder subgroup models.

Chapter 13

List of Publicly Available Data

Raw RNA-seq data are deposited at the Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo). Bioinformatics pipelines and model code are deposited at GitHub with DOI-registered Zenodo archives.

Chapter 3 – VitDHiD

GEO GSE260981: 50 *in vivo* PBMC RNA-seq samples from 25 VitDHiD participants (d0 and d1).

GEO GSE261699: 28 *in vitro* PBMC RNA-seq samples from 14 VitDHiD participants.

Chapter 5 – VitDPAS

GEO GSE278885: VitDPAS PBMC RNA-seq samples at days 0 and 1 from 45 participants.

Chapter 6 – Manuscript

GEO GSE278885: VitDPAS RNA-seq days 0 and 1 (shared with Chapter 5).

GEO GSE319989: VitDPAS RNA-seq longitudinal samples at days 28, 29, 56, 57, 84 from all participants.

GitHub: [ranjinighoshdastidar/VitDPAS-DGE_RNAseq](https://github.com/ranjinighoshdastidar/VitDPAS-DGE_RNAseq). DOI: 10.5281/zenodo.18929051

GitHub: [ranjinighoshdastidar/VitDPAS-Classification](https://github.com/ranjinighoshdastidar/VitDPAS-Classification). DOI: 10.5281/zenodo.18929070

GitHub: [ranjinighoshdastidar/VitDPAS-Modeling](https://github.com/ranjinighoshdastidar/VitDPAS-Modeling). DOI: 10.5281/zenodo.18929074

Chapter 14

References

References within each appended paper appear in their original published form. The following list covers all references cited in the thesis narrative, numbered as they appear in the text. An asterisk (*) denotes equal contribution.

1. Adams JS & Hewison M (2008) Unexpected actions of vitamin D: new perspectives on the regulation of innate and adaptive immunity. *Nat Clin Pract Endocrinol Metab* 4:80–90. DOI: 10.1038/ncpendmet0716
2. AlGhamdi SA*, Ghosh Dastidar R*, Rybiński M, Alsufiani HM, Khoja SO, Enaibsi NN, Saif SF & Carlberg C (2024) Evaluation of the vitamin D response index in a Saudi cohort. *Saudi Pharm J* 32:102137. DOI: 10.1016/j.jsps.2024.102137
3. Baeke F, Takiishi T, Korf H, Gysemans C & Mathieu C (2010) Vitamin D: modulator of the immune system. *Curr Opin Pharmacol* 10:482–496. DOI: 10.1016/j.coph.2010.04.001
4. Barbarawi M, Kheiri B, Zayed Y, Barbarawi O, Dhillon H, Swaid B, Yelangi A, Sundus S, Bachuwa G, Alkotob ML & Manson JE (2019) Vitamin D supplementation and cardiovascular disease risks in more than 83,000 individuals in 21 randomized clinical trials: a meta-analysis. *JAMA Cardiol* 4:765–776. DOI: 10.1001/jamacardio.2019.1870
5. Bikle DD (2014) Vitamin D metabolism, mechanism of action, and clinical applications. *Chem Biol* 21:319–329. DOI: 10.1016/j.chembiol.2013.12.016
6. Bishop EL, Ismailova A, Dimeloe S, Hewison M & White JH (2020) Vitamin D and immune regulation: antibacterial, antiviral, anti-inflammatory. *JBM R Plus* 5:e10405. DOI: 10.1002/jbm4.10405
7. Bouillon R, Manousaki D, Rosen C, Trajanoska K, Rivadeneira F & Richards JB (2022) The health effects of vitamin D supplementation: evidence from human studies. *Nat Rev Endocrinol* 18:96–110. DOI: 10.1038/s41574-021-00593-z
8. Buenrostro JD, Wu B, Chang HY & Greenleaf WJ (2015) ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* 109:21.29.1–9. DOI: 10.1002/0471142727.mb2129s109

9. Carlberg C (2014) Genome-wide (over)view on the actions of vitamin D. *Front Physiol* 5:167. DOI: 10.3389/fphys.2014.00167
10. Carlberg C (2018) Vitamin D genomics: from in vitro to in vivo. *Front Endocrinol* 9:250. DOI: 10.3389/fendo.2018.00250
11. Carlberg C (2019) Nutrigenomics of vitamin D. *Nutrients* 11:676. DOI: 10.3390/nu11030676
12. Carlberg C & Haq A (2018) The concept of the personal vitamin D response index. *J Steroid Biochem Mol Biol* 175:12–17. DOI: 10.1016/j.jsbmb.2016.12.011
13. Carlberg C & Molnár F (2025) Nutrition and immune responses. In: *Nutrigenomics*. Springer, Cham. DOI: 10.1007/978-3-031-85881-9_7
14. Carlberg C & Mycko MP (2023) Linking mechanisms of vitamin D signaling with multiple sclerosis. *Cells* 12:2391. DOI: 10.3390/cells12192391
15. Carlberg C & Velleuer E (2022) Vitamin D and the risk for cancer: a molecular analysis. *Biochem Pharmacol* 196:114735. DOI: 10.1016/j.bcp.2021.114735
16. Carlberg C & Velleuer E (2024) Vitamin D and aging: central role of immunocompetence. *Nutrients* 16:398. DOI: 10.3390/nu16030398
17. Carlberg C, Raczyk M & Zawrotna N (2023) Vitamin D: a master example of nutrigenomics. *Redox Biol* 62:102695. DOI: 10.1016/j.redox.2023.102695
18. Carlberg C, Seuter S, de Mello VD, Schwab U, Voutilainen S, Pulkki K, Nurmi T, Virtanen J, Tuomainen TP & Uusitupa M (2013) Primary vitamin D target genes allow a categorization of possible benefits of vitamin D3 supplementation. *PLoS ONE* 8:e71042. DOI: 10.1371/journal.pone.0071042
19. Carlberg C, Seuter S, Nurmi T, Tuomainen TP, Virtanen JK & Neme A (2018) In vivo response of the human epigenome to vitamin D: a proof-of-principle study. *J Steroid Biochem Mol Biol* 180:142–148. DOI: 10.1016/j.jsbmb.2018.01.002
20. Chen S, Sims GP, Chen XX, Gu YY, Chen S & Lipsky PE (2007) Modulatory effects of 1,25-dihydroxyvitamin D3 on human B cell differentiation. *J Immunol* 179:1634–1647. DOI: 10.4049/jimmunol.179.3.1634

21. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21. DOI: 10.1093/bioinformatics/bts635
22. Domínguez-Hüttinger E, Christodoulides P, Miyauchi K, Irvine AD, Okada-Hatakeyama M, Kubo M & Tanaka RJ (2017) Mathematical modeling of atopic dermatitis reveals “double-switch” mechanisms underlying 4 common disease phenotypes. *J Allergy Clin Immunol* 139:1861–1872.e7. DOI: 10.1016/j.jaci.2016.10.026
23. Domínguez-Hüttinger E, Ono M, Barahona M & Tanaka RJ (2012) Mathematical model of the development of atopic dermatitis. *FEBS J* 279:523. DOI: 10.1111/j.1742-4658.2012.08477.x
24. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D & Saez-Rodriguez J (2019) Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res* 29:1363–1375. DOI: 10.1101/gr.240663.118
25. Geiger B, Spatz JP & Bershadsky AD (2009) Environmental sensing through focal adhesions. *Nat Rev Mol Cell Biol* 10:21–33. DOI: 10.1038/nrm2593
26. Geiger B & Yamada KM (2011) Molecular architecture and function of matrix adhesions. *Cold Spring Harb Perspect Biol* 3:a005033. DOI: 10.1101/cshperspect.a005033
27. Ghosh Dastidar R, Jaroslawska J, Malinen M, Tuomainen TP, Virtanen JK, Bendik I & Carlberg C (2024) In vivo vitamin D targets reveal the upregulation of focal adhesion-related genes in primary immune cells of healthy individuals. *Sci Rep* 14:17552. DOI: 10.1038/s41598-024-68741-9
28. Ghosh Dastidar R, Gospodarska E, Tokarczyk-Malesa K, Romaszko J, Domínguez-Hüttinger E & Carlberg C (2025) Dynamic system identification defines a minimal transcription factor control architecture of human endocrine gene regulation. Manuscript. Under review.
29. Gospodarska E, Ghosh Dastidar R & Carlberg C (2023) Intervention approaches in studying the response to vitamin D3 supplementation. *Nutrients* 15:3382. DOI: 10.3390/nu15153382
30. Gospodarska E*, Ghosh Dastidar R*, Jaroslawska J, Rybiński M, Raczyk M, Tokarczyk-Malesa K, Romaszko J & Carlberg C (2025) Transcriptomic profiling of

immune modulation induced by vitamin D3 in the VitDPAS and VitDHiD cohort studies. *Sci Rep* 15:17334. DOI: 10.1038/s41598-025-02495-w

31. Hanel A & Carlberg C (2020) Vitamin D and evolution: pharmacologic implications. *Biochem Pharmacol* 173:113595. DOI: 10.1016/j.bcp.2019.07.024

32. Hanel A, Neme A, Malinen M, Hämäläinen E, Malmberg HR, Etheve S, Tuomainen TP, Virtanen JK, Bendik I & Carlberg C (2020) Common and personal target genes of the micronutrient vitamin D in primary immune cells from human peripheral blood. *Sci Rep* 10:21051. DOI: 10.1038/s41598-020-78288-0

33. Hanel A, Veldhuizen C & Carlberg C (2022) Gene-regulatory potential of 25-hydroxyvitamin D₃ and D₂. *Front Nutr* 9:910601. DOI: 10.3389/fnut.2022.910601

34. Hewison M, Freeman L, Hughes SV, Evans KN, Bland R, Eliopoulos AG, Kilby MD, Moss PA & Chakraverty R (2003) Differential regulation of vitamin D receptor and its ligand in human monocyte-derived dendritic cells. *J Immunol* 170:5382–5390. DOI: 10.4049/jimmunol.170.11.5382

35. Holick MF (2007) Vitamin D deficiency. *N Engl J Med* 357:266–281. DOI: 10.1056/NEJMr070553

36. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, Murad MH & Weaver CM (2011) Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab* 96:1911–1930. DOI: 10.1210/jc.2011-0385

37. Hosseini-Nezhad A & Holick MF (2013) Vitamin D for health: a global perspective. *Mayo Clin Proc* 88:720–755. DOI: 10.1016/j.mayocp.2013.05.011

38. Hyppönen E & Power C (2007) Hypovitaminosis D in British adults at age 45 y: nationwide cohort study of dietary and lifestyle predictors. *Am J Clin Nutr* 85:860–868. DOI: 10.1093/ajcn/85.3.860

39. Jolliffe DA, Greiller CL, Mein CA, Hoti M, Bakhsoliani E, Telcian AG, Simpson A, Barnes NC, Curtin JA, Custovic A, Johnston SL, Griffiths CJ, Walton RT & Martineau AR (2018) Vitamin D receptor genotype influences risk of upper respiratory infection. *Br J Nutr* 120:891–900. DOI: 10.1017/S000711451800209X

40. Kim JR, Kim J, Kwon YK, Lee HY, Heslop-Harrison P & Cho KH (2011) Reduction of complex signaling networks to a representative kernel. *Sci Signal* 4:ra35. DOI: 10.1126/scisignal.2001390
41. Laird E, Ward M, McSorley E, Strain JJ & Wallace J (2010) Vitamin D and bone health: potential mechanisms. *Nutrients* 2:693–724. DOI: 10.3390/nu2070693
42. Liu PT, Stenger S, Li H, Wenzel L, Tan BH, Krutzik SR, Ochoa MT, Schaubert J, Wu K, Meinken C, Kamen DL, Wagner M, Bals R, Steinmeyer A, Zügel U, Gallo RL, Eisenberg D, Hewison M, Hollis BW, Adams JS, Bloom BR & Modlin RL (2006) Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response. *Science* 311:1770–1773. DOI: 10.1126/science.1123933
43. Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. DOI: 10.1186/s13059-014-0550-8
44. Manson JE, Cook NR, Lee IM, Christen W, Bassuk SS, Mora S, Gibson H, Gordon D, Copeland T, D’Agostino D, Friedenberg G, Ridge C, Bubes V, Giovannucci EL, Willett WC & Buring JE (2019) Vitamin D supplements and prevention of cancer and cardiovascular disease. *N Engl J Med* 380:33–44. DOI: 10.1056/NEJMoA1809944
45. Neale RE, Baxter C, Romero BD, McLeod DSA, English DR, Armstrong BK, Ebeling PR, Hartel G, Kimlin MG, O’Connell R, van der Pols JC, Venn AJ, Webb PM, Whiteman DC & Waterhouse M (2022) The D-Health Trial: a randomised controlled trial of the effect of vitamin D on mortality. *Lancet Diabetes Endocrinol* 10:120–128. DOI: 10.1016/S2213-8587(21)00345-4
46. Neme A, Seuter S & Carlberg C (2016) Vitamin D-dependent chromatin association of CTCF in human monocytes. *Biochim Biophys Acta Gene Regul Mech* 1859:1380–1388. DOI: 10.1016/j.bbagr.2016.08.008
47. Pike JW, Meyer MB & Bishop KA (2012) Regulation of target gene expression by the vitamin D receptor: an update on mechanisms. *Rev Endocr Metab Disord* 13:45–55. DOI: 10.1007/s11154-011-9198-9
48. Pike JW & Meyer MB (2010) The vitamin D receptor: new paradigms for the regulation of gene expression by 1,25-dihydroxyvitamin D₃. *Endocrinol Metab Clin North Am* 39:255–269. DOI: 10.1016/j.ecl.2010.02.007

49. Ramagopalan SV, Heger A, Berlanga AJ, Maugeri NJ, Lincoln MR, Burrell A, Handunnetthi L, Handel AE, Disanto G, Orton SM, Watson CT, Morahan JM, Giovannoni G, Ponting CP, Ebers GC & Knight JC (2010) A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. *Genome Res* 20:1352–1360. DOI: 10.1101/gr.107920.110
50. Robinson MD, McCarthy DJ & Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. DOI: 10.1093/bioinformatics/btp616
51. Rybiński M, Ghosh Dastidar R, Zawrotna N, Gospodarska E & Carlberg C (2025) Chromatin and transcriptional dynamics underlying the immune-modulatory effects of vitamin D₃ in vivo. *Sci Rep* 16:2997. DOI: 10.1038/s41598-025-32831-z
52. Sawatsubashi S, Nishimura K, Mori J, Kouzmenko A & Kato S (2019) The function of the vitamin D receptor and a possible role of enhancer RNA in epigenomic regulation of target genes: implications for bone metabolism. *J Bone Metab* 26:3–12. DOI: 10.11005/jbm.2019.26.1.3
53. Scragg R, Stewart AW, Waayer D, Lawes CMM, Toop L, Sluyter J, Murphy J, Khaw KT & Camargo CA Jr (2017) Effect of monthly high-dose vitamin D supplementation on cardiovascular disease in the Vitamin D Assessment Study: a randomized clinical trial. *JAMA Cardiol* 2:608–616. DOI: 10.1001/jamacardio.2017.0175
54. Seuter S, Neme A & Carlberg C (2017) Epigenomic PU.1-VDR crosstalk modulates vitamin D signaling. *Biochim Biophys Acta Gene Regul Mech* 1860:405–415. DOI: 10.1016/j.bbagr.2017.02.005
55. Seuter S, Virtanen JK, Nurmi T, Pihlajämäki J, Mursu J, Voutilainen S, Tuomainen TP, Neme A & Carlberg C (2017) Molecular evaluation of vitamin D responsiveness of healthy young adults. *J Steroid Biochem Mol Biol* 174:314–321. DOI: 10.1016/j.jsbmb.2016.06.003
56. Storn R & Price K (1997) Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces. *J Glob Optim* 11:341–359. DOI: 10.1023/A:1008202821328

57. Strimmer K (2008) fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics* 24:1461–1462. DOI: 10.1093/bioinformatics/btn209
58. Sutherland JP, Zhou A & Hyppönen E (2022) Vitamin D deficiency increases mortality risk in the UK Biobank: a nonlinear Mendelian randomization study. *Ann Intern Med* 175:1552–1559. DOI: 10.7326/M21-3324
59. Wacker M & Holick MF (2013) Sunlight and vitamin D: a global perspective for health. *Dermatoendocrinol* 5:51–108. DOI: 10.4161/derm.24494
60. Wang Z, Gerstein M & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63. DOI: 10.1038/nrg2484
61. Yang R, Han C, Xie Y, Qiu S, Zhang S, He J, Wang Z, Zhang Z, Liu H, Chen L & Li B (2025) The pathogenic roles of local vitamin D metabolism defect in valve inflammation and calcification. *Adv Sci* 12:e2501250. DOI: 10.1002/advs.202501250

Chapter 15

Acknowledgements

This doctoral thesis was made possible by the exceptional support, expertise, and generosity of many people across four countries.

My deepest gratitude goes to not only my supervisor but also a father-figure, Prof. Carsten Carlberg, whose scientific vision, intellectual generosity, and unfailing encouragement shaped every aspect of this work. He gave me both the freedom to explore and the rigour to ensure that exploration was worthwhile. He motivated me professionally as well as personally and reeled me back when my overthinking took over. Words fall short for the sense of pride I take in being his student.

I am deeply grateful to the entire Nutrigenomics Laboratory team at the InLife Institute of Animal Reproduction and Food Research: Dr. Emilia Gospodarska (equal-contribution co-author on Chapters 5 and 6, whose biological expertise and collaborative generosity were irreplaceable); Dr. Julia Jarosławska-Miszkiewicz, Dr. Marianna Raczyk, Maciej Rybiński, Mariusz Jankowski, Tanya Tripathi, Arleta Janysz, Priman Alfred Fau, and Piotr Lekowski.

I owe special gratitude to Dr. Elisa Domínguez-Hüttinger (Instituto de Investigaciones Biomédicas, UNAM, Mexico City), whose expertise in ODE-based immune systems modelling and whose intellectual generosity in sharing it were indispensable to the most technically demanding part of this thesis.

The VitDHiD study was made possible through Dr. Marjo Malinen (South-Eastern Finland University of Applied Sciences), Prof. Tomi-Pekka Tuomainen, Prof. Jyrki K. Virtanen (University of Eastern Finland, Kuopio), and Dr. Igor Bendik (DSM Nutritional Products, Basel).

For the Saudi cohort study, I thank Dr. Shareefa A. AlGhamdi, Dr. Hadeil M. Alsufiani, Dr. Sawsan O. Khoja, Nusaibah N. Enaibsi, and Safa F. Saif at King Abdulaziz University, Jeddah.

The VitDPAS clinical study depended on Dr. Kornelia Tokarczyk-Malesa and Dr. Jerzy Romaszko at the University of Warmia and Mazury, whose clinical leadership and ethical integrity enabled participant recruitment across two years of sample collection.

I gratefully acknowledge the Gene Core Facility at EMBL Heidelberg for RNA sequencing; the CI TASK supercomputing centre in Gdańsk for computational resources; and Procter & Gamble Poland for institutional support.

I am always indebted to Maciej Cieřlik who provided invaluable support with patience and warmth often going above and beyond his administrative duties. Agnieszka Sikorska provided administrative support that made the most intimidating bureaucracy bearable. I also thank Aleksandra Mrozek for patiently handling all the financial bureaucracies I a breeze.

This work was financially supported by the WELCOME2 ERA Chair (Horizon 2020, No. 952601), the David & Amy Fulton Foundation, the National Science Centre Poland (2023/49/B/NZ9/00402), and King Abdulaziz University, Jeddah (RG: 21-130-38).

Beyond the academic and institutional support that made this work possible, I am equally grateful for the personal connections and friendships that sustained me throughout this journey.

I would like to begin by expressing my heartfelt gratitude to Daniela Milheiro Oliveira and Pedro Navarro. As the first people I connected with upon moving from India to Poland, they became my first family away from home. Sharing the same hostel at Biofarma, they offered me warmth, kindness, and a sense of belonging during a period of significant transition. Through shared meals, conversations, and their generosity in introducing me to Portuguese culture, they made an unfamiliar place feel welcoming. I will always remain deeply grateful for their support.

Building on that initial sense of belonging, I am sincerely thankful to Dr. Mamata Savanagoudar, Divya Das, Dr. Abhipsa Panda, Dr. Manohara Mahadeva, Dr. Rossella Debernardis, Dr. Yashaswi Sharma, Ankur Singh, Matteo Pacewicz, Martin, and Piotr. As my seniors and close companions in this journey, they provided invaluable guidance, encouragement, and emotional support during some of the most challenging periods of my personal and academic life. They formed a second family and a vital support system without which navigating these years would have been far more difficult.

I would especially like to acknowledge Yashaswi, who, as my flatmate, became much more than that. She became a sister whose presence, care, and companionship I will always cherish. I would also like to extend a special note of thanks to Matteo, with whom I have

shared many meaningful discussions on psychiatric and philosophical topics. These conversations have been both grounding and intellectually enriching, and I will always value that connection.

I thank Julia for always lending me an understanding and patient ear for my rants and woes. I found in her the guiding elderly presence that most PhD students seek yet lack. Not only professionally but also personally, she has been a strong anchor to me.

Equally important to this journey have been the friendships that grew into my closest support system. I am deeply grateful to Tanya Tripathi, Srinjoyee Singh Pawar and Aleksandra Maria Kuliga, who have been my pillars of strength throughout this time. Their unwavering support, understanding, and trust have meant more to me than words can convey. They have been my closest confidantes, standing by me through every challenge, no matter how small. Tanya, in particular, has been a constant reminder of home, a connection to my roots in India, and someone I deeply admire for her strength and character.

Alongside them, I would like to thank Reshma Susan Babu and Sandra Brigida Mirończuk, who, although they joined my journey later, have become an integral part of my inner circle. The time spent together, especially while sharing a workspace and moments of cooking and dining, has been one of the most joyful and comforting aspects of my PhD experience. These moments created a sense of home that I will always hold close.

I am also sincerely thankful to Mariusz Jankowski for his companionship and unwavering support. From grocery runs and travels to long night walks and conversations, he has been a constant presence, offering honesty, perspective, and steadfast encouragement through all circumstances.

In a similar spirit of camaraderie, I am grateful to Maciej Rybiński for being a true companion throughout this academic journey. As fellow batchmates, we navigated similar challenges, and I deeply appreciate that our friendship remained strong and unaffected by competition.

I would also like to express my sincere gratitude to Roberto, Juan, Iliana, Adan, Daniel, and Abril for their exceptional warmth and gracious hospitality during my time in Mexico City. They welcomed me with a generosity that made a new environment feel immediately familiar and comfortable, and I remain deeply appreciative of the care and openness with which they introduced me to life in Mexico.

I am especially grateful to Juan and Roberto, who went above and beyond to support me during this transition. Their guidance, kindness, and constant presence ensured that I never felt alone while adapting to a new working environment, and their support made a lasting difference to my experience.

Looking beyond these years, I would like to acknowledge my childhood friends, Manjima Bhattacharjee and Parvathi V Nair, for their enduring presence in my life. They have continually reminded me of my resilience, strength, and the confidence I once carried so effortlessly. Their belief in me, even from afar, has been a source of reassurance and motivation.

Above all, I owe my deepest gratitude to my family, whose influence and sacrifices have shaped every part of who I am today. I come from a family where education, discipline, and intellectual curiosity have always been deeply valued. The example set by my extended family, many of whom are accomplished in their respective academic and professional fields, gave me an early sense of direction and aspiration. This foundation has given me a significant head start, and I remain profoundly grateful for it.

My parents, in particular, have been the driving force behind every opportunity I have had. They built their lives through determination and resilience, establishing themselves in a new city from the ground up while carrying the responsibility of creating a better future for us. In doing so, they often set aside their own aspirations and comforts so that we would not have to. Their sacrifices are woven into every achievement of mine, including this thesis. I am deeply aware that what I have been able to accomplish rests on the strength of everything they have endured and built. It is a debt I can never truly repay, but one I will always strive to honour through my work and the life I lead.

My sister has been an irreplaceable part of my life and a constant source of strength. From the very beginning, she has brought a sense of completeness and joy that I cannot imagine life without. Despite being younger, she has consistently supported and protected me, often helping me navigate even my most difficult thoughts. Her presence has given me the strength to move forward during moments when it felt most challenging.

My mother has been one of the most defining influences in my life. Beyond everything she has given me, she was also my very first friend. Some of my most cherished childhood memories are of her reading stories to me and the simple joy of time spent together. She

instilled in me core values, discipline, and a sense of curiosity that continue to guide me, and I will carry these with me always.

My father has been my wisest advisor and a constant source of stability. In moments when all other avenues seemed uncertain, he has always been the person I could turn to. His perspective, calm guidance, and unwavering support have provided me with a sense of direction and reassurance that I deeply value.

Finally, I would like to thank my boyfriend, Sarthak Chakrabarti, for being a source of positivity and stability in my life. He came into my life at a time when I needed support the most and has since been a grounding presence. His care, understanding, and the way he consistently shows up for me have meant more than I ever expected, and I am deeply grateful for his presence in my life.

*“There is no such thing as a self-made person. You will reach
your goals only with the help of others.”*

— George Shinn

Chapter 16

Author's Statements

Candidate's Contribution to Each Publication

The candidate's scientific contribution to this thesis was exclusively computational and bioinformatic. All wet-laboratory work was performed by co-authors and clinical collaborators as specified below.

Chapter 3 – Scientific Reports (2024)

Wet-lab (not by candidate): RNA isolation and library preparation by Dr. Marjo Malinen; clinical design by Prof. Tuomainen and Prof. Virtanen.

Candidate's contribution: RNA-seq preprocessing and QC; differential gene expression analysis; pathway enrichment; response index algorithm; k-means clustering; figure preparation; manuscript review and editing.

Chapter 4 – Saudi Pharmaceutical Journal (2024)

Wet-lab (not by candidate): blood sampling, ELISA cytokine measurement, and qPCR performed by King Abdulaziz University collaborators.

Candidate's contribution: cytokine and qPCR response scoring algorithms; k-means clustering; concordance analysis; figure preparation; manuscript review and editing.

Chapter 5 – Scientific Reports (2025)

Wet-lab (not by candidate): participant recruitment, blood sampling, PBMC isolation, RNA extraction, and library preparation by the VitDPAS InLife and UWM teams.

Candidate's contribution: full VitDPAS RNA-seq analysis (QC, alignment, differential expression); cross-cohort comparison with VitDHiD; 26-gene core identification; ATAC-seq analysis; response index classification; manuscript writing (original draft, equal contribution with E. Gospodarska).

Chapter 6 – Under review (2026)

Candidate's contribution: RNA-seq differential expression across seven time points; MDS and temporal clustering; pharmacokinetic model fitting; DoRothEA network construction; iterative kernel reduction; ODE formulation; differential evolution optimisation in Julia; cross-cohort validation in VitDHiD; manuscript writing (original draft, lead author).

Chapter 17

Official Documents

Oświadczenie o autorstwie i oryginalności / *Declaration of Authorship and Originality*

Niniejszym oświadczam, że niniejsza rozprawa doktorska została napisana przeze mnie samodzielnie, że wszystkie wykorzystane źródła zostały prawidłowo zacytowane, a praca ta nie była wcześniej przedmiotem rozprawy w ramach innego stopnia naukowego.

Ponadto oświadczam, że publikacje dołączone jako artykuły I–IV stanowią oryginalne prace naukowe, do których mój wkład został opisany w rozdziale 15, a ich wykorzystanie w niniejszej rozprawie zbiorczej zostało zatwierdzone przez wszystkich współautorów.

I hereby declare that this doctoral thesis was written independently by me, that all sources used have been correctly cited, and that this work has not been previously submitted for any other academic degree.

I further declare that the publications included as Appended Papers I–IV represent original scientific work to which my contribution is described in Chapter 15, and that their use in this cumulative thesis has been approved by all co-authors.

Signed: 

Ranjini Ghosh Dastidar

Olsztyn, 2026

Załącznik nr 10 / Annex No. 10 • Paper 1 (Chapter 3) — VitDHiD Sci Rep 2024	
<i>Oświadczenie kandydata o jego merytorycznym udziale w powstanie pracy</i>	
<i>Statement of the candidate on their substantive contribution to the creation of a work</i>	
Imię i nazwisko / Name and surname:	Ranjini Ghosh Dastidar
ORCID:	0000-0001-6180-2523
Miejscowość, data / City, date:	Olsztyn, 11 May 2026

Oświadczenie

Statement

Niniejszym oświadczam, że w pracy „*In vivo vitamin D targets reveal the upregulation of focal adhesion-related genes in primary immune cells of healthy individuals*” (Ranjini Ghosh Dastidar, Julia Jaroslawska, Marjo Malinen, Tomi-Pekka Tuomainen, Jyrki K. Virtanen, Igor Bendik & Carsten Carlberg, *Scientific Reports*, 2024, vol. 14, 17552, DOI: 10.1038/s41598-024-68741-9) mój udział polegał na konceptualizacji, analizie danych, oprogramowaniu, wizualizacji, przygotowaniu pierwszej wersji manuskryptu oraz przygotowaniu odpowiedzi na recenzje.

I hereby declare that in the scientific paper „*In vivo vitamin D targets reveal the upregulation of focal adhesion-related genes in primary immune cells of healthy individuals*” (Ranjini Ghosh Dastidar, Julia Jaroslawska, Marjo Malinen, Tomi-Pekka Tuomainen, Jyrki K. Virtanen, Igor Bendik & Carsten Carlberg, *Scientific Reports*, 2024, vol. 14, 17552, DOI: 10.1038/s41598-024-68741-9) my contribution consisted of conceptualization, formal analysis, software, visualization, writing – original draft, and writing – review & editing.

Konceptualizacja / Conceptualization	Współtworzyłam strategię analizy bioinformatycznej oraz definicję indeksu odpowiedzi na witaminę D opartego na danych RNA-seq. / <i>Contributed to formulation of the bioinformatic analysis strategy and the definition of the RNA-seq-based vitamin D response index.</i>
Analiza danych / Formal analysis	Prowadziłam całą analizę obliczeniową: kontrolę jakości RNA-seq (FastQC v0.11.9), przycinanie adapterów (Cutadapt v3.5), wyrównanie odczytów do GRCh38/hg38 (STAR v2.7.10a), zliczanie odczytów na poziomie genów (featureCounts / subread v2.0.3), normalizację wielkości bibliotek (TMM) oraz analizę różnicowej ekspresji w R/edgeR (FDR < 0,05, n = 25 uczestników). Opracowałam pierwszy oparty na RNA-seq indeks odpowiedzi na witaminę D metodą k-means na profilach log2FC uczestników. Surowe dane sekwencjonowania zdeponowano w NCBI GEO. / <i>Led all computational analysis: RNA-seq quality control (FastQC v0.11.9), adapter trimming (Cutadapt v3.5), read alignment to GRCh38/hg38 (STAR v2.7.10a), gene-level read counting (featureCounts / subread v2.0.3), library-size normalisation (TMM), and differential expression analysis in R/edgeR</i>

		<i>(FDR < 0.05, n = 25 participants). Derived the first RNA-seq-based vitamin D response index via k-means clustering of participant-level log-fold-change profiles. Raw sequencing data deposited to NCBI GEO (GSE accession submitted).</i>
Oprogramowanie Software	/	R (v4.2): edgeR, limma, ggplot2, pheatmap, RColorBrewer; Python (v3.10): pandas, numpy, matplotlib, seaborn; Shell (bash): FastQC, MultiQC, Cutadapt, STAR, Samtools (v1.16), featureCounts; Adnotacja genomu: Ensembl GRCh38.p13. / R (v4.2): <i>edgeR, limma, ggplot2, pheatmap, RColorBrewer</i> ; Python (v3.10): <i>pandas, numpy, matplotlib, seaborn</i> ; Shell (bash): <i>FastQC, MultiQC, Cutadapt, STAR, Samtools (v1.16), featureCounts</i> ; Genome annotation: <i>Ensembl GRCh38.p13</i> .
Wizualizacja Visualization	/	Przygotowałam większość rycin: wykresy MDS/PCA, volcano, MA, mapy cieplne klastrowania k-means oraz wykresy korelacji. / <i>Prepared most figures: MDS/PCA plots, volcano plots, MA plots, k-means clustering heatmaps, and correlation scatter plots.</i>
Redakcja i przygotowanie pierwszej wersji manuskryptu / Writing – Original Draft		Pierwszy autor. Napisałam sekcję metod bioinformatycznych; wniosłam wkład w wyniki i dyskusję. / <i>First Author. Wrote the bioinformatics methods section; contributed to results and discussion.</i>
Przygotowanie odpowiedzi na recenzje / Writing – Review & Editing		Krytyczny przegląd i rewizja pełnego manuskryptu przed złożeniem i po recenzji. / <i>Critical review and revision of the full manuscript prior to submission and after peer review.</i>

Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej zgodnie z taksonomią CRediT (<https://credit.niso.org/>).⁷

podpis kandydata / signature of the candidate:



RANJINI GHOSH DASTIDAR

Ranjini Ghosh Dastidar

Podpisy współautora/współautorów / Signatures of the co-author/co-authors:

Julia Jaroslawska:



Marjo Malinen:



Tomi-Pekka Tuomainen:

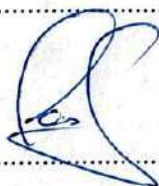
 8-MAY-2026

Jyrki K. Virtanen: Attached in next page

Igor Bendik:



Carsten Carlberg:



Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej zgodnie z taksonomią CRediT (<https://credit.niso.org/>).

podpis kandydata / signature of the candidate:

.....

Ranjini Ghosh Dastidar

Podpisy współautora/współautorów / Signatures of the co-author/co-authors:

Julia Jaroslawska:

.....

Marjo Malinen:

.....

Tomi-Pekka Tuomainen:

.....

Jyrki K. Virtanen:

Jyrki Virtanen

.....

Igor Bendik:

.....

Carsten Carlberg:

.....

Tämä dokumentti on allekirjoitettu sähköisesti UEF Sign-järjestelmällä / This document has been electronically signed using UEF Sign
Annex10_ViDiHD.pdf / bf9c5e3e-3955-4060-b6c2-69228a2a38c7 / sivu/page 2(3)

Tämä dokumentti on allekirjoitettu sähköisesti UEF Sign-järjestelmällä
This document has been electronically signed using UEF Sign



Päiväys / Date: 11.05.2026 09:04:47 (UTC +0300)

Jyrki Virtanen
Tutkimusjohtaja
Kansanterveystiede ja kliininen ravitsemustiede
jyrki.virtanen@uef.fi

Organisation varmentama (UEF-käyttäjätunnus) (eIDAS-tunnistamisen varmuustaso: korotettu)
Certified by organization (UEF user account) (eIDAS level of assurance: substantial)

Tämä dokumentti on allekirjoitettu sähköisesti UEF Sign-järjestelmällä / This document has been electronically signed using UEF Sign
Annex10_VicDHD.pdf / 35aba3a-3602-4e41-a72c-bf6bb1b4236 / sivu/page 4(4)

⁷Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej (<https://credit.niso.org/>)

W celu jednoznacznego określenia wkładu kandydatów w powstanie publikacji naukowej, przyjmuje się następującą klasyfikację:

- **Konceptualizacja:** Tworzenie idei; formułowanie lub rozwijanie głównych celów, hipotez i założeń badawczych.
- **Opracowanie danych:** Zarządzanie danymi, w tym ich adnotacja (tworzenie metadanych), przechowywanie danych badawczych do użytku bieżącego i/lub ponownej analizy.
- **Analiza danych:** Zastosowanie technik statystycznych, matematycznych, obliczeniowych lub innych metod do analizy danych badawczych.
- **Pozyskanie finansowania:** Zdobycie środków finansowych na realizację projektu, który doprowadził do powstania publikacji.
- **Badania:** Prowadzenie procesu badawczego, w tym przeprowadzanie eksperymentów oraz zbieranie danych.
- **Metodologia:** Opracowanie metod badawczych; tworzenie modeli badawczych.
- **Zarządzanie projektem:** Koordynacja i zarządzanie planowaniem oraz realizacją działań badawczych.
- **Oprogramowanie:** Programowanie; projektowanie programów komputerowych; implementacja kodu komputerowego i wspierających algorytmów; testowanie istniejących komponentów kodu.
- **Zasoby:** Zapewnienie materiału badawczego, odczynników, narzędzi badawczych, pacjentów, próbek laboratoryjnych, zwierząt doświadczalnych, sprzętu, zasobów obliczeniowych lub innych narzędzi analitycznych.
- **Walidacja:** Weryfikacja powtarzalności i odtwarzalności wyników badań, eksperymentów oraz innych rezultatów naukowych.
- **Wizualizacja:** Przygotowanie, opracowanie i/lub prezentacja wyników badawczych, w szczególności wizualizacja danych.
- **Redakcja i przygotowanie pierwszej wersji manuskryptu:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej, w szczególności napisanie pierwszej wersji manuskryptu.
- **Przygotowanie odpowiedzi na recenzje:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej przez autorów manuskryptu, obejmująca krytyczną analizę, komentarze oraz poprawki.

To clearly define the candidate's contribution to the publication, the following taxonomy (as per credit.niso.org) is recommended:

- **Conceptualization:** Ideas; formulation or evolution of overarching research goals and aims.
- **Data curation:** Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- **Formal analysis:** Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.
- **Funding acquisition:** Acquisition of the financial support for the project leading to this publication.
- **Investigation:** Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- **Methodology:** Development or design of methodology; creation of models.
- **Project administration:** Management and coordination responsibility for the research activity planning and execution.
- **Software:** Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- **Resources:** Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- **Validation:** Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- **Visualization:** Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- **Writing – Original Draft:** Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- **Writing – Review & Editing:** Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

Załącznik nr 10 / Annex No. 10 • Paper 2 (Chapter 4) — Saudi Cohort Saudi Pharm J 2024	
<i>Oświadczenie kandydata o jego merytorycznym udziale w powstanie pracy</i>	
<i>Statement of the candidate on their substantive contribution to the creation of a work</i>	
Imię i nazwisko / Name and surname:	Ranjini Ghosh Dastidar
ORCID:	0000-0001-6180-2523
Miejscowość, data / City, date:	Olsztyn, 11 May 2026

Oświadczenie

Statement

Niniejszym oświadczam, że w pracy „*Evaluation of the vitamin D response index in a Saudi cohort*” (Shareefa A. AlGhamdi*, Ranjini Ghosh Dastidar*, Maciej Rybiński, Hadeil M. Alsufiani, Sawsan O. Khoja, Nusaibah N. Enaibsi, Safa F. Saif & Carsten Carlberg, *Saudi Pharmaceutical Journal*, 2024, vol. 32, 102137, DOI: 10.1016/j.jsps.2024.102137)mój udział polegał na koncepcualizacji, analizie danych, oprogramowaniu, wizualizacji, przygotowaniu pierwszej wersji manuskryptu oraz przygotowaniu odpowiedzi na recenzje.

I hereby declare that in the scientific paper „*Evaluation of the vitamin D response index in a Saudi cohort*” (Shareefa A. AlGhamdi*, Ranjini Ghosh Dastidar*, Maciej Rybiński, Hadeil M. Alsufiani, Sawsan O. Khoja, Nusaibah N. Enaibsi, Safa F. Saif & Carsten Carlberg, *Saudi Pharmaceutical Journal*, 2024, vol. 32, 102137, DOI: 10.1016/j.jsps.2024.102137)my contribution consisted of conceptualization, formal analysis, software, visualization, writing – original draft, and writing – review & editing.

Koncepcualizacja / Conceptualization	Współtworzyłam algorytm punktacji indeksu odpowiedzi na witaminę D oparty na cytokinach oraz schemat klasyfikacji uczestników. / <i>Co-conceptualised the cytokine-based vitamin D response index scoring algorithm and participant classification framework.</i>
Analiza danych / Formal analysis	Prowadziłam całą analizę statystyczną: opracowanie ważonego algorytmu punktacji odpowiedzi cytokinowej; klastrowanie k-means (pakiet R stats) 100 uczestników saudyjskich na respondentów wysokich, średnich i niskich; analiza zgodności między klasyfikacjami opartymi na cytokinach i qPCR (88,6% zgodności, n = 37 uczestników). / <i>Led all statistical analysis: development of a weighted cytokine-response scoring algorithm; k-means clustering (R stats package) of 100 Saudi participants into high, mid, and low responders; concordance analysis between cytokine- and qPCR-based classifications (88.6% concordance, n = 37 participants).</i>
Oprogramowanie / Software	R (v4.2): stats (k-means, cor.test), ggplot2, dplyr, tidyr; Python (v3.10): pandas, scipy, matplotlib; Microsoft Excel (zestawianie wyników czytnika płytek ELISA). / <i>R (v4.2): stats (k-means, cor.test), ggplot2, dplyr, tidyr;</i>

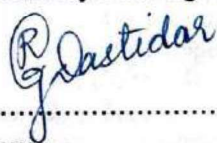
	<i>Python (v3.10): pandas, scipy, matplotlib; Microsoft Excel (collation of ELISA plate-reader outputs).</i>
Wizualizacja Visualization	Przygotowałam wszystkie ryciny: wykresy słupkowe odpowiedzi cytokinowej, wykresy skumulowanego indeksu odpowiedzi uczestników oraz wykresy klasyfikacji zgodności/niezgodności. / <i>Prepared all figures: cytokine response bar plots, participant-ranked cumulative response index plots, concordance/discordance classification scatter plots.</i>
Redakcja i przygotowanie pierwszej wersji manuskryptu / Writing – Original Draft	Wniosłam wkład jako współpierwszy autor (równy udział, oznaczony *): napisałam sekcje metod statystycznych i wyników. / <i>Contributed as joint first author (equal contribution, denoted *): wrote statistical methods and results sections.</i>
Przygotowanie odpowiedzi na recenzje / Writing – Review & Editing	Krytyczny przegląd i rewizja pełnego manuskryptu przed złożeniem i po recenzji. / <i>Critical review and revision of the full manuscript prior to submission and after peer review.</i>

Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej zgodnie z taksonomią CRediT (<https://credit.niso.org/>).⁷

	<i>cytokine response bar plots, participant-ranked cumulative response index plots, concordance/discordance classification scatter plots.</i>
Redakcja i przygotowanie pierwszej wersji manuskryptu / Writing – Original Draft	Wniosłam wkład jako współpierwszy autor (równy udział, oznaczony *): napisałam sekcje metod statystycznych i wyników. / <i>Contributed as joint first author (equal contribution, denoted *): wrote statistical methods and results sections.</i>
Przygotowanie odpowiedzi na recenzje / Writing – Review & Editing	Krytyczny przegląd i rewizja pełnego manuskryptu przed złożeniem i po recenzji. / <i>Critical review and revision of the full manuscript prior to submission and after peer review.</i>

Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej zgodnie z taksonomią CRediT (<https://credit.niso.org/>).⁷

podpis kandydata / signature of the candidate:



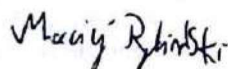
Ranjini Ghosh Dastidar

Podpisy współautora/współautorów / Signatures of the co-author/co-authors:

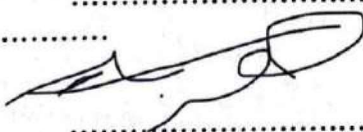
Shareefa A. AlGhamdi:



Maciej Rybiński:



Hadeil M. Alsufiani:



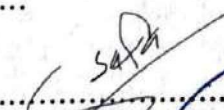
Sawsan O. Khoja:



Nusaibah N. Enaibsi:



Safa F. Saif:



Carsten Carlberg:



⁷Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej (<https://credit.niso.org/>)

W celu jednoznacznego określenia wkładu kandydatów w powstanie publikacji naukowej, przyjmuje się następującą klasyfikację:

- **Konceptualizacja:** Tworzenie idei; formułowanie lub rozwijanie głównych celów, hipotez i założeń badawczych.
- **Opracowanie danych:** Zarządzanie danymi, w tym ich adnotacja (tworzenie metadanych), przechowywanie danych badawczych do użytku bieżącego i/lub ponownej analizy.
- **Analiza danych:** Zastosowanie technik statystycznych, matematycznych, obliczeniowych lub innych metod do analizy danych badawczych.
- **Pozyskanie finansowania:** Zdobycie środków finansowych na realizację projektu, który doprowadził do powstania publikacji.
- **Badania:** Prowadzenie procesu badawczego, w tym przeprowadzanie eksperymentów oraz zbieranie danych.
- **Metodologia:** Opracowanie metod badawczych; tworzenie modeli badawczych.
- **Zarządzanie projektem:** Koordynacja i zarządzanie planowaniem oraz realizacją działań badawczych.
- **Oprogramowanie:** Programowanie; projektowanie programów komputerowych; implementacja kodu komputerowego i wspierających algorytmów; testowanie istniejących komponentów kodu.
- **Zasoby:** Zapewnienie materiału badawczego, odczynników, narzędzi badawczych, pacjentów, próbek laboratoryjnych, zwierząt doświadczalnych, sprzętu, zasobów obliczeniowych lub innych narzędzi analitycznych.
- **Walidacja:** Weryfikacja powtarzalności i odtwarzalności wyników badań, eksperymentów oraz innych rezultatów naukowych.
- **Wizualizacja:** Przygotowanie, opracowanie i/lub prezentacja wyników badawczych, w szczególności wizualizacja danych.
- **Redakcja i przygotowanie pierwszej wersji manuskryptu:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej, w szczególności napisanie pierwszej wersji manuskryptu.
- **Przygotowanie odpowiedzi na recenzje:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej przez autorów manuskryptu, obejmująca krytyczną analizę, komentarze oraz poprawki.

To clearly define the candidate's contribution to the publication, the following taxonomy (as per credit.niso.org) is recommended:

- **Conceptualization:** Ideas; formulation or evolution of overarching research goals and aims.
- **Data curation:** Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- **Formal analysis:** Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.
- **Funding acquisition:** Acquisition of the financial support for the project leading to this publication.
- **Investigation:** Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- **Methodology:** Development or design of methodology; creation of models.
- **Project administration:** Management and coordination responsibility for the research activity planning and execution.
- **Software:** Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- **Resources:** Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- **Validation:** Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- **Visualization:** Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- **Writing – Original Draft:** Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- **Writing – Review & Editing:** Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

Załącznik nr 10 / Annex No. 10 • Paper 3 (Chapter 5) — VitDPAS Sci Rep 2025	
<i>Oświadczenie kandydata o jego merytorycznym udziale w powstanie pracy</i>	
<i>Statement of the candidate on their substantive contribution to the creation of a work</i>	
Imię i nazwisko / Name and surname:	Ranjini Ghosh Dastidar
ORCID:	0000-0001-6180-2523
Miejscowość, data / City, date:	Olsztyn, 11 May 2026

Oświadczenie

Statement

Niniejszym oświadczam, że w pracy „*Transcriptomic profiling of immune modulation induced by vitamin D3 in the VitDPAS and VitDHiD cohort studies*” (Emilia Gospodarska*, Ranjini Ghosh Dastidar*, Julia Jaroslawska, Maciej Rybiński, Marianna Raczyk, Kornelia Tokarczyk-Malesa, Jerzy Romaszko & Carsten Carlberg, *Scientific Reports*, 2025, vol. 15, 17334, DOI: 10.1038/s41598-025-02495-w)mój udział polegał na konceptualizacji, analizie danych, oprogramowaniu, wizualizacji, przygotowaniu pierwszej wersji manuskryptu oraz przygotowaniu odpowiedzi na recenzje.

I hereby declare that in the scientific paper „*Transcriptomic profiling of immune modulation induced by vitamin D3 in the VitDPAS and VitDHiD cohort studies*” (Emilia Gospodarska*, Ranjini Ghosh Dastidar*, Julia Jaroslawska, Maciej Rybiński, Marianna Raczyk, Kornelia Tokarczyk-Malesa, Jerzy Romaszko & Carsten Carlberg, *Scientific Reports*, 2025, vol. 15, 17334, DOI: 10.1038/s41598-025-02495-w)my contribution consisted of conceptualization, formal analysis, software, visualization, writing – original draft, and writing – review & editing.

Konceptualizacja / Conceptualization	Współtworzyłam strategię analizy porównawczej między kohortami (kohorta polska VitDPAS vs. kohorta fińska VitDHiD) oraz podejście oparte na zachowanym rdzeniu sygnatury ekspresji genów witaminy D. / <i>Co-conceptualised the cross-cohort comparative analysis strategy (VitDPAS Polish cohort vs. VitDHiD Finnish cohort) and the conserved vitamin D gene-expression core-signature approach.</i>
Analiza danych / Formal analysis	Główny bioinformatyk: kontrola jakości RNA-seq (FastQC v0.11.9), przycinanie adapterów (Cutadapt v3.5), wyrównanie do GRCh38/hg38 (STAR v2.7.10a), zliczanie genów (featureCounts / subread v2.0.3), normalizacja TMM, analiza różnicowej ekspresji (R/edgeR, FDR < 0,05) dla 45 uczestników VitDPAS (287 bibliotek podłużnych); analiza Venna między kohortami identyfikująca zachowaną sygnaturę 26 genów. Surowe dane zdeponowane w NCBI GEO. / <i>Lead bioinformatician: RNA-seq QC (FastQC v0.11.9), adapter trimming (Cutadapt v3.5), alignment to GRCh38/hg38 (STAR v2.7.10a), gene counting (featureCounts / subread v2.0.3), TMM normalisation, differential expression (R/edgeR, FDR < 0.05) for 45 VitDPAS participants (287 longitudinal libraries); cross-cohort Venn</i>

		<i>analysis identifying the conserved 26-gene core signature. Raw sequencing data deposited to NCBI GEO (GSE accession submitted).</i>
Oprogramowanie Software	/	R (v4.2): edgeR, limma, ggplot2, VennDiagram, pheatmap; Python (v3.10): pandas, numpy, matplotlib, seaborn; Shell (bash): FastQC, MultiQC, Trim Galore, Cutadapt, STAR, Samtools, featureCounts; Adnotacja genomu: Ensembl GRCh38.p13. / R (v4.2): <i>edgeR, limma, ggplot2, VennDiagram, pheatmap</i> ; Python (v3.10): <i>pandas, numpy, matplotlib, seaborn</i> ; Shell (bash): <i>FastQC, MultiQC, Trim Galore, Cutadapt, STAR, Samtools, featureCounts</i> ; Genome annotation: <i>Ensembl GRCh38.p13</i> .
Wizualizacja Visualization	/	Przygotowałam większość rycin: wykresy MDS, volcano, diagramy Venna nakładania się genów między kohortami oraz podłużne mapy cieplne ekspresji. / <i>Prepared most figures: MDS plots, volcano plots, Venn diagrams of cross-cohort gene overlaps, and longitudinal expression heatmaps.</i>
Redakcja i przygotowanie pierwszej wersji manuskryptu / Writing – Original Draft		Wniosłam wkład jako współpierwszy autor (równy udział, oznaczony *): napisałam sekcję metod bioinformatycznych; wniosłam wkład w wyniki i dyskusję. / <i>Contributed as joint first author (equal contribution, denoted *): wrote the bioinformatics methods section; contributed to results and discussion.</i>
Przygotowanie odpowiedzi na recenzje / Writing – Review & Editing		Krytyczny przegląd i rewizja pełnego manuskryptu przed złożeniem i po recenzji. / <i>Critical review and revision of the full manuscript prior to submission and after peer review.</i>

Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej zgodnie z taksonomią CRediT (<https://credit.niso.org/>).⁷

podpis kandydata / signature of the candidate:



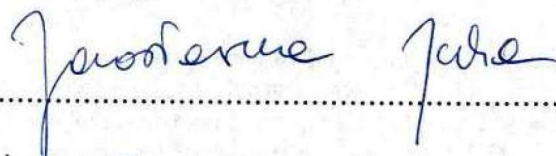
Ranjini Ghosh Dastidar

Podpisy współautora/współautorów / Signatures of the co-author/co-authors:

Emilia Gospodarska:



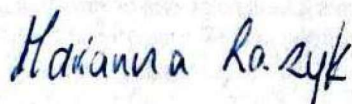
Julia Jaroslawska:



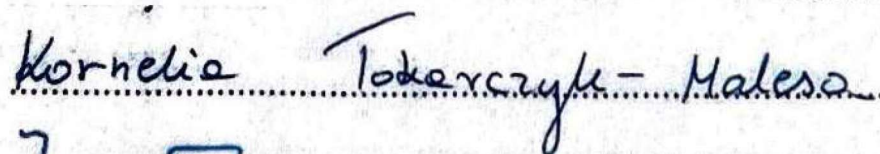
Maciej Rybiński:



Marianna Raczyk:



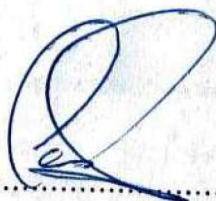
Kornelia Tokarczyk-Malesa:



Jerzy Romaszko:



Carsten Carlberg:



⁷Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej (<https://credit.niso.org/>)

W celu jednoznacznego określenia wkładu kandydatów w powstanie publikacji naukowej, przyjmuje się następującą klasyfikację:

- **Konceptualizacja:** Tworzenie idei; formułowanie lub rozwijanie głównych celów, hipotez i założeń badawczych.
- **Opracowanie danych:** Zarządzanie danymi, w tym ich adnotacja (tworzenie metadanych), przechowywanie danych badawczych do użytku bieżącego i/lub ponownej analizy.
- **Analiza danych:** Zastosowanie technik statystycznych, matematycznych, obliczeniowych lub innych metod do analizy danych badawczych.
- **Pozyskanie finansowania:** Zdobycie środków finansowych na realizację projektu, który doprowadził do powstania publikacji.
- **Badania:** Prowadzenie procesu badawczego, w tym przeprowadzanie eksperymentów oraz zbieranie danych.
- **Metodologia:** Opracowanie metod badawczych; tworzenie modeli badawczych.
- **Zarządzanie projektem:** Koordynacja i zarządzanie planowaniem oraz realizacją działań badawczych.
- **Oprogramowanie:** Programowanie; projektowanie programów komputerowych; implementacja kodu komputerowego i wspierających algorytmów; testowanie istniejących komponentów kodu.
- **Zasoby:** Zapewnienie materiału badawczego, odczynników, narzędzi badawczych, pacjentów, próbek laboratoryjnych, zwierząt doświadczalnych, sprzętu, zasobów obliczeniowych lub innych narzędzi analitycznych.
- **Walidacja:** Weryfikacja powtarzalności i odtwarzalności wyników badań, eksperymentów oraz innych rezultatów naukowych.
- **Wizualizacja:** Przygotowanie, opracowanie i/lub prezentacja wyników badawczych, w szczególności wizualizacja danych.
- **Redakcja i przygotowanie pierwszej wersji manuskryptu:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej, w szczególności napisanie pierwszej wersji manuskryptu.
- **Przygotowanie odpowiedzi na recenzje:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej przez autorów manuskryptu, obejmująca krytyczną analizę, komentarze oraz poprawki.

To clearly define the candidate's contribution to the publication, the following taxonomy (as per credit.niso.org) is recommended:

- **Conceptualization:** Ideas; formulation or evolution of overarching research goals and aims.
- **Data curation:** Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- **Formal analysis:** Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.
- **Funding acquisition:** Acquisition of the financial support for the project leading to this publication.
- **Investigation:** Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- **Methodology:** Development or design of methodology; creation of models.
- **Project administration:** Management and coordination responsibility for the research activity planning and execution.
- **Software:** Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- **Resources:** Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- **Validation:** Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- **Visualization:** Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- **Writing – Original Draft:** Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- **Writing – Review & Editing:** Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

Załącznik nr 10 / Annex No. 10 • Paper 4 (Chapter 6) — ODE Modelling Under Review	
<i>Oświadczenie kandydata o jego merytorycznym udziale w powstanie pracy</i>	
<i>Statement of the candidate on their substantive contribution to the creation of a work</i>	
Imię i nazwisko / Name and surname:	Ranjini Ghosh Dastidar
ORCID:	0000-0001-6180-2523
Miejscowość, data / City, date:	Olsztyn, 11 May 2026

Oświadczenie

Statement

Niniejszym oświadczam, że w pracy „*Dynamic system identification defines a minimal transcription factor control architecture of human endocrine gene regulation*” (Ranjini Ghosh Dastidar, Emilia Gospodarska, Kornelia Tokarczyk-Malesa, Jerzy Romaszko, Elisa Domínguez-Hüttinger & Carsten Carlberg, Manuscript under review, 2025/2026, DOI: Submitted; DOI pending)mój udział polegał na koncepcualizacji, metodologii, analizie danych, oprogramowaniu, wizualizacji, przygotowaniu pierwszej wersji manuskryptu oraz przygotowaniu odpowiedzi na recenzje.

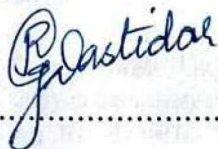
I hereby declare that in the scientific paper „*Dynamic system identification defines a minimal transcription factor control architecture of human endocrine gene regulation*” (Ranjini Ghosh Dastidar, Emilia Gospodarska, Kornelia Tokarczyk-Malesa, Jerzy Romaszko, Elisa Domínguez-Hüttinger & Carsten Carlberg, Manuscript under review, 2025/2026, DOI: Submitted; DOI pending)my contribution consisted of conceptualization, methodology, formal analysis, software, visualization, writing – original draft, and writing – review & editing.

Koncepcualizacja / Conceptualization	Główny autor; opracowałam podejście oparte na jądrze regulatorowym ODE, iteracyjne ramy redukcji sieci (od 1 633 interakcji do minimalnego jądra 117 interakcji / 11 TF) oraz strategię walidacji między kohortami. / <i>Lead author; conceptualised the ODE-based regulatory kernel approach, the iterative network reduction framework (from 1,633 interactions to a minimal 117-interaction / 11-TF kernel), and the cross-cohort validation strategy.</i>
Analiza danych / Formal analysis	Cała analiza bioinformatyczna i matematyczna: różnicowa ekspresja w siedmiu podłużnych punktach czasowych i trzech cyklach suplementacji (VitDPAS, R/edgeR); budowa sieci TF–cel na podstawie regulonu DoRothEA (R/dorothea, viper); modelowanie farmakokinetyczne indywidualnych trajektorii 25(OH)D3 w surowicy (Python/scipy curve_fit, równanie Hilla); formułowanie modelu ODE z kinetyką aktywacji/represji typu Hilla; dopasowanie parametrów metodą ewolucji różnicowej (Julia v1.9: DifferentialEquations.jl, Optim.jl); walidacja ODE między kohortami na VitDHiD ($R^2 > 0,94$). / <i>All bioinformatic and mathematical analysis: differential expression at seven longitudinal time points and three supplementation cycles (VitDPAS, R/edgeR); TF–target network construction from the DoRothEA regulon (R/dorothea, viper);</i>

		<i>pharmacokinetic modelling of individual serum 25(OH)D3 trajectories (Python/scipy curve_fit, Hill equation); ODE model formulation with Hill-type activation/repression kinetics; parameter fitting via differential evolution (Julia v1.9: DifferentialEquations.jl, Optim.jl); cross-cohort ODE validation on VitDHiD ($R^2 > 0.94$).</i>
Metodologia Methodology	/	Zaprojektowałam kompletny potok identyfikacji systemów: (i) redukcja sieci regulacji genów oparta na wnioskach o aktywności TF (VIPER / DoRothEA); (ii) montaż jądra ODE; (iii) iteracyjne przycinanie kierowane analizą wrażliwości. / <i>Designed the complete systems identification pipeline: (i) TF-activity inference-based gene regulatory network reduction (VIPER / DoRothEA); (ii) ODE kernel assembly; (iii) iterative sensitivity-analysis-guided pruning.</i>
Oprogramowanie Software	/	Julia (v1.9): DifferentialEquations.jl, Optim.jl (ewolucja różnicowa), Plots.jl; R (v4.2): edgeR, limma, dorothea, viper, igraph, ggplot2; Python (v3.10): pandas, numpy, scipy (całkowanie ODE, dopasowanie farmakokinetyczne), matplotlib; Shell: FastQC, Cutadapt, STAR, featureCounts, Samtools. / <i>Julia (v1.9): DifferentialEquations.jl, Optim.jl (differential evolution), Plots.jl; R (v4.2): edgeR, limma, dorothea, viper, igraph, ggplot2; Python (v3.10): pandas, numpy, scipy (ODE integration, pharmacokinetic fitting), matplotlib; Shell: FastQC, Cutadapt, STAR, featureCounts, Samtools.</i>
Wizualizacja Visualization	/	Przygotowałam wszystkie ryciny: wykresy volcano amplifikacji czasowej, panele dopasowania modelu ODE dla wszystkich 11 TF w siedmiu punktach czasowych, diagramy osi regulatorowej TP53–AP1–MYC, schematy redukcji sieci oraz wykresy walidacji między kohortami. / <i>Prepared all figures: temporal amplification volcano plots, ODE model fit panels for all 11 TFs across seven time points, TP53–AP1–MYC regulatory axis diagrams, network reduction flow charts, and cross-cohort validation scatter plots.</i>
Redakcja i przygotowanie pierwszej wersji manuskryptu / Writing – Original Draft		Główny autor; wniosłam istotny wkład we wszystkie sekcje manuskryptu — w tym metody, wyniki i dyskusję — reprezentując znacznie większy wkład autorski niż w poprzednich publikacjach. / <i>Lead author; made substantial contributions to all sections of the manuscript — including methods, results, and discussion — representing significantly greater authorial input than in previous publications.</i>
Przygotowanie odpowiedzi na recenzje / Writing – Review & Editing		Krytyczny przegląd i rewizja pełnego manuskryptu. / <i>Critical review and revision of the full manuscript.</i>

Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej zgodnie z taksonomią CRediT (<https://credit.niso.org/>).⁷

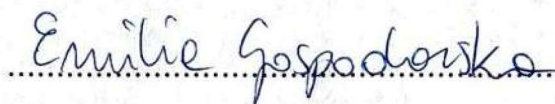
podpis kandydata / signature of the candidate:



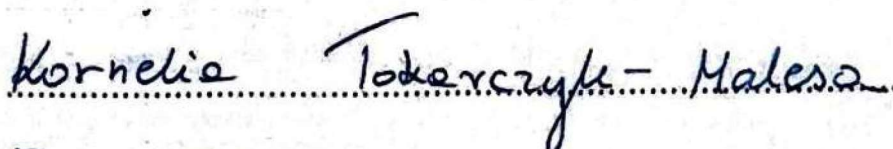
Ranjini Ghosh Dastidar

Podpisy współautora/współautorów / Signatures of the co-author/co-authors:

Emilia Gospodarska:



Kornelia Tokarczyk-Malesa:



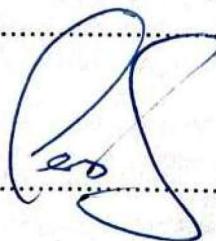
Jerzy Romaszko:



Elisa Domínguez-Hüttinger:



Carsten Carlberg:



⁷Określenie merytorycznego udziału kandydata w powstanie artykułu naukowego/monografii naukowej (<https://credit.niso.org/>)

W celu jednoznacznego określenia wkładu kandydatów w powstanie publikacji naukowej, przyjmuje się następującą klasyfikację:

- **Konceptualizacja:** Tworzenie idei; formułowanie lub rozwijanie głównych celów, hipotez i założeń badawczych.
- **Opracowanie danych:** Zarządzanie danymi, w tym ich adnotacja (tworzenie metadanych), przechowywanie danych badawczych do użytku bieżącego i/lub ponownej analizy.
- **Analiza danych:** Zastosowanie technik statystycznych, matematycznych, obliczeniowych lub innych metod do analizy danych badawczych.
- **Pozyskanie finansowania:** Zdobywanie środków finansowych na realizację projektu, który doprowadził do powstania publikacji.
- **Badania:** Prowadzenie procesu badawczego, w tym przeprowadzanie eksperymentów oraz zbieranie danych.
- **Metodologia:** Opracowanie metod badawczych; tworzenie modeli badawczych.
- **Zarządzanie projektem:** Koordynacja i zarządzanie planowaniem oraz realizacją działań badawczych.
- **Oprogramowanie:** Programowanie; projektowanie programów komputerowych; implementacja kodu komputerowego i wspierających algorytmów; testowanie istniejących komponentów kodu.
- **Zasoby:** Zapewnienie materiału badawczego, odczynników, narzędzi badawczych, pacjentów, próbek laboratoryjnych, zwierząt doświadczalnych, sprzętu, zasobów obliczeniowych lub innych narzędzi analitycznych.
- **Walidacja:** Weryfikacja powtarzalności i odtwarzalności wyników badań, eksperymentów oraz innych rezultatów naukowych.
- **Wizualizacja:** Przygotowanie, opracowanie i/lub prezentacja wyników badawczych, w szczególności wizualizacja danych.
- **Redakcja i przygotowanie pierwszej wersji manuskryptu:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej, w szczególności napisanie pierwszej wersji manuskryptu.
- **Przygotowanie odpowiedzi na recenzje:** Opracowanie, stworzenie i/lub prezentacja pracy naukowej przez autorów manuskryptu, obejmująca krytyczną analizę, komentarze oraz poprawki.

To clearly define the candidate's contribution to the publication, the following taxonomy (as per credit.niso.org) is recommended:

- **Conceptualization:** Ideas; formulation or evolution of overarching research goals and aims.
- **Data curation:** Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- **Formal analysis:** Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.
- **Funding acquisition:** Acquisition of the financial support for the project leading to this publication.
- **Investigation:** Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- **Methodology:** Development or design of methodology; creation of models.
- **Project administration:** Management and coordination responsibility for the research activity planning and execution.
- **Software:** Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- **Resources:** Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- **Validation:** Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- **Visualization:** Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- **Writing – Original Draft:** Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- **Writing – Review & Editing:** Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

Chapter 18

Appended Publications

The four publications constituting this cumulative doctoral thesis are appended in full in the order they appear in the thesis narrative.

Appended Paper I

In Vivo Vitamin D Targets Reveal the Upregulation of Focal Adhesion-Related Genes in Primary Immune Cells of Healthy Individuals

Ranjini Ghosh Dastidar, Julia Jaroslawska, Marjo Malinen, Tomi-Pekka Tuomainen, Jyrki K. Virtanen, Igor Bendik & Carsten Carlberg

Scientific Reports (2024) 14:17552 • DOI: 10.1038/s41598-024-68741-9

Appended Paper II

Evaluation of the Vitamin D Response Index in a Saudi Cohort

Shareefa A. AlGhamdi*, Ranjini Ghosh Dastidar*, Maciej Rybinski, Hadeil M. Alsufiani, Sawsan O. Khoja, Nusaibah N. Enaibsi, Safa F. Saif & Carsten Carlberg

Saudi Pharmaceutical Journal (2024) 32:102137 • DOI: 10.1016/j.jsps.2024.102137

Appended Paper III

Transcriptomic Profiling of Immune Modulation Induced by Vitamin D₃ in the VitDPAS and VitDHiD Cohort Studies

Emilia Gospodarska*, Ranjini Ghosh Dastidar*, Julia Jaroslawska, Maciej Rybinski, Marianna Raczky, Kornelia Tokarczyk-Malesa, Jerzy Romaszko & Carsten Carlberg

Scientific Reports (2025) 15:17334 • DOI: 10.1038/s41598-025-02495-w

Appended Paper IV

Dynamic System Identification Defines a Minimal Transcription Factor Control Architecture of Human Endocrine Gene Regulation

Ranjini Ghosh Dastidar, Emilia Gospodarska, Kornelia Tokarczyk-Malesa, Jerzy Romaszko, Elisa Domínguez-Hüttinger & Carsten Carlberg

Under review

Fwd: Nucleic Acids Research - Manuscript ID NAR-01039-X-2026

From: Carsten Carlberg <c.carlberg@pan.olsztyn.pl>
To: Ranjini Ghosh Dastidar <r.dastidar@pan.olsztyn.pl>

Sat, 09-05-2026 12:53

Attachments: In-Life-stopka-logo.png (7 KB)

This is the last e-mail I got from NAR

Prof. Dr. Carsten Carlberg

ERA chair
Nutrigenomics team

Institute of Animal Reproduction
and Food Research
Polish Academy of Sciences



ul. Trylińskiego 18, 10-683 Olsztyn, Poland
www.pan.olsztyn.pl

Save the date for the 22. annual conference

NuGOweek 2026 (<https://nugo2026.pan.olsztyn.pl>)
'Healthy aging through diet'
From 31 August to 2 September 2026, Olsztyn, Poland

Begin forwarded message:

From: Nucleic Acids Research <onbehalf@manuscriptcentral.com>
Subject: Nucleic Acids Research - Manuscript ID NAR-01039-X-2026
Date: 1. April 2026 at 08:53:54 CEST
To: c.carlberg@pan.olsztyn.pl
Reply-To: [REDACTED]

01-Apr-2026

NAR-01039-X-2026

Dynamic system identification defines a minimal transcription factor control architecture of human endocrine gene regulation
Ghosh Dastidar, Ranjini ; Gospodarska, Emilia ; Tokarczyk-Malesa, Kornelia ; Romaszko, Jerzy ; Domínguez-Hüttinger, Elisa ; Carlberg, Carsten

Dear Dr Carlberg,

This email confirms that your paper has been received and potential peer reviewers are now being invited to examine the study. We hope to have reviews back and further information to you in about three weeks, if the reviewers follow the requested schedule. If we have difficulties in obtaining willing reviewers, this schedule may be extended.

Best wishes,

Prof. Andrew Sharrocks
Executive Editor
Nucleic Acids Research

Professor of Molecular Biology
Faculty of Biology, Medicine and Health
University of Manchester
UK

Realizując wymogi Rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE prosimy o zapoznanie się z udostępnioną na stronie internetowej informacją na temat przetwarzania danych osobowych przez InLife Instytut Rozrodu Żwierząt i Badań Żywności Polskiej Akademii Nauk <http://pan.olsztyn.pl/instytut/informacje-ogolne/rodo/>



OPEN

In vivo vitamin D targets reveal the upregulation of focal adhesion-related genes in primary immune cells of healthy individuals

Ranjini Ghosh Dastidar¹, Julia Jaroslawska¹, Marjo Malinen², Tomi-Pekka Tuomainen³, Jyrki K. Virtanen³, Igor Bendik⁴ & Carsten Carlberg^{1,5}✉

Vitamin D modulates innate and adaptive immunity, the molecular mechanisms of which we aim to understand under human in vivo conditions. Therefore, we designed the study VitDHiD (NCT03537027) as a human investigation, in which 25 healthy individuals were supplemented with a single vitamin D₃ bolus (80,000 IU). Transcriptome-wide differential gene expression analysis of peripheral blood mononuclear cells (PBMCs), which were isolated directly before and 24 h after supplementation, identified 452 genes significantly (FDR < 0.05) responding to vitamin D. In vitro studies using PBMCs from the same individuals confirmed 138 of these genes as targets of 1 α ,25-dihydroxyvitamin D₃. A subset of the 91 most regulated in vivo vitamin D target genes indicated focal adhesion as the major pathway being upregulated by vitamin D₃ supplementation of healthy individuals. Differences in the individual-specific responsiveness of in vivo vitamin D target genes in relation to the increase of the person's vitamin D status allowed a segregation of the VitDHiD participants into 9 high, 12 mid and 4 low responders. The expression profile of nearly 600 genes elucidate the difference between high and low vitamin D responders, the most prominent of which is the HLA-C (major histocompatibility complex, class I, C) gene.

Keywords Vitamin D, Vitamin D intervention study, Transcriptome, Vitamin D target genes, Vitamin D response index, Focal adhesion

Abbreviations

1,25(OH) ₂ D ₃ or 1,25D	1 α ,25-Dihydroxyvitamin D ₃
25(OH)D ₃	25-Hydroxyvitamin D ₃
AKT	AKT murine thymoma viral oncogene homolog
ALOX5	Arachidonate 5-lipoxygenase
ANPEP	Alanyl aminopeptidase, membrane
AQP9	Aquaporin 9
ASAP2	ArfGAP with SH3 domain, ankyrin repeat and PH domain 2
BMI	Body mass index
CD	Cluster of differentiation
CDKN1A	Cyclin dependent kinase inhibitor 1A
CES1	Carboxylesterase 1
ChIP-seq	Chromatin immunoprecipitation sequencing
CPM	Counts per million
CRISPLD2	Cysteine rich secretory protein LCCL domain containing 2
CXCL5	C-X-C motif chemokine ligand 5
CYP	Cytochrome P450

¹Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Ul. J. Tuwima 10, 10-748, Olsztyn, Poland. ²Department of Forestry and Environmental Engineering, South-Eastern Finland University of Applied Sciences, Kuopio, Finland. ³Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland. ⁴dsm-Firmenich AG, Health, Nutrition and Care (HNC), Kaiseraugst, Switzerland. ⁵Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland. ✉email: c.carlberg@pan.olsztyn.pl

DGKK	Diacylglycerol kinase kappa
DHCR7	7-Dehydrocholesterol reductase
ECM	Extracellular matrix
EDAR	Ectodysplasin A receptor
EPPK1	Epiplakin 1
FAIRE-seq	Formaldehyde-assisted identification of regulatory elements followed by sequencing
FC	Fold change
FDR	False discovery rate
FPR1	Formyl peptide receptor 1
GADD45A	Growth arrest and DNA damage inducible alpha
GC	Gc vitamin D binding protein
GNG11	G protein subunit gamma 11
HGNC	Human genome nomenclature committee
HIF1A	Hypoxia inducible factor 1 subunit alpha
HLA-C	Major histocompatibility complex, class I, C
ITG	Integrin
KEGG	Kyoto Encyclopedia of Genes and Genomes
MDS	Multidimensionality scaling
MPEG1	Macrophage expressed 1
MYL9	Myosin light chain 9
MYLK	Myosin light chain kinase
NAMPT	Nicotinamide phosphoribosyltransferase
NK	Natural killer
PBMC	Peripheral blood mononuclear cell
PER1	Period circadian regulator 1
PFKFB3	6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
RNA-seq	RNA sequencing
S100	S100 calcium binding protein
SLC11A1	Solute carrier family 11 member 1
TGFB1	Transforming growth factor beta induced
THBS1	Thrombospondin 1
TOR	Target of rapamycin
TREM1	Triggering receptor expressed on myeloid cells 1
TSS	Transcription start site
VDR	Vitamin D receptor
VEGFA	Vascular endothelial growth factor A

Vitamin D influences both innate and adaptive immunity, enhancing the body's ability to combat infectious diseases like tuberculosis¹² while preventing immune overreactions associated with autoimmune diseases such as multiple sclerosis¹³, rheumatoid arthritis¹⁴ and type 1 diabetes¹⁵. The biologically inert vitamin D₃ is converted in the liver to 25-hydroxyvitamin D₃ (25(OH)D₃), which, due to its relatively long half-life of two weeks, serves as a biomarker for vitamin D status¹⁶.

The most biologically active vitamin D metabolite is 1 α ,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃), produced from 25(OH)D₃ primarily in the kidneys and also in some immune cells¹⁷. This metabolite functions as a nuclear hormone that binds with high affinity to the vitamin D receptor (VDR), a transcription factor¹⁸. Mechanistically, vitamin D's primary molecular effect involves the ligand-triggered binding of VDR to enhancer regions near the transcription start sites (TSSs) of primary vitamin D target genes. This binding alters chromatin accessibility¹⁹, increases H3K27ac marks at active chromatin regions like enhancers, H3K4me3 marks at TSS regions²⁰, and facilitates DNA looping of enhancer regions to TSS regions²¹. Thus, vitamin D modulates the epigenome and transcriptome of VDR-expressing tissues and cell types^{22,23}.

The genome-wide effects of vitamin D have been comprehensively investigated in the human monocytic leukemia cell line THP-1²⁴. However, these principles need validation under human *in vivo* conditions. PBMCs comprise B cells, CD4⁺ and CD8⁺ T cells, natural killer (NK) cells and monocytes, and can be obtained with minimal harm to human volunteers²⁵. Studies on vitamin D have typically used these cells either *in vitro*, with stimulation by supraphysiological doses of 1,25(OH)₂D₃^{26,27}, or *in vivo* over months of daily vitamin D₃ supplementation^{28,29}. Contrarily, the VitDbol study (NCT02063334, ClinicalTrials.gov)^{30,31} isolated PBMC samples 1 and 2 days^{32,33} after a vitamin D₃ bolus (80,000 IU) supplementation, confirming several well-known *in vitro* vitamin D target genes under *in vivo* conditions and categorizing participants into high, mid, and low responders. This led to the concept of the vitamin D response index³⁴, based on the differential activity of vitamin D target genes in individuals^{35,36}.

In this study, we present the molecular findings from the VitDHiD intervention trial, which follows the VitDbol study design, isolating samples one day after vitamin D₃ bolus supplementation. Transcriptome analysis of PBMCs identified 452 *in vivo* vitamin D target genes and categorized the 25 participants into high (9), mid (12), and low (4) responders, providing insight into the molecular basis of the vitamin D response index.

Material and methods

Sample collection

The VitDHiD intervention study had been performed in May 2018 in Kuopio, Finland. The inclusion criteria of the study were an age between 20 and 60 years and a body mass index (BMI) of 20–25 kg/m². Exclusion criteria were a history of kidney stones, renal failure or dialysis, hypercalcemia, hypo- or hyperparathyroidism, severe liver disease (cirrhosis), sarcoidosis or other granulomatous diseases, such as active chronic tuberculosis or Wegener's granulomatosis. Moreover, further reasons for an exclusion were continuous use of anti-inflammatory medicines and regular use of supplements containing over 800 IU vitamin D₃. The ethics committee of the Northern Savo Hospital District had approved the study protocol (#515/2018). In the context of the study, 25 healthy individuals (12 females and 13 males, age 21–54, BMI 21.4–25.6, Table 1) had been supplemented with a vitamin D₃ bolus of in total 80,000 IU, which was given in two equal doses together with breakfast and lunch. The bolus represents a monthly dose of vitamin D₃, which was applied only once to the volunteers, while other vitamin D interventions, such as ViDA²⁵, used even 100,000 IU vitamin D₃/month over a period of more than 2 years. Thus, 80,000 IU vitamin D₃ used in this study represent a safe dose. Blood was drawn directly before supplementation (d0) and 24 h later (d1) for serum and PBMC isolation. All participants gave written informed consent to participate in the study. All experiments were performed in accordance with relevant guidelines and regulations.

Serum measurements

Serum total 25(OH)D (i.e., 25(OH)D₂ and 25(OH)D₃) levels were measured using high-performance liquid chromatography (1290 Infinity II LC System, Agilent) coupled with mass spectrometry detection (API 4000 LC-MS/MS System, SCIEX)²⁶.

PBMC isolation and culture

Without any *in vitro* culture, PBMCs of the 25 VitDHiD study participants (#1–18, 20–26) were isolated within one hour after the collection of 24 ml peripheral blood at d0 and d1 using Vacutainer CPT Cell Preparation Tubes with sodium citrate (Becton Dickinson) according to manufacturer's instructions. The PBMCs of the 25 study participants have the composition 7.4 ± 2.1% B cells, 24.5 ± 6.4% CD4⁺ T cells, 14.9 ± 4.1% CD8⁺ T cells, 13.8 ± 4.5% NK cells and 38.6 ± 6.5% monocytes, while there are only 0.8 ± 0.4% neutrophils as determined

Participant	Age	Gender	BMI	25OHD (d0)	25OHD (d1)
1	21	F	22.6	49.80	64.06
2	21	F	21.7	26.18	33.51
3	33	M	22.7	22.76	31.23
4	23	F	23.9	34.45	44.54
5	27	M	23.0	30.93	37.14
6	23	F	22.2	31.33	40.48
7	21	M	23.5	44.81	56.04
8	21	F	24.5	38.07	49.53
9	41	M	23.2	32.09	38.28
10	21	M	21.4	21.52	29.86
11	36	F	22.4	19.94	29.46
12	24	M	24.1	34.43	39.57
13	26	M	25.6	24.53	36.57
14	54	M	25.1	47.34	55.46
15	23	F	21.7	25.78	35.17
16	44	M	21.6	18.87	30.68
17	26	F	20.8	15.60	30.59
18	27	F	21.4	34.72	44.11
20	22	F	20.9	25.96	34.70
21	21	F	23.2	26.27	35.66
22	22	F	21.1	40.89	61.16
23	24	M	24.1	18.72	31.73
24	32	M	21.6	26.45	36.15
25	26	M	22.6	22.89	29.54
26	29	M	22.8	32.59	38.58
Mean	27.52		22.7	29.88	39.75
Standard deviation	8.39		1.32	9.15	10.10

Table 1. VitDHiD study participants. Age, gender, BMI and total 25(OH)D serum levels (in ng/ml) at d0 and d1 are indicated. Gene expression changes were investigated *in vivo* for all 25 study participant and *in vitro* for participants #1–12, 23 and 26.

by Cbersort²⁷ (<https://cbersortx.stanford.edu>) from RNA-seq (RNA-sequencing) data. After washing with phosphate-buffered saline, PBMCs were aliquoted (4 million cells/ml) and stored in liquid nitrogen until use for either RNA isolation (d0 and d1) or in vitro cell culture (only d0 samples of the 14 study participants #1–12, 23 and 26). PBMC in vitro culture was performed in RPMI 1640 medium (0.5 million cells/ml) supplemented with 10% charcoal-stripped fetal calf serum, 2 mM L-glutamine, 0.1 mg/ml streptomycin and 100 U/ml penicillin. The cells were exposed for 24 h to either 1,25(OH)₂D₃ (10 nM, Sigma-Aldrich) or solvent (0.1% EtOH) at 37 °C in a humidified 95% air/5% CO₂ incubator.

RNA-seq analysis

Total RNA was extracted from PBMCs using the High Pure RNA Isolation kit (Roche) following the manufacturer's protocol. RNA quality was assessed on an Agilent Bioanalyzer and library preparation was performed after rRNA depletion applying kits and protocols from New England Biolabs. RNA-seq libraries were sequenced at 75 bp read length on a NextSeq500 system (Illumina) using standard protocols at the Gene Core of the EMBL (Heidelberg, Germany). The quality of the files was evaluated using FastQC (version 0.11.9, www.bioinformatics.babraham.ac.uk/projects/fastqc) (Table S1). Read alignment to GRCh38 reference genome (Ensembl version 111.38) was done using STAR²⁸ (version 2.7.2a) and quantification was carried out using FeatureCounts²⁹ (version 2.0.1) with default parameters. Human genome nomenclature committee (HGNC) gene symbols of all datasets were brought up to date using the R package HGNC helper (version 0.8.1, <https://CRAN.R-project.org/package=HGNCHelper>) and annotated with gene identifiers, description, genomic location and biotype by accessing the Ensembl database (release 100) via the R package BiomaRt³⁰ (version 2.42.1). Entrez Gene identifiers were added with the R package org.Hs.gene.db (version 3.10.0) and incomplete mappings of target genes manually retrieved from NCBI (www.ncbi.nlm.nih.gov/home/genomes). Genes missing genomic position information or being mitochondrially encoded were removed. In this study, we focused exclusively on protein coding genes.

Differential gene expression analysis

Differential gene expression analysis was computed in R (version 4.3.1) in the MacOS 13 (Ventura) operating system using the tool EdgeR³¹ (version 4.0.14). The analysis focused on 19,287 protein-coding genes, in order to mitigate transcriptional noise expected by non-coding genes. Read counts were normalized for differences in library size to counts per million (CPM). Genes with very low expression were filtered out by applying the function *FilterByExpr()*, in order to mitigate the multiple testing problem and to not interfere with the statistical approximations of the EdgeR pipeline. After filtering, library sizes were recomputed and trimmed mean of M-value normalization was applied. The transcriptome data structure was explored via the method MDS (multidimensionality scaling) (Fig. S1), which was computed via EdgeR's function *plotMDS()*, in which distances approximate the typical log₂ fold change (FC) between the samples. This distance was calculated as the root mean square deviation (Euclidean distance) of the log₂FC values of genes with significant (false discovery rate (FDR) < 0.05) changes after vitamin D₃ supplementation or 1,25(OH)₂D₃ treatment. The gene-wise statistical test for differential expression was computed using the generalized linear model quasi-likelihood pipeline³². Trended negative binomial dispersion estimate was calculated using the method CoxReid profile-adjusted likelihood and together with empirical Bayes-moderated quasi-likelihood gene-wise dispersion estimates used for generalized linear model fitting. The empirical Bayes shrinkage was robustified against outlier dispersions as recommended³³. The *glmTreat* approach was used to test for differential expression relative to absolute log₂FC > 1 for in vitro experiments. Genes with a Benjamini–Hochberg corrected p-value, i.e., FDR < 0.05, and a total trimmed mean of M value-normalized logCPM > log₂10 (average gene expression level for all samples) were considered as significant vitamin D target genes (Tables S2 and S3). Venn diagrams were created via the webtool *jvenn*³⁴ (<http://jvenn.toulouse.inrae.fr/app>).

Characterization of VDR enhancers

The genomic regions of key in vivo vitamin D target genes were investigated for VDR binding enhancers by using epigenome-wide data from THP-1 cells, which had been stimulated for 24 h with 10 nM 1,25(OH)₂D₃ or solvent (0.1% EtOH). These were ChIP-seq (chromatin immunoprecipitation sequencing) data of (i) the histone marker of active promoter regions, H3K4me3³¹, (ii) the histone marker of active chromatin, H3K27ac³¹ and (iii) VDR³⁴ as well as FAIRE-seq (formaldehyde-assisted identification of regulatory elements followed by sequencing) data³⁵. Comparable epigenetic data for PBMCs are not publicly available. However, a comparison of the VDR cisrome of 6 different in vitro models indicated that there is large overlap in the genome-wide VDR binding pattern of cells deriving from the same lineage³⁵. With help of the IGV-browser³⁶ these data sets were displayed for the genomic regions of the vitamin D target genes and VDR enhancers showing responsiveness to 1,25(OH)₂D₃ were marked. One Mb of the genomic region up- and downstream of the target gene's TSS were screened but only essential regions are shown.

Relating vitamin D target genes to pathways

Gene enrichment analysis was performed via the webtool EnrichR³⁷ (<https://maayanlab.cloud/Enrichr>). All genes/proteins of the indicated pathways were inspected for their function, such as type of encoded protein and its primary biological function, with the help of databases like Human Protein Atlas (www.proteinatlas.org) and GeneCards (www.genecards.org). The latter two databases were also used for the categorization of proteins encoded by vitamin D target genes. The network of interacting genes was plotted according to rules of KEGG (Kyoto Encyclopedia of Genes and Genomes)³⁸ signaling pathway diagrams. In a pathway a set of genes (nodes) was connected, in order to reflect the functional relationships between them. Arrows connecting one node with another indicate positive downstream effects, while lines ending with a shorter perpendicular line

display inhibitory effects. The nodes were colored based on the logFC in gene expression between d1 and d0 (downregulation in red, upregulation in green). White nodes indicate genes expressed in PBMCs but not being significantly regulated by vitamin D₃ bolus supplementation. The arrow next to the gene name in each colored node indicates the overall net effect of vitamin D₃ supplementation on the expression of the respective gene.

Classification of study participants

The classification of participants into high, mid or low responders was performed through plotting for each study participant the average absolute logFC values of all 452 *in vivo* vitamin D target genes (or subsets of them) (Table S2) over changes of total 25(OH)D concentrations (d1/d0) (Table 1) of the respective person. A trendline passing through 1/0 divided the plot and the orthogonal distance of each data point from the line was determined. These distances were ranked in a descending order and the algorithm *K-means* (KMeans function from sklearn, cluster library (<https://pypl.org/project/scikit-learn/>)) was used for clustering the individuals into three groups (Fig. S2A). High responders are found at the upper left of the plot (green), mid responders around the trendline (yellow) and low responder at the lower right corner (red). The distance from the trendline is proportional to the vitamin D response index, which negatively correlates very well (-0.94 (Pearson correlation test), P -value 2.1×10^{-12} (paired *T*-test)) with the ratio of the vitamin D status at d1 and d0 (Fig. S2B), but less well (0.51 , P -value 0.01) with the vitamin D status at d0 (Fig. S2C). In contrast, the vitamin D response index does not correlate with the BMI (0.24 , P -value 0.26) of the study participants (Fig. S2D) or their age (0.04 , P -value 0.84) (Fig. S2E). Please note that our segregation approach is not a linear regression analysis.

Results

The VitDHiD intervention trial

The VitDHiD intervention study had been designed as a human investigation, in order to study the effects of vitamin D₃ supplementation on gene expression in human immune cells. In order to allow direct comparisons, the *in vivo* gene expression assessment was as similar as possible to that of *in vitro* cell culture experiments. In the context of the VitDHiD trial, 25 healthy individuals were supplemented with a vitamin D₃ bolus (80,000 IU) and blood samples were taken directly before supplementation (d0) and 24 h later (d1). This protocol follows our previous trial VitDBoL^{26,27}. Total 25(OH)D serum levels indicated that the basal vitamin D status of the study participants ranged from 15.6 to 49.8 ng/ml and increased through vitamin D₃ bolus supplementation within 24 h by 14.9 – 96.1% (Table 1). Thus, a single vitamin D₃ dose of 80,000 IU led within one day to a significant (P -value 1.48×10^{-13} , paired *T*-test) increase of the vitamin D status of all study participants into vitamin D sufficiency (29.5 – 64.1 ng/ml).

VitDHiD identified *in vivo* vitamin D target genes

RNA-seq analysis was executed for the 50 *in vivo* PBMC samples as well as for the 28 *in vitro* samples. The inspection of the MDS plot of the *in vivo* data (Fig. S1A) displayed for each of the 25 individuals in dimension 1 the d0 data point left of the respective d1 data point. The MDS plot of the *in vitro* data (Fig. S1B) clearly distinguished in dimension 1 the solvent samples (left) from the 1,25(OH)₂D₃-treated samples (right). Differential gene expression of the total set of 9374 expressed protein coding genes (average normalized expression of all 50 samples > 10) indicated 452 *in vivo* vitamin D target genes ($FDR < 0.05$ and $\log_2 CPM > \log_{10} 10$) (Fig. S3A, Table S2), of which 181 are upregulated and 271 downregulated. For comparison, the 9628 commonly expressed protein coding genes in the 28 *in vitro* samples revealed 2661 target genes ($FDR < 0.05$) (Fig. S3B, Table S3). However, when the additional threshold of absolute logFC > 1 is applied, only 230 *in vitro* vitamin D target genes are remaining. From these *in vitro* target genes, 113 are upregulated and 117 are downregulated.

The list of 452 *in vivo* vitamin D target genes was compared with different *in vitro* datasets (all absolute logFC > 1, treatment for 24 h with 10 nM 1,25(OH)₂D₃), which had been obtained from the same group of 25 VitDHiD study participants. We observed an overlap with 36 of 230 target genes of the 14 participants investigated only once (Fig. S3C), with 96 of 918 target genes of the five individuals (#5, 9, 12, 13 and 14) tested in triplicate²⁸ as well as with 65 of 466 target genes²⁷ and with 89 of 634 target genes²⁹ from a single person (#14) studied in triplicate (Fig. 1A). In total, 138 *in vivo* vitamin D target genes could be verified by *in vitro* data.

In summary, the setup of the VitDHiD study allowed the identification of 452 genes that within 24 h respond under *in vivo* conditions in PBMCs to vitamin D₃ supplementation. Some 30% of these *in vivo* vitamin D target genes could be confirmed under *in vitro* conditions using cells of the same cohort of participants.

In vivo function of vitamin D based on target gene subsets

The 16 genes *ALOX5* (arachidonate 5-lipoxygenase), *ANPEP* (alanyl aminopeptidase, membrane), *AQP9* (aquaporin 9), *ASAP2* (ArfGAP with SH3 domain, ankyrin repeat and PH domain 2), *CD300E* (CD300E molecule), *CD93*, *CRISPLD2* (cysteine rich secretory protein LCCL domain containing 2), *CXCL5* (C-X-C motif chemokine ligand 5), *FPRI* (formyl peptide receptor 1), *MPEG1* (macrophage expressed 1), *S100A12* (S100 calcium binding protein A12), *S100A9*, *SLC11A1* (solute carrier family 11 member 1), *TGFBI* (transforming growth factor beta induced), *THBS1* (thrombospondin 1) and *TREM1* (triggering receptor expressed on myeloid cells 1) are vitamin D targets not only under *in vivo* conditions but also in all four *in vitro* studies (Fig. 1A). The genes *ASAP2* and *CXCL5* are upregulated by vitamin D, while the remaining 14 genes are downregulated. A screening for VDR-bound enhancers using published ChIP-seq and FAIRE-seq datasets from THP-1 cells identified for all 16 genes one to three VDR-binding enhancers in a distance ranking from very close to the TSS (*ANPEP*, *FPRI*, *S100A9* and *TREM1*) up to 246 kb apart (*CD93*) from the TSS (Fig. S4). Accordingly, all 16 genes seem to be directly regulated by ligand-triggered VDR, but they can be distinguished based on the intensity of the VDR ChIP-seq peaks into strong binders (*CD93*, *ASAP2*, *CRISPLD2*, *TREM1* and *AQP9*, Fig. S4A), medium binders (*ALOX5*,

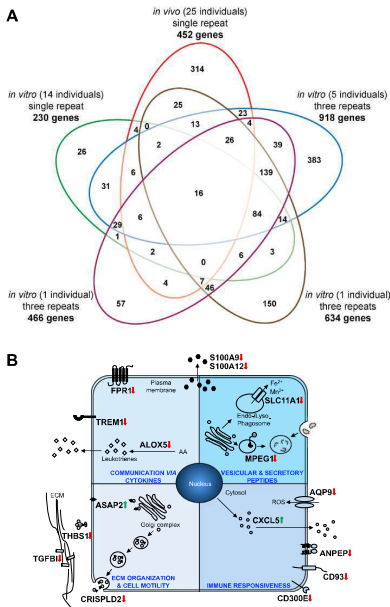


Figure 1. In vivo and in vitro vitamin D target genes. A Venn diagram was used, in order to display the overlap between the 452 in vivo vitamin D target genes identified in PBMCs isolated from 25 participants of the VitDHiD study and in vitro target genes found by different approaches from the pool of the same individuals (A). The identity of the overlapping genes can be obtained from Table S2. A schematic model indicates the location (within or outside the cell) and function (sorted into four categories) of the proteins encoded by the 16 vitamin D target genes that were overlapping between in vivo and in vitro datasets (B). Arrows indicate the direction of regulation of the respective genes/proteins.

SLC11A1, *TGFBI*, *THBS1* and *ANPEP*, Fig. S4B) and weak binders (*S100A9*, *S100A12*, *FPR1*, *CD300E*, *MPEG1* and *CXCL5*, Fig. S4C). As expected for immune cells like PBMCs, the functions of the proteins encoded by the 16

genes relates to immunity, such as “communication via cytokines”, “vesicular and secretory peptides”, “immune responsiveness” and “extracellular matrix (ECM) organization and cell motility” (Fig. 1B).

Another interesting subset of the 452 *in vivo* vitamin D targets are the 91 genes that in average of all 25 study participants have an absolute $\log_{2}FC > 0.25$ (Table S2). These are 66 upregulated and 26 downregulated genes, but only the genes *AQP9*, *ASAP2*, *CXCL5* and *THBS1* are in common with list of 16 confirmed targets (Fig. 1A). Well-known members of the list of the 91 genes are *CD14*, *CDKN1A* (cyclin dependent kinase inhibitor 1A), *GADD45A* (growth arrest and DNA damage inducible alpha), *HIF1A* (hypoxia inducible factor 1 subunit alpha), *NAMPT* (nicotinamide phosphoribosyltransferase), *PER1* (period circadian regulator 1) and *VEGFA* (vascular endothelial growth factor A). These candidates suggest an *in vivo* role of vitamin D in response to bacteria, cell cycle regulation, energy metabolism, circadian rhythms as well as angiogenesis. However, gene enrichment analysis via EnrichR suggested most reliable results from the pathway database Wikipathways⁴⁰ (www.wikipathways.org) (Table 2). As expected, FDR ranking of the pathways identified the VDR pathway (WP2877) as most significant followed by the focal adhesion PI3K (phosphatidylinositol-4,5-bisphosphate 3-kinase)-AKT (AKT murine thymoma viral oncogene homolog)-TOR (target of rapamycin) signaling pathway (WP3932) and the regular focal adhesion pathway (WP306). Taken the latter two pathways together, 11 of the 91 genes (*CDKN1A*, *GNG11* (G protein subunit gamma 11), *HIF1A*, *ITGA2B* (integrin subunit alpha 2B), *ITGB3*, *ITGB5*, *MYL9* (myosin light chain 9), *MYLK* (myosin light chain kinase), *PFKFB3* (6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3), *THBS1* and *VEGFA*), are involved in the process of focal adhesion.

Taken together, the 16 *in vivo* vitamin D target genes that were confirmed by all *in vitro* PBMC datasets can be categorized into (i) communication via cytokines, (ii) vesicular and secretory peptides, (iii) immune responsiveness as well as (iv) ECM organization and cell motility. In contrast, gene enrichment analysis of the 91 most responsive *in vivo* vitamin D target genes indicated that the process of focal adhesion is the major pathway reflecting the *in vivo* function of vitamin D in immune cells of healthy individuals.

Vitamin D upregulates genes related to focal adhesion in healthy individuals

In order to visualize the effect of vitamin D, supplementation on genes related to focal adhesion, we displayed the pathway following the rules of KEGG and color-coded the nodes representing up- and downregulated *in vivo* vitamin D target genes (Fig. 2A). From the 11 vitamin D-triggered genes, *GNG11*, *ITGA2B*, *ITGB3*, *ITGB5*, *MYL9* and *MYLK* are upregulated and they promote focal adhesion. In contrast, the genes *CDKN1A*, *HIF1A*, *PFKFB3*, *THBS1* and *VEGFA* are downregulated by vitamin D, but in case of *CDKN1A* this leads to a supportive effect on the pathway. Thus, in average of 25 study participants, the response of 7 of 11 *in vivo* vitamin D target genes activates focal adhesion.

The participant-specific regulation of the 11 target genes monitors that there is a large interindividual variation of the effect of vitamin D₃ supplementation on the focal adhesion pathway (Fig. 2B). For example, in PBMCs of individuals #2, 12, 18 and 22 the genes *MYL9*, *GNG11*, *ITGB3*, *ITGA2B*, *MYLK* and *ITGB5* (in the order of their responsiveness) are upregulated, while in other participants (#4, 5, 8 and 26) the same genes are downregulated. In contrast, the *THBS1* gene is significantly downregulated in individuals #7, 12, 14 and 21 but upregulated in persons #2, 6, 11 and 18. Thus, in comparison of the study participants, *in vivo* vitamin D target genes can vary both in the intensity of their response as well as in the direction of their regulation.

In summary, on average of the investigated healthy subjects, genes supporting the process of focal adhesion are more upregulated through vitamin D₃ supplementation than downregulated. However, the study participants differ significant in the response of their *in vivo* vitamin D target genes, which in part has opposite effects on pathway regulation.

Segregation of study participants into high, mid and low vitamin D responders

In order to evaluate the vitamin D responsiveness of the 25 individuals, we plotted for each person the average absolute $\log_{2}FC$ of all 452 *in vivo* vitamin D target genes over the ratio of the vitamin D status at d1 compared to that at d0 (Fig. 3A). *K-means* clustering was performed, in order to segregate the individuals into 9 high (green), 12 mid (yellow) and 4 low (red) responders (Fig. S2A). When the list of investigated genes was reduced to 91 targets, basically the same result was obtained, but individuals #7 and 26 were classified only as mid responders (Fig. 3B). Furthermore, when the gene list was restricted to the 16 *in vivo* vitamin D targets that were confirmed by all 4 *in vitro* datasets, individual #2 was graded as mid responder but participant #21 as additional high

Term	Overlap	FDR	Genes
VDR pathway WP2877	8/186	0.00000000046	<i>CDKN1A</i> ; <i>GADD45A</i> ; <i>TNFSF4</i> ; <i>TNFAIP3</i> ; <i>ASAP2</i> ; <i>CD14</i> ; <i>HIF1A</i> ; <i>IRF6</i>
Focal adhesion PI3K-AKT-TOR signaling pathway WP3932	9/302	0.00105169096	<i>CDKN1A</i> ; <i>PFKFB3</i> ; <i>ITGB5</i> ; <i>ITGB3</i> ; <i>ITGA2B</i> ; <i>HIF1A</i> ; <i>THBS1</i> ; <i>GNG11</i> ; <i>VEGFA</i>
Focal adhesion WP306	7/199	0.00253360440	<i>ITGB5</i> ; <i>ITGB3</i> ; <i>ITGA2B</i> ; <i>MYL9</i> ; <i>THBS1</i> ; <i>MYLK</i> ; <i>VEGFA</i>
Hepatitis C and hepatocellular carcinoma WP3646	4/49	0.00397767602	<i>CDKN1A</i> ; <i>CTTN</i> ; <i>HIF1A</i> ; <i>VEGFA</i>
lncRNA-mediated mechanisms of therapeutic resistance WP3672	2/6	0.0132946032	<i>CDKN1A</i> ; <i>HIF1A</i>

Table 2. Gene enrichment analysis. The webtool EnrichR was used for enrichment analysis of the 91 *in vivo* vitamin D target genes with a $\log_{2}FC > 0.25$ and provided most meaningful results with Wikipathways. The top 5 results (ranked by FDR) are shown.

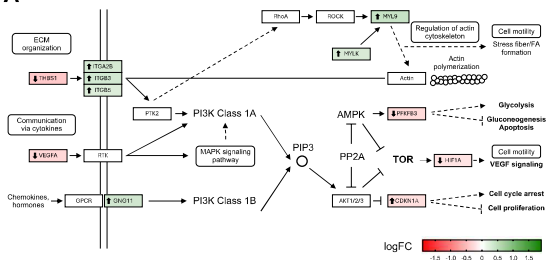
A**B**

Figure 2. Regulation of genes related to focal adhesion by vitamin D₃ supplementation. Representation of the focal adhesion pathway following the design of KEGG (A). Upregulated vitamin D target genes are labeled green and downregulated red. Color intensity is proportional to logFC of gene expression between d1 and d0. A heatmap following the same color scheme indicates the regulation of 11 in vivo vitamin D target genes involved in the pathway for each of the 25 study participants (B). The individuals are ranked by sum of regulation of all 11 genes and the genes are sorted by descending intensity of upregulation.

responder (Fig. 3C). We made similar observations, when we used other subsets of 452 in vivo vitamin D target genes, such as the 181 upregulated genes, the 271 downregulated genes, the 138 genes that were confirmed by at least one in vitro dataset or the 11 genes involved in focal adhesion (Fig. S5). Thus, for some individuals in the

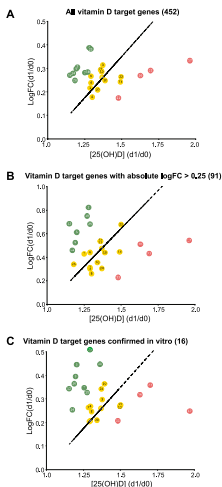


Figure 3. Segregation of VitDHiD study participants. The segregation of the 25 individuals into high (green), mid (yellow) and low (red) responders was based on all 452 in vivo vitamin D target genes (A), a subset of 91 genes with a $\log FC > 0.25$ (B) and 11 genes involved in focal adhesion (C). Numbers indicate the individuals listed in Table 1.

border of high to mid as well as mid to low responders the choice of the list of vitamin D target genes is critical for their classification. However, individuals #9, 12, 14 and 25 were always found as high responders, #4, 6, 8, 10, 15, 20 and 24 constantly as mid responders and #11, 16, 17 and 23 each time as low responders.

Taken together, the responsiveness of in vivo vitamin D target genes can be used to segregate the VitDHiD study participants into high, mid and low responders. For some 40% of individuals that are at the border between high and mid or mid and low responders, the choice of target genes subset is critical for their exact classification.

Molecular distinction between high and low vitamin D responders

The observation of individual-specific responsiveness to vitamin D_3 supplementation raises the question of its molecular explanation. Therefore, we compared the gene expression pattern of the 4 stable high responders (#9, 12, 14 and 25) with that of the 4 stable low responders (#11, 16, 17 and 23) at both d0 and d1. Before supplementation 588 genes were significantly ($FDR < 0.05$) different in their expression when comparing high with low responders (Table S4). One day of a vitamin D_3 bolus reduced the total number of differently expressed genes to 265 (Table S5). Interestingly, only 107 genes of both lists are identical, i.e., vitamin D_3 supplementation equalized 481 genes between high and low responders but additional 158 genes became differently expressed between the two groups (Fig. 4A). When we restrict the gene lists to an absolute $\log FC$ larger than 1, only 15 genes are differently expressed between high and low responders at d0 and even only 12 genes at d1 (Fig. 4B). Interestingly, from the five overlapping genes *HLA-C*, *DGKK* (diacylglycerol kinase kappa) and *CES1* (carboxylesterase

1) were higher expressed in high responders, while *EDAR* (ectodysplasin A receptor) and *EPPK1* (epiplakin 1) showed higher expression in low responders. The most remarkable difference in the gene expression profile of high and low responders shows the *HLA-C* gene, which is in high responders on an average more than 4-times higher expressed than in low responders. Interestingly, *HLA-C* has been reported as a vitamin D target gene¹⁷, which has a strong VDR-binding enhancer 286 kb upstream of its TSS (Fig. 4C).

In summary, the gene expression profile of PBMCs from high and low vitamin D responders differ in nearly 600 genes, but vitamin D₃ supplementation is cutting the number of differently expressed genes to a half. The vitamin D target gene *HLA-C* shows the most prominent differences between high and low responders.

Discussion

The main goal of this study was to obtain a better understanding of the gene regulatory effects of vitamin D₃ supplementation on immune cells of healthy individuals. With the exception of a few scenarios of non-genomic rapid signaling⁴¹, vitamin D acts via its nuclear receptor VDR. Therefore, a central question is, which genes show in response to vitamin D stimulation a statistically significant change in their expression. In average of the 25 VitDHiD participants, this resulted in 452 candidate genes, 138 of which could be confirmed by four in vitro studies with PBMCs of the same individuals and the identical treatment time (24 h). Compared with other studies^{16,19,42}, a validation rate of 30.5% of the in vivo vitamin D target genes is a good score. However, if we had selected an even larger in vitro cohort, such as all 25 VitDHiD participants, we would probably have obtained an even higher validation rate. The in vitro studies had the clear advantage to use solvent treatment as a negative control and to apply with 10 nM 1,25(OH)₂D₃ a more than 100-times higher concentrated stimulus as could be reached in vivo by a bolus with 80,000 IU vitamin D₃⁴³. For the same reasons in vitro experiments show far higher inducibilities of vitamin D target genes than observed in the in vivo setting.

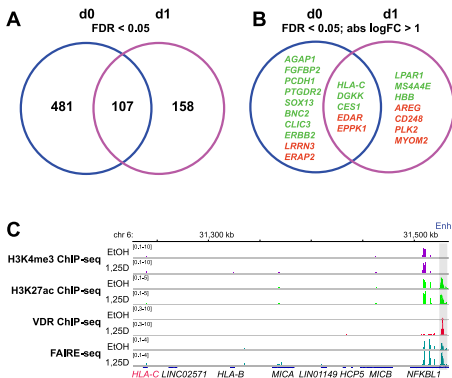


Figure 4. Gene expression profile of high and low responders. Venn diagrams display the number (A) and name (B) of significantly (FDR < 0.05) differently expressed genes in high and low responders before (d0) and after (d1) vitamin D₃ supplementation. In (B) only genes with an absolute logFC > 1 are shown. Genes being expressed higher in high responders are indicated in green, while red indicates higher expression in low responders. The IGV browser was used to visualize ChIP-seq results for H3K4me3 (purple)⁴¹, H3K27ac (green)⁴¹ and VDR (red)⁴⁴ as well as FAIRE-seq data (turquoise)⁴⁵ obtained in THP-1 cells that had been treated for 24 h with solvent (EIOH) or 1,25(OH)₂D₃ (1,25D) (C). The peak tracks display merged data from three biological repeats. A VDR-binding enhancer is shaded in grey and gene structures are shown in blue.

Due to the known large interindividual variation in vitamin D target genes³⁸, only a shortlist of 16 genes responded significantly to vitamin D both in vivo as well as in all four in vitro studies. Interestingly, all 16 genes have at least one VDR-binding enhancer in the vicinity of 0–246 kb to their TSS region, i.e., they are all primary vitamin D targets⁴⁵. Although the 16 genes represent rather divergent molecular processes, such as communication via cytokines, vesicular and secretory peptides, immune responsiveness as well as ECM organization and cell motility, they all relate to immunity. A more focused gene enrichment analysis on the basis of the 91 genes with a $\log_{2}FC > 0.25$ highlighted 11 targets being involved in the process of focal adhesion. Thus, this study concludes that a key physiological effect of vitamin D₃ supplementation of healthy individuals may be the promotion of focal adhesion. The main role of focal adhesion is the communication of a cell with the ECM, i.e., with its direct environment⁴⁴. Integrins, such as encoded by the upregulated in vivo vitamin D target genes *ITGA2B*, *ITGB3* and *ITGB5*, form the core of large adhesion complexes that connect the cytoskeleton with proteins of the ECM. In this way, cells can adapt their shape in response to the mechanical structure of their environment⁴⁵. For example, this is important for the movement of monocytes from the blood stream to inflamed tissues and their differentiation to macrophages⁴⁶. Since monocytes are the most vitamin D responsive cell type of PBMCs¹⁵, the upregulation of genes involved in the process of focal adhesion may prepare healthy individuals for a possible attack by pathogenic microbes. This enhances the potency of innate immunity. Interestingly, in the context of the vitamin D-triggered differentiation of mesenchymal stem cells into osteoblasts, ITGAs and ITGBs have already been shown to be upregulated in vitro by $1,25(\text{OH})_2\text{D}_3$ ⁴⁷.

The individual-specific analysis of the effects of vitamin D₃ supplementation on the 11 in vivo vitamin D target genes, which are involved in the pathway of focal adhesion, indicated larger variations between the study participants. In some individuals vitamin D seems to inhibit focal adhesion suggesting that these persons' immune system is less prepared to respond to a possible infection. Accordingly, the classification of the VitDHiD participants based on the responsiveness of all 452 in vivo vitamin D target genes in relation to the increase of the vitamin D status after vitamin D₃ bolus supplementation allowed a distinction into 9 high, 12 mid and 4 low responders. The use of subsets of these target genes, such as the 91 genes with a $\log_{2}FC > 0.25$ or the 16 genes being confirmed by all four in vitro studies, resulted in basically the same classification but indicated some deviations for those individuals who are at the border between high and mid or mid and low responders. Nevertheless, 4 high, 7 mid and 4 low responders turned out to be stable in their classification irrespective of the used subset of in vivo vitamin D target genes.

We described the concept of the vitamin D response index already for other Finnish cohorts and found that some 25% of the participants are low responders^{18,22,23}. The 4 stable low responders of this study suggests a rate of only 16% in the Finnish population. However, some gene lists, such as those of the 138 in vitro confirmed candidates or the 271 downregulated targets, suggested 6 low responders, which stands for a rate of 24%, i.e., a rate very comparable to previous studies. We propose that low vitamin D responders should take special care on their vitamin D₃ supplementation, in order to prevent diseases, such as musculoskeletal disorders like osteoporosis⁴⁸ and sarcopenia⁴⁹, autoimmune diseases like multiple sclerosis⁵⁰ and different types of cancer like colon cancer⁵¹. Since a daily vitamin D₃ supplementation up to 4000 IU is considered safe⁵², the percentage of low responders in the population should rather be overestimated than underestimated.

The comparison of the gene expression profile of stable high and low responders provided some insight into the molecular basis of the vitamin D response index. Interestingly, at basal conditions the transcriptome of PBMCs from high and low responders significantly ($FDR < 0.05$, $\text{abs log}_{2}FC > 1$ and $\log_{2}CPM > \log_{2}10$) differed in the expression strength of 588 genes, i.e., in some 5% of all expressed genes. Gene expression is primarily based on the epigenetic stage of the respective cell, since effective gene regulation requires accessible chromatin at enhancers and TSS regions⁵³. Therefore, the difference in the gene expression between high and low responders is rather based on epigenetics and signals from the environment than on genetic variations. However, genome-wide association studies indicated that key genes of vitamin D endocrinology, such as *DHCR7* (7-dehydrocholesterol reductase), *GC* (GC vitamin D binding protein), *CYP2R1* (cytochrome P450 family 2 subfamily R member 1) and *CYP24A1*, significantly contribute to $25(\text{OH})\text{D}_3$ serum levels⁵⁴. Therefore, in analogy to the trait "vitamin D status", it is likely that there is also a genetic contribution for the trait "vitamin D responsiveness". This needs to be further investigated by future studies involving a significantly higher number of participants.

Interestingly, vitamin D₃ bolus supplementation can be considered as a major environmental trigger, since it reduces within one day the difference in the gene expression profile of high and low responders to 265 genes, i.e., to less than half. *HLA-C* stands out from the differently expressed genes, since it is more than 4-times higher expressed in high responders than in low responders. Moreover, it is a vitamin D target gene in PBMCs^{17,55} and has a prominent VDR-binding enhancer 286 kb upstream of its TSS. Importantly, the *HLA* gene cluster in chromosome 6 is known as the genetically most divergent region of the human genome⁵⁶. Therefore, it is possible that single nucleotide variants of the *HLA-C* gene as well as that of other members of the *HLA* gene family, most of which are also vitamin D targets⁵⁷, contribute to the genetic variations that are the basis for interindividual differences in the vitamin D response index.

This study has a few limitations, the major of which is the rather low number of investigated VitDHiD participants, i.e., at present we are lacking statistical power to prove our claims concerning the molecular differences of high and low vitamin D responders. Thus, future studies, such as the presently ongoing VitDPAS trial (NCT06104111), will need to refine the results presented here. This concerns both the epigenetic as well as the genetic contribution to the vitamin D response index. Another limitation of this study is that PBMCs represent a mixture of at least five types of primary immune cells, i.e., the effects of vitamin D described here cannot be assigned to one single cell type. However, from other studies we know that monocytes are most responsive to vitamin D¹⁵.

In conclusion, in this study we present 452 in vivo vitamin D target genes that show a higher overlap with genes identified by in vitro studies than previous reports. Moreover, the process of focal adhesion, which is

central for the migration and differentiation of monocytes was identified as a major pathway being upregulated by vitamin D₃ supplementation in healthy individuals. This sharpens the response of the innate immune system against possible microbial infections. Furthermore, this study provides some insight into the molecular difference of high and low vitamin D responders. On one hand high responders seem to differ from low responders primarily by epigenetically and environmentally driven differences in their transcriptome. On the other hand, variations of some of the differentially expressed genes, such as *HLA-C* and possibly other members of this gene family, may represent the genetic contribution of the trait "vitamin D response index".

Data availability

Fastq files of the raw data can be found at Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) with accession numbers GSE260981 (50 in vivo samples) and GSE261699 (28 in vitro samples).

Received: 10 April 2024; Accepted: 26 July 2024

Published online: 30 July 2024

References

- Huang, S. J. et al. Vitamin D deficiency and the risk of tuberculosis: A meta-analysis. *Drug Des. Dev. Ther.* **11**, 91–102. <https://doi.org/10.2147/DDDT.S79070> (2017).
- Chun, R. F., Liu, P. T., Modlin, R. L., Adams, J. S. & Hewison, M. Impact of vitamin D on immune function: Lessons learned from genome-wide analysis. *Front. Physiol.* **5**, 151. <https://doi.org/10.3389/fphys.2014.00151> (2014).
- Stintzel, M. B., Rametta, M. & Roder, A. T. Vitamin D and multiple sclerosis: A comprehensive review. *Neurol. Ther.* **7**, 59–85. <https://doi.org/10.1007/s40120-017-0086-4> (2018).
- Lu, M., McComish, B. J., Burdon, K. P., Taylor, B. V. & Körner, H. The association between vitamin D and multiple sclerosis risk: 1,25(OH)₂D₃ induces super-enhancers bound by VDR. *Front. Immunol.* **10**, 488. <https://doi.org/10.3389/fimmu.2019.00488> (2019).
- Jeffery, L. E., Raza, K. & Hewison, M. Vitamin D in rheumatoid arthritis-towards clinical application. *Nat. Rev. Rheumatol.* **12**, 201–210. <https://doi.org/10.1038/nrrheum.2015.140> (2016).
- Infante, M. et al. Influence of vitamin D on islet autoimmunity and beta-cell function in type 1 diabetes. *Nutrients* **11**, 2185. <https://doi.org/10.3390/nu11092185> (2019).
- Hollis, B. W. Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: Implications for establishing a new effective dietary intake recommendation for vitamin D. *J. Nutr.* **135**, 317–322 (2005).
- Whitfield, G. K. et al. Cloning of a functional vitamin D receptor from the lamprey (*Petromyzon marinus*), an ancient vertebrate lacking a calcified skeleton and teeth. *Endocrinology* **144**, 2704–2716. <https://doi.org/10.1210/en.2002-221101> (2003).
- Hausler, M. R. et al. Vitamin D receptor: Molecular signaling and actions of nutritional ligands in disease prevention. *Nutr. Rev.* **66**, 598–112. <https://doi.org/10.1111/1753-4887.2008.00093.x> (2008).
- 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. <https://doi.org/10.1093/nar/gkv1519> (2016).
- Nurminen, V., Neme, A., Seuter, S. & Carlberg, C. The impact of the vitamin D-modulated epigenome on VDR target gene regulation. *Biochim. Biophys. Acta* **1861**, 697–905. <https://doi.org/10.1016/j.bbagen.2018.05.006> (2018).
- Rivera, C. M. & Ren, B. Mapping human epigenomes. *Cell* **155**, 39–55. <https://doi.org/10.1016/j.cell.2013.09.011> (2013).
- Campbell, M. J. Vitamin D and the RNA transcriptome: More than mRNA regulation. *Front. Physiol.* **5**, 181. <https://doi.org/10.3389/fphys.2014.00181> (2014).
- Carlberg, C. Vitamin D genomics: From in vitro to in vivo. *Front. Endocrinol.* **9**, 250. <https://doi.org/10.3389/fendo.2018.00250> (2018).
- Carlberg, C. Vitamin D signaling in the context of innate immunity: Focus on human monocytes. *Front. Immunol.* **10**, 2211. <https://doi.org/10.3389/fimmu.2019.02211> (2019).
- Kariuki, S. N. et al. Mapping variation in cellular and transcriptional response to 1,25-dihydroxyvitamin D₃ in peripheral blood mononuclear cells. *PLoS ONE* **11**, e0159779. <https://doi.org/10.1371/journal.pone.0159779> (2016).
- Hanel, A. & Carlberg, C. Time-resolved gene expression analysis monitors the regulation of inflammatory mediators and attenuation of adaptive immune response by vitamin D. *Int. J. Mol. Sci.* **23**, 911. <https://doi.org/10.3390/ijms23020911> (2022).
- Carlberg, C. et al. Primary vitamin D target genes allow a categorization of possible benefits of vitamin D₃ supplementation. *PLoS ONE* **8**, e71042. <https://doi.org/10.1371/journal.pone.0071042> (2013).
- Hosseini-nezhad, A., Spira, A. & Holick, M. F. Influence of vitamin D status and vitamin D₃ supplementation on genome wide expression of white blood cells: A randomized double-blind clinical trial. *PLoS ONE* **8**, e58725. <https://doi.org/10.1371/journal.pone.0058725> (2013).
- Vukic, M. et al. Relevance of vitamin D receptor target genes for monitoring the vitamin D responsiveness of primary human cells. *PLoS ONE* **10**, e0124339. <https://doi.org/10.1371/journal.pone.0124339> (2015).
- Seuter, S. et al. Molecular evaluation of vitamin D responsiveness of healthy young adults. *J. Steroid Biochem. Mol. Biol.* **174**, 314–321 (2017).
- Carlberg, C. & Haq, A. The concept of the personal vitamin D response index. *J. Steroid Biochem. Mol. Biol.* **175**, 12–17. <https://doi.org/10.1016/j.jsbmb.2016.12.011> (2018).
- Seuter, S. et al. Molecular evaluation of vitamin D responsiveness of healthy young adults. *J. Steroid Biochem. Mol. Biol.* **174**, 314–321. <https://doi.org/10.1016/j.jsbmb.2016.06.003> (2017).
- Gospodarska, E., Ghosh Dastidar, R. & Carlberg, C. Intervention approaches in studying the response to vitamin D₃ supplementation. *Nutrients* **15**, 3382. <https://doi.org/10.3390/nu15153382> (2023).
- Scragg, R. et al. Effect of monthly high-dose vitamin D supplementation on cardiovascular disease in the vitamin D assessment study: A randomized clinical trial. *JAMA Cardiol.* **2**, 608–616. <https://doi.org/10.1001/jamacardio.2017.0175> (2017).
- Hanel, A. et al. Common and personal target genes of the micronutrient vitamin D in primary immune cells from human peripheral blood. *Sci. Rep.* **10**, 21051. <https://doi.org/10.1038/s41598-020-78288-0> (2020).
- Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457. <https://doi.org/10.1038/nmeth.3337> (2015).
- Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21. <https://doi.org/10.1093/bioinformatics/btt053> (2013).
- Liao, Y., Smyth, G. K. & Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930. <https://doi.org/10.1093/bioinformatics/btt656> (2014).
- Durbin, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protocols* **4**, 1184–1191. <https://doi.org/10.1038/nprot.2009.97> (2009).
- 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. <https://doi.org/10.1093/bioinformatics/btp166> (2010).

32. Chen, Y., Lun, A. T. & Smyth, G. K. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research* **5**, 1438. <https://doi.org/10.12688/f1000research.9861.2> (2016).
33. Bardou, P., Mariette, J., Escudie, F., Djemiel, C. & Klopp, C. Jvarkit: An interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 293. <https://doi.org/10.1186/1471-2105-15-293> (2014).
34. Neme, A., Seuter, S. & Carlberg, C. Selective regulation of biological processes by vitamin D based on the spatio-temporal cistrome of its receptor. *Biochim. Biophys. Acta* **1860**, 952–962. <https://doi.org/10.1016/j.bbagr.2017.07.002> (2017).
35. Turossi, P., Väisänen, S., Neme, A., Helkkinen, S. & Carlberg, C. Patterns of genome-wide VDR locations. *PLoS ONE* **9**, e96105. <https://doi.org/10.1371/journal.pone.0096105> (2014).
36. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192. <https://doi.org/10.1093/bib/bbs017> (2013).
37. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128. <https://doi.org/10.1186/1471-2105-14-128> (2013).
38. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361. <https://doi.org/10.1093/nar/gkx1092> (2017).
39. Malmberg, H. R., Hanel, A., Taipale, M., Helkkinen, S. & Carlberg, C. Vitamin D treatment sequence is critical for transcriptome modulation of immune-challenged primary human cells. *Front. Immunol.* **12**, 754056. <https://doi.org/10.3389/fimmu.2021.754056> (2021).
40. Pico, A. R. et al. WikiPathways: Pathway editing for the people. *PLoS Biol.* **6**, e184 (2008).
41. Zmijewski, M. A. & Carlberg, C. Vitamin D receptor(s): In the nucleus but also at membranes? *Exp. Dermatol.* **29**, 876–884. <https://doi.org/10.1111/exd.14147> (2020).
42. Yeh, W. Z. et al. The immune cell transcriptome is modulated by vitamin D₃ supplementation in people with a first demyelinating event participating in a randomized placebo-controlled trial. *Clin. Immunol.* **262**, 110183. <https://doi.org/10.1016/j.clim.2024.110183> (2024).
43. Carlberg, C. Vitamin D and its target genes. *Nutrients* **14**, 1354. <https://doi.org/10.3390/nu14071354> (2022).
44. Geiger, B., Spatz, J. P. & Bershadsky, A. D. Environmental sensing through focal adhesions. *Nat. Rev. Mol. Cell Biol.* **10**, 21–33. <https://doi.org/10.1038/nrm2593> (2009).
45. DuFort, C. C., Paszek, M. J. & Weaver, V. M. Balancing forces: Architectural control of mechanotransduction. *Nat. Rev. Mol. Cell Biol.* **12**, 308–319. <https://doi.org/10.1038/nrm3112> (2011).
46. Shi, C. & Pamer, E. G. Monocyte recruitment during infection and inflammation. *Nat. Rev. Immunol.* **11**, 762–774. <https://doi.org/10.1038/nri3070> (2011).
47. Posa, E. et al. Vitamin D promotes MSC osteogenic differentiation stimulating cell adhesion and alphaVbeta3 expression. *Stem Cells Int.* **2018**, 6958713. <https://doi.org/10.1155/2018/6958713> (2018).
48. Sunyecz, J. A. The use of calcium and vitamin D in the management of osteoporosis. *Ther. Clin. Risk Manag.* **4**, 827–836. <https://doi.org/10.2147/tcrm.s3552> (2008).
49. Uchitomi, R., Ouyab, M. & Kamei, Y. Vitamin D and sarcopenia: potential of vitamin D supplementation in sarcopenia prevention and treatment. *Nutrients* **12**, 3189. <https://doi.org/10.3390/nu12103189> (2020).
50. Feige, J. et al. Vitamin D supplementation in multiple sclerosis: a critical analysis of potentials and threats. *Nutrients* **12**, 783. <https://doi.org/10.3390/nu12030783> (2020).
51. Vaughan-Shaw, P. G. et al. The effect of vitamin D supplementation on survival in patients with colorectal cancer: Systematic review and meta-analysis of randomised controlled trials. *Br. J. Cancer* **123**, 1705–1712. <https://doi.org/10.1038/s41416-020-01060-8> (2020).
52. Johnson, K. C. et al. Safety and tolerability of high-dose daily vitamin D₃ supplementation in the vitamin D and type 2 diabetes (D2d) study: a randomized trial in persons with prediabetes. *Eur. J. Clin. Nutr.* **76**, 1117–1124. <https://doi.org/10.1038/s41430-022-01068-8> (2022).
53. Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Heredity (Edinb)* **105**, 4–13. <https://doi.org/10.1038/s41430-020-01068-8> (2010).
54. Wang, T. J. et al. Common genetic determinants of vitamin D insufficiency: A genome-wide association study. *Lancet* **376**, 180–188. [https://doi.org/10.1016/S0140-6736\(10\)60588-0](https://doi.org/10.1016/S0140-6736(10)60588-0) (2010).
55. Neme, A. et al. In vivo transcriptome changes of human white blood cells in response to vitamin D. *J. Steroid Biochem. Mol. Biol.* **188**, 71–76. <https://doi.org/10.1016/j.jsbmb.2018.11.019> (2019).
56. Trowsdale, J. & Knight, J. C. Major histocompatibility complex genomics and human disease. *Annu. Rev. Genom. Hum. Genet.* **14**, 301–323. <https://doi.org/10.1146/annurev-genom.091212-153455> (2013).

Acknowledgements

The authors thank Dr. Antonio Neme and Andrea Hanel for initial data analysis as well as Emmi Hämäläinen, Ella Janzen and Retu Haikonen for PBMC isolation. Kind thanks to the Gene Core Facility at the EMBL in Heidelberg, Germany, for massively parallel sequencing services and the Centre of Informatics Tricity Academic Supercomputer and Network of the Technical University of Gdansk (<https://task.gda.pl/en>) for the use of their supercomputers for sequence alignment.

Author contributions

RGD, JJ and CC performed data analysis, while MM did RNA isolation and RNA-seq library preparation. CC wrote the manuscript, which was reviewed by all authors. The VitDHiD study had been designed and conducted by CC, JKV, TPT and IB.

Funding

This publication is part of a project that has received funding from the European Union's Horizon2020 research and innovation program under grant agreement no. 952601, from the David and Amy Fulton Foundation, Seattle, US and from the National Science Centre (Poland), project number 2023/49/B/NZ9/00402. The dsM-firmenich AG sponsored the VitDHiD study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-68741-9>.

Correspondence and requests for materials should be addressed to C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024



Evaluation of the vitamin D response index in a Saudi cohort

Shareefa A. AlGhamdi^{a,b,c,1,*}, Ranjini Ghosh Dastidar^{d,1}, Maciej Rybiński^d,
Hadeil M. Alsufiani^{a,b,c}, Sawsan O. Khoja^{a,b}, Nusaibah N. Enaibsi^a, Safa F. Saif^a,
Carsten Carlberg^{a,c,1**}

^a Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

^b Vitamin D Pharmacogenomics Research Group, King Abdulaziz University, Jeddah, Saudi Arabia

^c Experimental Biochemistry Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

^d Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn, Poland

^e Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland

ARTICLE INFO

Keywords:

Vitamin D₃ supplementation
Proinflammatory cytokines
Vitamin D target genes
Vitamin D response index
Saudi cohort

ABSTRACT

The concept of the vitamin D response index was developed based on vitamin D intervention studies conducted with Finnish cohorts. In this study, we challenged the concept by performing a single vitamin D₃ bolus (80,000 IU) intervention with a cohort of 100 native Saudis. The change of serum levels of the proinflammatory cytokines interleukin 6, interleukin 8 and tumor necrosis factor measured directly before intervention in comparison to samples taken one and thirty days after vitamin D₃ supplementation were used as biomarkers for distinguishing low, mid and high responders. Interestingly, we identified 39 % of the study participants as low responders. In contrast, when we used in a subset of 37 study participants whole blood expression changes of seven well-known vitamin D target genes one and thirty days after supplementation as alternative biomarkers, only 9 persons (24 %) were identified as low responders. In conclusion, in Saudi Arabia the rate of low vitamin D responders is equal or even higher than that in Finland. Therefore, similar to Nordic countries also in Saudi Arabia appropriate vitamin D₃ supplementation is essential, in order to fulfill the needs of low responders.

1. Introduction

Vitamin D₃ is an important signaling molecule, because its metabolite 1,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃) acts as a high affinity ligand to the transcription factor vitamin D receptor (VDR) (Pike et al., 2017; Zmijewski and Carlberg, 2020). This makes vitamin D a direct regulator of gene expression (Carlberg 2022) and a molecule with a pleiotropic physiological profile (Dimitrov et al., 2021). The most important functions of vitamin D₃ are the regulation of calcium homeostasis (Fleet, 2022), which is essential for appropriate bone formation (van Driel and van Leeuwen, 2023), and the modulation of the immune system, such as boosting innate immunity and preventing overreactions of adaptive immunity (Bishop et al., 2020). Accordingly, insufficient circulating levels of 25-hydroxyvitamin D₃ (25(OH)D₃), which is the most stable vitamin D₃ metabolite and serves as a biomarker for the vitamin D status, are associated with rickets in children,

osteomalacia in adults and osteoporosis in elderly (Uday and Hogler, 2017) as well as with inflammatory and autoimmune diseases, such as rheumatoid arthritis and multiple sclerosis (Harrison et al., 2020).

When 7-dehydrocholesterol, a direct precursor of cholesterol, is exposed in human skin to the ultraviolet-B (UVB) component of sunlight, vitamin D₃ is synthesized in a non-enzymatic reaction (Wacker and Holick, 2013). However, changes in lifestyle like preferential indoor activities as well as living at higher latitudes with large seasonal variations in sun exposure, drastically reduced the endogenous vitamin D₃ production in nearly all human populations (Bendik et al., 2014; Spiro and Buttriss, 2014). This made vitamin D₃ a vitamin that needs to be taken up by diet or directly supplemented via pills. There is an ongoing discussion about the dose of daily vitamin D₃ supplementation as well as on the threshold levels distinguishing vitamin D sufficiency from insufficiency and deficiency (Pilz et al., 2019). In this study, we follow the well-accepted recommendation of the US Endocrine Society that a

* Corresponding author at: Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.

** Corresponding author at: Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn, Poland.

E-mail addresses: snaalghamdi1@kau.edu.sa (S.A. AlGhamdi), c.carlberg@pan.olsztyn.pl (C. Carlberg).

¹ Both authors contributed equally to the manuscript.

25(OH) D_3 level above 75 nM (30 ng/ml) is considered as sufficient, while vitamin D deficiency is defined as a 25(OH) D_3 concentration below 30 nM (12 ng/ml) (Holick et al., 2011).

Studies with Finnish cohorts being supplemented either daily with 1600 or 3200 IU vitamin D_3 (Carlberg et al., 2013) or once with a monthly dose of 80,000 IU suggested that individuals can be segregated into low, mid and high responders to vitamin D. This vitamin D response index concept (Carlberg and Haq, 2018) implies that there should be supplementation with personalized doses of vitamin D_3 , because low responders need a higher daily vitamin D_3 dose than high responders. Determining the vitamin D response index requires at least two samples from the same individual and measurements of circulating vitamin D-triggered proteins, such as the cytokines IL6 (interleukin 6), IL8, TNF (tumor necrosis factor) and the peptide hormone PTH (parathyroid hormone) (Carlberg et al., 2013), or the expression of vitamin D target genes in immune cells, such as PBMCs (peripheral blood mononuclear cells). The molecular basis of the vitamin D response index is not well explored, but it is likely based both on genetic and epigenetic variations of individuals (Hanel and Carlberg, 2020). However, the vitamin D response index is independent from the vitamin D status, since persons with high 25(OH) D_3 serum levels can be low responders, while individuals with a low vitamin D status may be high responders (Carlberg and Haq, 2018).

Recently, we reported on single vitamin D_3 bolus (80,000 IU) supplementation of 50 males (AlGhamdi et al., 2022) and 50 females (Alsuifiani et al., 2022) from Saudi Arabia. We found in both cohorts a significant reduction in the circulating levels of the proinflammatory cytokines IL6, IL8 and TNF. In this study, we performed a meta-analysis of the data from both cohorts and used them for determining the vitamin D response index. Furthermore, we used changes in the expression of vitamin D target genes in a subgroup of 37 study participants for the verification of response index determination.

2. Materials and methods

2.1. Study design and participants

Each of the 50 male and 50 female adults (18–53 years old) had been recruited from the staff of King Abdul Aziz University and King Fahad Medical Research Center. Details of the study have been described previously (AlGhamdi et al., 2022, Alsuifiani et al., 2022). In brief, all participants were native Saudis, free of cancer, liver or kidney diseases and did not take vitamin D supplements within three months before the study. They orally received a single vitamin D_3 (80,000 IU) bolus and gave fasting blood samples directly before supplementation (d0) as well as one and thirty days after (d1 and d30). The study was approved by the ethical committee of the Faculty of Medicine, King Abdulaziz University (ref. no. 30–18) and all participants provided written informed consent. The individuals were considered healthy as per Saudi standards (set by the Saudi Central Board for Accreditation of Healthcare Institutions (CBAHI, <https://portal.cbahi.gov.sa/english/cbahi-standards>) and the Saudi Commission for Health Specialties (SCFHS, <https://scfhs.org.sa/en/>), which defines health as being free of serious ailments or life-long medications.

2.2. Biomarker measurements

Quantitative measurements of serum levels of 25(OH) D_3 were done via the Architect 25-OH vitamin D assay kit (Abbott), which has a detection limit of 4 ng/mL, with an intra-assay coefficient of variation (CV) of approximately 5.0 % and an inter-assay CV of approximately 7.0 %. The cytokines IL6, IL8 and TNF were determined by fully automated ELISA (enzyme-linked immunosorbent assay, Bioss Technology Laboratory). The detection limits for IL6, IL8, and TNF were 2 pg/mL, 1 pg/mL, and 2 pg/mL, respectively, with intra-assay CVs of less than 10 % and inter-assay CVs of less than 12 % as previously described (AlGhamdi

et al., 2022, Alsuifiani et al., 2022) (Table S1).

2.3. Gene expression measurements

RNA was extracted from fresh blood sample by using QIAamp RNA blood mini kit following the instructions of the manufacturer (Qiagen). RNA quantity was determined via a Nanodrop 2000 spectrophotometer and RNA integrity was assessed on an Agilent 2100 bioanalyzer. RNA with RIN > 8 was used for cDNA synthesis through the SuperScript™ IV VILIO™ Master Mix (Invitrogen) according to manufacturer's protocol. For measuring the expression of the vitamin D target genes *ASAP2* (*ArfGAP* with SH3 domain, ankyrin repeat and PH domain 2), *CAMP* (cathelicidin antimicrobial peptide), *CD14* (CD14 molecule), *FBP1* (fructose-bisphosphatase 1), *ITGAM* (integrin subunit alpha M), *LRRC25* (leucine rich repeat containing 25) and *NINJ1* (ninjurin 1) in relation to *RPL13A* (ribosomal protein L13a) housekeeping gene, real-time PCR was performed using the EverGreen Universal qPCR Master Mix (Haven Scientific) and the primer oligonucleotides listed in Table 1. Triplicate reactions were done on an ABI7300 system applying the following protocol: 3 min at 95 °C followed by 40 cycles of 95 °C for 15 s and then 60 °C for 60 s (data collection). To eliminate inter-run variabilities, all assays for a given sample were run on the same plate. PCR product specificity was monitored using post-PCR melt curve analysis. Relative mRNA expression levels were determined using the formula $2^{-\Delta Ct}$, where ΔCt is $Ct(\text{target gene}) - \text{mean of } Ct(\text{reference gene})$.

2.4. Determination of the vitamin D response index

In order to determine the vitamin D response index, we scored all 100 study participants based on the downregulation of proinflammatory cytokines (IL6, IL8, and TNF) at days 1 (d1) and 30 (d30) after a single vitamin D_3 bolus (80,000 IU) supplementation at day 0 (d0). Absolute log fold changes (logFC) in cytokine serum levels between d1 and d0 and between d30 and d0 were compared to changes in 25(OH) D_3 serum concentrations. This scoring system was also applied to a subgroup of 37 individuals for the upregulation of vitamin D target genes (*ASAP2*, *CAMP*, *CD14*, *FBP1*, *ITGAM*, *LRRC25* and *NINJ1*).

For the visualization of the data, we plotted cytokine expression changes against 25(OH) D_3 serum levels and added a trend line to understand the pattern of responses. This line divided the participants into low responders (score 0), mid responders (score 1), and high responders (score 3) based on their relative positions to the line. Further, we utilized the *K-means* clustering algorithm to classify participants into low, mid and high responders based on their cumulative scores. We performed correlation tests to rule out confounding factors such as age, gender and BMI, confirming that the observed changes in cytokine serum levels and target gene expressions were primarily influenced by 25(OH) D_3 serum

Table 1
Oligonucleotide primers used for real-time PCR.

Gene name	Orientation	Oligonucleotide sequence	Amplicon size
ASAP2	Forward	TACGATCTTCAACAGATCAAC	104
	Reverse	CCTCTTTCTGTTCACCTGCA	
CAMP	Forward	GATGCTAACTCTACCGCTCA	61
	Reverse	CTGGTCCCATCATCTGCT	
CD14	Forward	GCGCTGTGTAGGAAGAAG	85
	Reverse	CGGCTCATGGTGATG	
FBP1	Forward	ACTCTGGTCTACGAGGAGAT	73
	Reverse	ACACGAGTCTCACTTTCAT	
ITGAM	Forward	GGTTCACCTCTCCAGTGT	81
	Reverse	ATGACATAAGGTCAAGCTGTT	
LRRC25	Forward	CCCTCCACTCCGACATGTA	83
	Reverse	GAAAGGCTGAGGCCCTGTTC	
NINJ1	Forward	GCGCTGGTGTTCATCATGGT	106
	Reverse	CAGGGTGTTCGCTGTCTACT	
RPL13A	Forward	GCTAAACAGGTACTGCTGGG	99
	Reverse	AGCCGAGTCTTCAACTGTGTT	

levels. For seven participants with dissimilar scores between the two methods, we noted a trend of higher basal IL6 and TNF serum levels but no statistically significant differences were found.

3. Results

In this study, we took advantage of two vitamin D intervention studies that were performed with 50 males (AlGhamdi et al., 2022) and 50 females (Alsufiani et al., 2022) from Saudi Arabia using an identical protocol, which had been established with Finnish cohorts (Seuter et al., 2017). The protocol involves taking blood samples directly before (d0), one day after (d1) and 30 days after (d30) a single oral vitamin D₃ bolus (80,000 IU) supplementation.

The meta-analysis of the data of both studies showed an age range of the 100 participants from 18 to 53 and a body mass index (BMI) spanning from 14.9 to 39.1 (Table S1). The individuals were considered healthy, because they were not diagnosed of any major disease and did not take any medication. However, at baseline (d0) the serum levels of IL6 ranged from 179 to 1079 ng/L, that of IL8 from 67 to 1311 ng/L and that of TNF from 69 to 342 ng/L, i.e., the study participants showed a large variation in basal chronic inflammation. Moreover, the basal vitamin D status ranged from 14.2 to 136.7 nM (5.7–54.7 ng/ml) 25 (OH)D₃, which diagnosed 45 % of the study participants vitamin D deficiency (serum 25(OH)D₃ levels less than 50 nM), for 49 % of them vitamin D insufficiency (serum 25(OH)D₃ levels between 50 to 75 nM) and only for 6 % of the persons vitamin D sufficiency (serum 25(OH)D₃ levels above 75 nM). At d1 only one person remained vitamin D deficient, 82 were still vitamin D insufficient but already 17 were vitamin D sufficient and at d30 1 % of the study participants were vitamin D deficient, 76 % were still vitamin D insufficient but 23 % became

vitamin D sufficient. The average 25(OH)D₃ serum concentration raised from 39.1 nM (15.6 ng/ml at d0) to 61.0 nM (24.4 ng/ml at d1) and even 64 nM (25.6 ng/ml at d30), i.e., by 56 and 64 %, respectively. In our previous studies (AlGhamdi et al., 2022; Alsufiani et al., 2022) we had reported that both in males as well as in female's vitamin D₃ bolus supplementation significantly reduced IL6, IL8 and TNF levels already after one day and even more after 30 days. As expected, the analysis of the samples of in total 100 study participants came to the same results (Fig. 1A).

The main focus of this study was to determine the vitamin D response index of the study participants on the level of vitamin D-triggered changes of cytokine levels. For each individual calculations was made with respect to the respective basal levels, i.e., without vitamin D₃ supplementation. We used the approach to display the changes in cytokine expression (in absolute log scales) between d1 and d0 (Fig. 1B) as well as between d30 and d0 (Fig. 1C) in relation to the respective changes in 25(OH)D₃ serum levels. A linear regression line through the zero point subdivided the graphs into low responders (red, score 0), mid responders (yellow, score 1) and high responders (green, score 3) for the respective parameter. In this way, each of 100 study participants got six scores, the sum of which is proportionally to the vitamin D response index. The index numbers varied from 0 to 16 and allowed the display of all study participants ascendingly (Fig. 2). K-means clustering was applied on this dataset and segregated the study participants into 39 low responders (red), 39 mid responders (yellow) and 22 high responders (green).

In order to challenge the scoring results based on cytokine level changes, we used expression changes of the vitamin D target genes *ASAP2*, *CAMP*, *CD14*, *FBP1*, *ITGAM*, *LRRC25* and *NINJ1* in whole blood cells as alternative biomarkers for the responsiveness to vitamin D₃

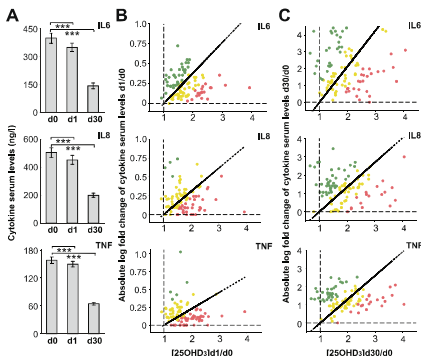


Fig. 1. Change of serum levels of the cytokines IL6, IL8 and TNF of the cohort of 100 participants in response to vitamin D₃ supplementation (A). Scoring of the study participants based on individual changes of the cytokines serum levels at d1 (B) and d30 (C) in relation to d0. Red (low responder); score 0, yellow (mid responder); score 1, green (high responder); score 3. $***, p < 0.001$ (Student's *t*-test). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

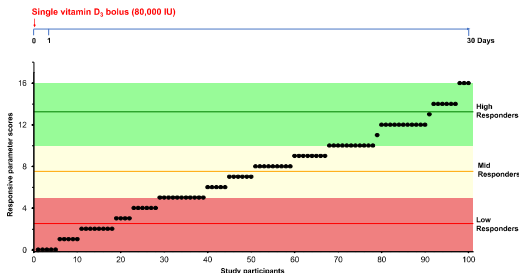


Fig. 2. Ranking of study participants. The *K*-means clustering algorithm was used to segregate the 100 study participants based on the sum of their scores in the responsiveness of cytokines at d1 and d30 (Fig. 1) into low (red), mid (yellow) and high (green) vitamin D responders. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

supplementation. All seven genes are well-known vitamin D target genes as shown by previous vitamin D intervention studies. The vitamin D target gene expression evaluation was done with a subgroup of 37 study participants (13 low, 16 mid and 8 high responders based on cytokine level changes) and resulted in 24 % low responders, 49 % mid responders and 27 % high responders (Fig. 3). Thus, this approach indicates for the chosen subgroup a lower percentage of low responders but more mid responders and high responders than the classification via cytokine serum levels.

The direct comparison of both methods indicated for 30 of 37 study participants (81 % of all) a similar scoring and classification. However, for 7 persons the classification based on target gene expression changes in whole blood differed more than 33 % from that calculated from changes in cytokine serum levels. In order to identify the reason for the difference in the classification, we compared the group of 30 persons with a similar score to the group of 7 individuals with a dissimilar score for their BMI and basal serum levels of IL6, IL8 and TNF (Fig. 4). For IL6 and TNF serum levels there is a trend that the medians of the values are higher for the persons with a dissimilar profile than for those with a similar profile, but no statistically significant difference was observed. To rule out other factors for the observed changes in serum levels of cytokines as well as the target gene expressions, we carried out

significance tests for other variables like BMI, gender and age, but only 25(OH)D₃ serum levels showed to cause these changes.

4. Discussion

This report combined the data of two recently published intervention studies having 50 native males (AlGhamdi et al., 2022) and 50 females (Alsufigi et al., 2022) from Saudi Arabia as participants. The in total 100 individuals were considered representative for the population in Saudi Arabia, since they were not diagnosed of any major disease. However, at baseline 45 % of the study participants were vitamin D deficient and only 6 % were vitamin D sufficient. Although this is not a good sign for the overall vitamin D health in Saudi Arabia, it provides optimal start conditions for vitamin D₃ intervention studies (Bouillon et al., 2022). The single oral vitamin D₃ bolus supplementation increased the average 25(OH)D₃ serum concentration already after one day by more than 20 nM (8 ng/ml). This increased vitamin D status was maintained 30 days later. The latter is not very surprising, since the bolus (80,000 IU) represents a monthly dose (Mazess et al., 2021). Thus, vitamin D₃ bolus supplementations are effective for rapidly elevating the vitamin D status and confirmed comparable previous studies (Witham et al., 2013). However, in clinical routine a vitamin D bolus should be

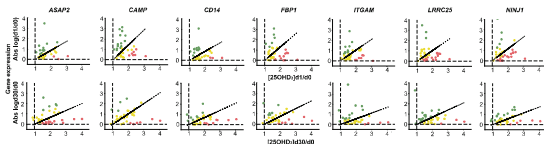


Fig. 3. Scoring of 37 study participants based on changes of the expression of the indicated upregulated vitamin D target genes. Red (low responder): score 0, yellow (mid responder): score 1, green (high responder): score 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

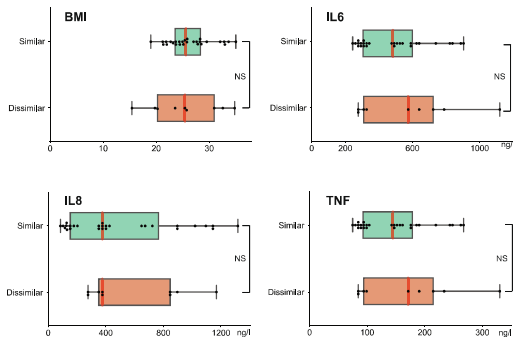


Fig. 4. Comparing similar with dissimilar study participants. Box plots were used to compare the 30 study participants with a similar score in both response index evaluations (green) with 7 persons showing a dissimilar result (red) concerning their BMI and basal serum levels of the cytokines IL6, IL8 and TNF. The red line indicates the median. NS, non-significant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

applied only in cases of severe vitamin D deficiency to individuals who do not comply daily supplementation routines (Hill et al., 2022).

The main aim of this study was to use the vitamin D₃ bolus treatment, in order to follow the protocol of the Finnish vitamin D intervention studies VitDol (NCT02063334) and VitDHD (NCT 03537027) (Hanel et al., 2020) for a fast determination of the vitamin D response index. In principle, the response index can also be evaluated after a longer-term daily vitamin D₃ supplementation, such as demonstrated by the VitD-met study (NCT01479933) (Carlberg et al., 2013), but there should be a few months between both measurements, i.e., the investigations take longer. With the combined data of both studies from Saudi Arabia we first applied the classical approach (Carlberg, 2016) and used the decrease in the serum levels of the proinflammatory cytokines IL6, IL8 and TNF one and thirty days after vitamin D₃ bolus supplementation as a biomarker for the effects of vitamin D. The combined score of the six single response evaluations allowed to segregate the 100 individuals into 39 % low responders, 39 % mid responders and 22 % high responders. The most critical subpopulation is that of the low responders and based on the vitamin D-triggered decrease in proinflammatory cytokines their proportion in this Saudi cohort is clearly higher than the rate in Finnish cohorts (25 %) (Carlberg et al., 2013; Seuter et al., 2017).

The measurement of the expression changes of vitamin D target genes like in this or previous studies is an alternative approach for determining the vitamin D response index. This study is the first one directly comparing both methods and demonstrated that for the majority of the tested persons the result is similar. This is an interesting result, since the populations in Saudi Arabia and Finland are genetically rather different and live in a distinct environment (Investigators et al., 2021). Therefore, this study provides evidence that the concept of the vitamin D response index may be extrapolated to all other populations.

Nevertheless, there is some discrepancy in the determination of vitamin D response index, which may result from higher rate of chronic inflammation detected by elevated basal levels of proinflammatory

cytokines, such as IL6 and TNF. Although the study participants did not have any major diseases, they differed significantly in their BMI, which ranged from 14.9 (underweight) to 39.1 (severe obesity). Accordingly, the basal serum levels of IL6, which may be majorly derived from adipocytes and associated macrophages (Han et al., 2020), also showed a large variation. This suggests that for overweight and obese persons the determination of the vitamin D response index via the downregulation of proinflammatory cytokines can lead to inaccurate results. Since previous assessment of vitamin D response indices had been performed with normal weight persons (BMI ranging from 20 to 25) (Carlberg et al., 2013; Seuter et al., 2017), this aspect before had not been an issue. Thus, vitamin D response index evaluations of overweight and/or diseased persons should be determined on the basis of gene expression rather than via changes of cytokine levels.

In a Finnish cohort of healthy, normal weight persons the rate of low vitamin D responders was determined as 25 % (Seuter et al., 2017). In contrast, 50 % of the participants of the Saudi cohort were overweight or obese and, based on cytokine levels, their rate of low responders was 39 %. This higher rate of low responders may be related to the fact that in the past in Saudi Arabia there was no evolutionary pressure leading to their outnumbering, since the traditional lifestyle provided enough sun exposure for endogenous vitamin D₃ production. However, it is also possible that in northern Europe with long dark winters, like in Finland, there was evolutionary selection of high vitamin D responders, for which a rather low vitamin D status is still sufficient for the activation of vitamin D target genes. Thus, the change of lifestyle in Saudi Arabia over the last two generations made the higher rate of low responders a challenge that needs to be overcome by more intensive vitamin D₃ supplementation (Al-Daghiri et al., 2019; Amer et al., 2022; Sabico et al., 2023). In this context also the significant rise in average BMI (Althumiri et al., 2021) needs to be considered, i.e., vitamin D₃ supplementation should be adjusted by weight.

Determining the vitamin D response index of a person is far more

effort than measuring the vitamin D status, since for the latter only a single measurement of the 25(OH)D₃ serum level is needed, while for the response index two samples and a number of parameters, such as 25 (OH)D₃, cytokines and/or gene expression have to be assessed (Carlberg, 2016). Therefore, it is unlikely that response index measurements will enter routine clinical diagnosis. However, more important is the public awareness of a vulnerable subpopulation of low responders, which need a higher dose of daily vitamin D₃ supplementation than indicated by national guidelines. In most countries the recommendations for daily vitamin D₃ supplementation are in the order of 10–20 µg (400–800 IU) (Spiro and Buttriss, 2014). This may be sufficient for high vitamin D responders but is insufficient for low responders. Since this study confirmed that low vitamin D responders are found not only in Finland and that their percentage may be even higher in countries closer to the equator, but the recommended doses of national guidelines should also be reconsidered. A supplementation with 1 µg vitamin D₃ (40 IU)/kg body weight may serve as a rule that can be easily memorized.

Progress in technology makes measuring gene expression on a transcriptome-wide level a standard for all types of future intervention studies. Moreover, in the context of many nutritional intervention studies 25(OH)D₃ serum level are determined, even if the primary focus of the trials may not be vitamin D. In this way, data are collected that allow the assessment of the vitamin D response index of the study participants. This will facilitate the segregation of the study participants in low, mid and high responders, which could lead to more specific presentations the trial results. The underlying hypothesis of this suggestion is that due to the pleiotropic actions of vitamin D, a low vitamin D response index may affect the results of a larger number of nutritional intervention studies.

This study has a few limitations. Although the cohort size is with 100 individual larger than that of comparable studies from Finland, which had 35–71 participants, an even larger number of studied persons would have been desirable, in order to obtain more accurate numbers on the percentage of low vitamin D responders. Moreover, the underlying studies (AlGhamdi et al., 2022, Alsuifani et al., 2022) are not classical controlled trials. The studies had a focus on adults in an age range of 18–53, but it would be interesting to obtain data also from children as well as from elderly. Finally, the availability of transcriptome-wide data would have further increased the accuracy of the vitamin D response index determination.

5. Conclusion

The concept of the vitamin D response index could be confirmed by a Saudi cohort. In Saudi Arabia the rate of low vitamin D responders is at least as high as in Finland but probably even 14 % larger when determined based on cytokine level changes. This suggests that under the present modern lifestyle, which also in sunny Saudi Arabia is characterized by low UV-B exposure, there should be daily vitamin D₃ supplementation in doses between 2000 and 4000 IU (depending on body weight), which are sufficient even for low vitamin D responders.

Funding

The authors thank the deanship of scientific research (DSR). This work was supported by the deanship of scientific research (DSR) – King Abdulaziz University – Saudi Arabia under Grant [RG: 21-130-38]. C.C. receives funding from the European Union's Horizon2020 research and innovation program under grant agreement no. 952601.

Institutional review Board Statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the ethical committee of the Faculty of Medicine, King Abdulaziz University (ref. no.30–18).

Informed consent Statement

Informed consent was obtained from all subjects involved in the study.

CRediT authorship contribution statement

Shareefa A. AlGhamdi: Writing – review & editing, Supervision, Methodology, Conceptualization, **Ranjini Ghosh Dasgupta:** Supervision, Formal analysis, **Maciej Rybicki:** Visualization, Formal analysis, **Hadeil M. Alsuifani:** Writing – review & editing, Methodology, Conceptualization, **Sawsan O. Khoja:** Conceptualization, **Nusaibah N. Enaib:** Methodology, Data curation, **Safa F. Saif:** Methodology, Data curation, **Carsten Carlberg:** Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jsps.2024.102137>.

References

- Al-Daghiri, N.M., Mohammed, A.K., Bukhari, L., Rikili, M., Abdi, S., Ansari, M.G.A., Sabico, S., Hussain, S.D., Alenad, A., Al-Saleh, Y., Alotaibi, M.S., 2019. Efficacy of vitamin D supplementation according to vitamin D-binding protein polymorphisms. *Nutrients* 11, 149–154. <https://doi.org/10.3390/nu11040149>.
- AlGhamdi, S.A., Enaili, N.M., Alsuifani, H.M., Alshaihi, H.F., Khoja, S.O., Carlberg, C., 2022. A single oral vitamin D₃ bolus reduces inflammatory markers in healthy Saudi males. *Int. J. Mol. Sci.* 23 <https://doi.org/10.3390/ijms23191992>.
- Alsuifani, H.M., AlGhamdi, S.A., Alshaihi, H.F., Khoja, S.O., Saif, S.F., Carlberg, C., 2022. A single vitamin D₃ bolus supplementation improves vitamin D status and reduces proinflammatory cytokines in healthy females. *Nutrients* 14. <https://doi.org/10.3390/nu14143065>.
- Altamimi, N.A., Bayoumi, M.H., AlMousa, A., Aluwaisay, M.F., Almhark, R.A., BinDhim, N.F., Alkhamali, Z., Alqahtani, S.A., 2021. Obesity in Saudi Arabia in 2020: prevalence, distribution, and its current association with various health conditions. *Healthcare (Basel)* 9. <https://doi.org/10.3390/healthcare9030311>.
- Amer, O.E., Sabico, S., Sheshah, E., Alotaibi, N.H., Aldidi, D.A., Enani, M.A., Aljohani, N., J., Alshingir, N., Alomar, S.Y., Hussain, S.D., Alnaimi, A.M., Elsaid, M.A., Al-Daghiri, N.M., 2022. Evaluation of 34 cytokines and vitamin D status reveal a sexually-dimorphic active immune response to SARS-CoV-2. *Healthcare (Basel)* 10. <https://doi.org/10.3390/healthcare10122571>.
- Bendik, L., Friedel, A., Roos, F.F., Weber, P., Eggensdorfer, M., 2014. Vitamin D: a critical and essential micronutrient for human health. *Front Physiol.* 5, 248. <https://doi.org/10.3389/fphys.2014.00248>.
- Bishop, E., Ismailova, A., Dimeloe, S.K., Hewison, M., White, J.H., 2020. Vitamin D and immune regulation: antibacterial, antiviral, anti-inflammatory. *JBMIR pub.* <https://doi.org/10.1002/jbm.10405>.
- Bouillon, R., Manosakul, D., Rosen, C., Trajanoska, K., Rivasencina, F., Richards, J.B., 2022. The health effects of vitamin D supplementation: evidence from human studies. *Nat. Rev. Endocrinol.* 18, 96–110. <https://doi.org/10.1038/s41574-021-00593-z>.
- Carlberg, C., 2016. Molecular approaches for optimizing vitamin D supplementation. *Vitam Horm.* 100, 255–271. <https://doi.org/10.1016/bv.2015.10.001>.
- Carlberg, C., Haq, A., 2018. The concept of the personal vitamin D response index. *J. Steroid Biochem. Mol. Biol.* 175, 12–17. <https://doi.org/10.1016/j.jsbmb.2016.12.011>.
- Carlberg, C., Seuter, S., de Mello, V.D., Schwab, U., Voutilainen, S., Pulkkinen, K., Nurmi, T., Viitanen, J., Tuominen, T.P., Uusitalo, M., 2013. Primary vitamin D target genes allow a categorization of possible benefits of vitamin D supplementation. *PLoS ONE* 8, e71042. <https://doi.org/10.1371/journal.pone.0071042>.
- Dimitrov, V., Rautava, J., Ismailova, A., Wang, Y., Dimoski, R., Salehi-Taher, R., Memari, B., Groussoulis, E., White, J.H., 2021. Vitamin D-regulated gene expression profiles: species-specificity and cell-specific effects on metabolism and immunity. *Endocrinology* 162. <https://doi.org/10.1210/endo/bqab218>.
- Fleet, J.C., 2022. Vitamin D-mediated regulation of intestinal calcium absorption. *Nutrients* 14. <https://doi.org/10.3390/nu14163351>.
- Han, M.S., White, A., Perry, R.J., Camporeale, J.P., Haldago, J., Shulman, G.I., Davis, R.J., 2020. Regulation of adipose tissue inflammation by interleukin-6. *Proc. Natl. Acad. Sci. U S A* 117, 2751–2760. <https://doi.org/10.1073/pnas.1920004117>.

- Hanel, A., Carlberg, C., 2020. Skin color and vitamin d: an update. *Exp. Dermatol.* 29, 866–875. <https://doi.org/10.1111/exd.14142>.
- Hanel, A., Neme, A., Mallon, M., Hansdottir, E., Malmberg, H.R., Eitene, S., Tuomimäki, T.P., Virtanen, J.K., Bendik, L., Carlberg, C., 2020. Common and personal target genes of the micronutrient vitamin d in primary immune cells from human peripheral blood. *Sci. Rep.* 10, 21051. <https://doi.org/10.1038/s41598-020-78289-0>.
- Harrison, S.R., Li, D., Jeffery, L.E., Raza, K., Hewison, M., 2020. Vitamin d, autoimmune disease and rheumatoid arthritis. *Calcif. Tissue Int.* 106, 56–75. <https://doi.org/10.1007/s00222-019-00577-2>.
- Hill, A., Starchel, C., Drosen, E., Stoppe, C., Amrein, K., 2022. An update of the effects of vitamins d and c in critical illness. *Front Med. (Lausanne)* 9, 1083760. <https://doi.org/10.3389/fmed.2022.1083760>.
- Holick, M.F., Binkley, N.C., Bischoff-Ferrari, H.A., Gordon, C.M., Hanley, D.A., Heaney, R.P., Murad, M.H., Weaver, C.M., Endocrine, S., 2011. Evaluation, treatment, and prevention of vitamin d deficiency: an endocrine society clinical practice guideline. *J. Clin. Endocrinol. Metab.* 96, 1911–1930. <https://doi.org/10.1210/nc.2011-0385>.
- Investigators, G.P.P., Smedley, D., Smith, K.R., Martin, A., Thomas, E.A., McDonagh, E. M., Cipriani, V., Ellingford, J.M., Arno, G., Tucci, A., Vandrovcova, J., Chan, G., Williams, H.J., Ramaikie, T., Wei, W., Strirups, K., Ibanez, K., Moutianas, L., Wielscher, M., Need, A., Barnes, M.R., Vesitio, L., Buchanan, J., Wordsworth, S., Ashford, S., Rehmstrom, K., Li, E., Fuller, G., Twiss, P., Spasic-Boskovic, O., Haskill, S., Jkono, K.A., Poole, K., Wagner, A., Mehia, S.G., Gurnell, M., Burrows, N., James, R., Penkett, C., Dewhurst, E., Graf, S., Mayana, R., Kassnicki, M., Haworth, A., Savage, H., Babcock, M., Reese, M.G., Blake, M., Barle, E., Boustred, C., Brittain, H., de Burca, A., Bleda, M., Devereaux, A., Haskil, D., Hatzikakioti, E., Hyder, Z., Kasperaviciute, D., Patch, C., Polychronopoulos, D., Matchan, A., Sultana, R., Ryten, M., Tavares, A.L.T., Tregidgo, C., Turnbull, C., Welland, M., Wood, S., Snow, C., Williams, E., Leigh, S., Foulger, R.E., Daugherty, L.C., Niblock, O., Leong, I. U.S., Wright, C.F., Davies, J., Crichton, C., Welch, J., Woods, K., Abulboul, L., Auerka, P., Bodenbauer, D., Bronsfield, A., Cleary, M.A., Lam, T., Duttani, M., Footitt, E., Ganesan, V., Grunewald, S., Compeyrol-Lacassagne, S., Muntoni, F., Pilkington, C., Quinlan, R., Thapar, N., Wallis, C., Wedderburn, L.R., Worth, A., Bueser, T., Compton, C., Deshpande, C., Fasih, H., Haque, E., Itzatt, L., Josifova, D., Mohammed, S., Robert, L., Rose, S., Ruddy, D., Sarkany, R., Say, G., Shaw, A.C., Wolejko, A., Habib, B., Burns, G., Hunter, S., Grocock, R.J., Humphray, S.J., Robinson, P., Haendel, M., Simpson, M.A., Banka, S., Clayton-Smith, J., Dougan, S., Hall, G., Thomas, H.B., O'Keefe, R.T., Michaelides, M., Moore, A.T., Malik, S., Puntikov, N., Browning, A.C., Straub, V., Gorman, G.S., Horvath, R., Quinton, R., Schaefer, A.M., Yu-Wai-Man, P., Turnbull, D.M., McFarland, R., Taylor, R.W., O'Connor, E., Yip, J., Newland, K., Morris, H.R., Polke, J., Wood, N.W., Campbell, C., Camps, C., Gibson, K., Koelling, N., Lester, T., Nemeth, A.H., Palles, C., Patel, S., Roy, N.B.A., Sen, A., Taylor, J., Cacheiro, P., Jacobsen, J.O., Seaby, E.G., Davison, V., Chitty, L., Douglas, A., Naveh, K., McMullan, P., Elford, S., Tompa, I.K., Mumford, A.D., Wilson, G., Beales, P., Ritner-Glindicic, M., Black, G., Bradley, J.R., Brennan, P., Burn, J., Chinnery, P.F., Elliott, P., Flinter, F., Houlden, H., Irving, M., Newman, W., Rahman, S., Sayer, J.A., Taylor, J.C., Webster, A.R., Wilkie, A.O.M., Ouweland, W.H., Raymond, F.L., Chisholm, J., Hill, S., Bentley, D., Scott, R.H., Fowler, T., Rendon, A., Caulfield, M., 2021. 100,000 genomes pilot on rare-disease diagnosis in health care – preliminary report. *N. Engl. J. Med.* 385, 1866–1880. <https://doi.org/10.1056/NEJMed2035790>.
- Mazess, R.B., Bischoff-Ferrari, H.A., Dawson-Hughes, B., 2021. Vitamin d: bonus is bogus – a narrative review. *JBMIR plus* 5, e10567. <https://doi.org/10.1002/jbm4.10567>.
- Pike, J.W., Meyer, M.B., Lee, S.M., Orsai, M., Benkovic, N.A., 2017. The vitamin d receptor: contemporary genomic approaches reveal new basic and translational insights. *J. Clin. Invest.* 127, 1146–1154. <https://doi.org/10.1172/JCI88887>.
- Pile, S., Zittermann, A., Trummer, C., Theiler-Schwetz, V., Lerchbaum, E., Keppel, M.H., Griebler, M.R., Marx, W., Pandis, M., 2019. Vitamin D testing and treatment: a narrative review of current evidence. *Endocr. Connect.* 8, R27–R43. <https://doi.org/10.1530/EC-18-0432>.
- Sabico, S., Wani, K., Grant, W.B., Al-Daghiri, N.M., 2023. Improved hdl cholesterol through vitamin d status correction substantially lowers 10-year atherosclerotic cardiovascular disease risk score in vitamin d-deficient arab adults. *Nutrients* 15. <https://doi.org/10.3390/nu15030551>.
- Seuter, S., Virtanen, J.K., Nurmi, T., Pihlakajärvi, J., Mursu, J., Voutilainen, S., Tuomimäki, T.P., Neme, A., Carlberg, C., 2017. Molecular evaluation of vitamin d responsiveness of healthy young adults. *J. Steroid Biochem. Mol. Biol.* 174, 314–321. <https://doi.org/10.1016/j.jsbmb.2016.06.003>.
- Spiro, A., Buttriss, J.L., 2014. Vitamin d: an overview of vitamin d status and intake in europe. *Nutr. Bull.* 39, 322–350. <https://doi.org/10.1111/nbu.12108>.
- Uday, S., Hogler, W., 2017. Nutritional rickets and osteomalacia in the twenty-first century: revised concepts, public health, and prevention strategies. *Curr. Osteoporos Rep.* 15, 293–302. <https://doi.org/10.1007/s11991-017-0383-y>.
- van Driel, M., van Leeuwen, J.P.T.M., 2023. Vitamin d and osteoblasts: an endo/autocrine story. *Nutrients* 15, 480. <https://doi.org/10.3390/nu15030480>.
- Wacker, M., Holick, M.F., 2013. Sunlight and vitamin d: a global perspective for health. *Dermatol. Immunol.* 5, 51–108. <https://doi.org/10.4161/derm.24494>.
- Witham, M.D., Adams, F., Babir, G., Kennedy, G., Belch, J.J., Khan, F., 2013. Effect of short-term vitamin d supplementation on markers of vascular health in south asian women living in the uk – a randomised controlled trial. *Atherosclerosis* 230, 293–299. <https://doi.org/10.1016/j.atherosclerosis.2013.08.005>.
- Zmijewski, M.A., Carlberg, C., 2020. Vitamin d receptor(s): in the nucleus but also at membranes? *Exp. Dermatol.* <https://doi.org/10.1111/exd.14147>.



OPEN Transcriptomic profiling of immune modulation induced by vitamin D₃ in the VitDPAS and VitDHiD cohort studies

Emilia Gospodarska^{1,4}, Ranjini Ghosh Dastidar^{1,4}, Julia Jarosławska¹, Maciej Rybiński¹, Marianna Raczky¹, Kornelia Tokarczyk-Malesa², Jerzy Romaszko² & Carsten Carlberg^{1,3,✉}

The VitDPAS study (NCT06104111) was designed as a medical experiment to assess the *in vivo* effects of vitamin D on immune responses. This study enrolled 45 healthy individuals from Olsztyn, Poland, who received a body weight-adjusted bolus dose of vitamin D₃ (1,000 IU/kg). Transcriptome-wide differential gene expression analysis of peripheral blood mononuclear cells, collected before and 24 h after supplementation, identified 758 significantly responsive genes ($p < 0.05$). By correlating individual gene expression changes with alterations in vitamin D status, participants were categorized into three response groups: 17 high responders, 19 mid responders, and 9 low responders. A comparative analysis with the VitDHiD study (NCT03537027), conducted on a Finnish cohort of 25 healthy participants, revealed 232 overlapping target genes, enabling an integrated assessment of vitamin D responsiveness across all 70 individuals. Applying a more stringent statistical threshold (false discovery rate < 0.05) highlighted 26 shared target genes, demonstrating a consistent *in vivo* response to vitamin D₃ across both cohorts. The modulation of inflammatory processes, mediated primarily via tumor necrosis factor and nuclear factor κ B signaling pathways, emerged as a shared effect, highlighting the immunomodulatory potential of vitamin D as a key function of the vitamin in healthy individuals.

Keywords Vitamin D, Vitamin D intervention studies, Transcriptome, Vitamin D target genes, Vitamin D response index, Inflammatory response

Abbreviations

1,25(OH) ₂ D ₃	1 α ,25-dihydroxyvitamin D ₃
25(OH)D ₃	25-hydroxyvitamin D ₃
ACSL1	Acyl-CoA synthetase long chain family member 1
ATAC-seq	assay for transposase accessible chromatin with high-throughput sequencing
BMI	Body mass index
CDKN1A	Cyclin dependent kinase inhibitor 1 A
CPM	Counts per million
CSRN1	cysteine and serine rich nuclear protein 1
CX3CR1	C-X3-C motif chemokine receptor 1
CXCR4	C-X-C motif chemokine receptor 4
ETS2	ETS proto-oncogene 2, transcription factor
FC	Fold change
FDR	False discovery rate
GADD45B	Growth arrest and DNA damage inducible beta
GIMAP6	GTPase, IMAF family member 6
GLUL	Glutamate-ammonia ligase
HERPUD1	Homocysteine inducible ER protein with ubiquitin like domain 1
HGNC	Human genome nomenclature committee

¹Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, ul. Trylińskiego 18, 10-683 Olsztyn, Poland. ²Department of Family Medicine and Infectious Diseases, School of Medicine, Collegium Medicum, University of Warmia and Mazury, Olsztyn, Poland. ³Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland. ⁴These authors contributed equally: Emilia Gospodarska and Ranjini Ghosh Dastidar. ✉email: c.carlberg@pan.olsztyn.pl

HIF1A	Hypoxia inducible factor 1 subunit alpha
IRS2	Insulin receptor substrate 2
KEGG	Kyoto Encyclopedia of Genes and Genomes
KLF	KLF transcription factor
MDS	Multidimensionality scaling
NAMPT	Nicotinamide phosphoribosyltransferase
NFkB	Nuclear factor kB
NFKB1	NFKB inhibitor
PBMC	Peripheral blood mononuclear cell
PEL1	Pellino E3 ubiquitin protein ligase 1
PER1	Period circadian regulator 1
PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
RBM7	RNA binding motif protein 7
RNA-seq	RNA sequencing
S1PR1	Sphingosine-1-phosphate receptor 1
SLC2A3	Solute carrier family 2 member 3
TNF	Tumor necrosis factor
TNFAIP3	TNF alpha induced protein 3
TNFSF14	TNF superfamily member 14
TSC22D3	TSC22 domain family member 3
TSS	Transcription start site
VDR	Vitamin D receptor

Vitamin D₃ is a vital micronutrient that can be synthesized endogenously in human skin when exposed to UV-B light¹. Both in vitro studies and animal model experiments suggest that vitamin D plays a significant role in modulating the activity of innate and adaptive immunity^{2–6}. Observational studies comparing individuals with vitamin D deficiency to those with sufficient levels suggest that adequate vitamin D may provide protective benefits against a range of diseases. These include musculoskeletal disorders such as osteoporosis⁷ and sarcopenia⁸, autoimmune diseases like multiple sclerosis⁹, and certain types of cancer, including colon cancer¹⁰. In vitamin D deficiency, the underlying issue often relates to insufficient immune modulation⁵. However, conclusive evidence of the beneficial effects of vitamin D supplementation in healthy, non-deficient individuals is lacking. Large-scale intervention studies with thousands of participants over periods up to five years was unable to demonstrate statistically significant primary benefits from vitamin D supplementation, partly due to ethical constraints against leaving control groups vitamin D deficient for extended periods^{11,12}.

Vitamin D's molecular mechanisms involve its nuclear receptor, VDR (vitamin D receptor), which binds the active vitamin D metabolite, 1 α ,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃)^{13,14}, with very high affinity ($K_D = 0.1$ nM). For endocrine signaling, this nuclear hormone is synthesized from 25-hydroxyvitamin D₃ (25(OH)D₃) in the kidneys but can also be locally produced in immune and skin cells for para- and autocrine functions¹⁵. The most abundant and stable metabolite, 25(OH)D₃, defines vitamin D status in serum¹⁶, with levels above 30 ng/ml (75 nM) considered sufficient and those below 12 ng/ml (30 nM) indicating severe deficiency¹⁷.

As a transcription factor, VDR mediates all genomic actions of vitamin D, with 1,25(OH)₂D₃ being the exclusive activator of VDR at physiological concentrations¹⁸. Based on data of the Genotype-Tissue Expression (GTEx) Portal¹⁹, VDR is expressed in most tissues, excluding the brain, with hundreds of target genes in each tissue²⁰, many of which exhibit tissue-specific and individual-specific responsiveness^{12,21}. This variation supports the concept of the vitamin D response index, categorizing individuals as high, mid, or low responders to vitamin D supplementation²². Notably, low responders may require higher vitamin D dosages to attain optimal physiological benefits²³.

To investigate vitamin D's molecular effects in healthy individuals under in vivo conditions, we previously conducted studies such as the VitDol (NCT02063334, [ClinicalTrials.gov](https://clinicaltrials.gov))^{24,25} and VitDHiD^{26,27}, involving healthy cohorts in Kuopio, Finland, who received a single 80,000 IU vitamin D₃ bolus. In the current study, we expanded on this approach by incorporating the VitDPAS study, which enrolled a cohort of 45 healthy individuals from Olsztyn, Poland. Located more than 1000 km south of Kuopio, the vitamin D winter in Olsztyn is about one month shorter than in Eastern Finland. Transcriptome-wide analysis of peripheral blood mononuclear cells (PBMCs) identified 758 significantly ($p < 0.05$) responsive vitamin D target genes, of which 232 overlapped with targets observed in the VitDHiD study. A core function of these common in vivo vitamin D target genes was the modulation of inflammatory responses.

Materials and methods

Sample collection

The VitDPAS intervention study commenced in November 2023 in Olsztyn, Poland. Participants eligible for inclusion were between 20 and 65 years old, with a body mass index (BMI) of 18.5–29.9 kg/m². Exclusion criteria included any disease or infection contracted within the last two weeks before the planned blood draw; severe visual or hearing impairment; gastrointestinal disorders like celiac disease or irritable bowel syndrome; chronic kidney disease; presence of kidney stones; chronic liver diseases; history of pathological bone fractures or fractures due to osteoporosis within the last five years; current diagnosis of cancer; treatment that affects calcium metabolism; diseases with a risk of recurring symptoms, including Parkinson's disease, post-stroke hemiplegia, epilepsy, recurrent dizziness, and repeated fainting; past or current diseases that may increase serum calcium levels, such as sarcoidosis, lymphoma, and hyperparathyroidism; use of medications that affect 25(OH)D₃ levels like phenobarbital, carbamazepine, and phenytoin; Paget's disease of bone (osteitis deformans), arthritis including

rheumatoid arthritis, Reiter's syndrome, and psoriatic arthritis; participation in research studies within the last 4 months, except for observational studies; known infection with HIV, hepatitis, and other infectious diseases; substance abuse like drug dependence, alcoholism, and tobacco smoking; mental illness; homeostasis disorders that may complicate the blood collection process; current pregnancy, breastfeeding, or planned pregnancy. The study protocol was approved by the Ethics Committee of the Olsztyn Chamber of Doctors (#31/2023/VIII).

The study involved a cohort of 45 healthy participants (23 females and 22 males, aged 24–64, BMI 19.3–29.3, Table 1) who received a single vitamin D₃ bolus of 1,000 IU/kg body mass with breakfast. This bolus equates to a monthly dose of vitamin D₃, administered only once, whereas other vitamin D interventions, such as the ViDA study²⁶, employed a monthly dose of 100,000 IU of vitamin D₃ over a period of more than two years.

Participant	Age	Gender	BMI	[25(OH)D ₃] at d0	[25(OH)D ₃] at d1
1	40	F	22.6	26.9	32.3
2	45	F	23.1	19.3	28.6
3	34	F	20.2	16.1	23.7
4	41	F	19.8	38.3	47.3
5	49	F	21.2	29.5	38.0
6	48	F	20.2	24.7	35.2
7	40	F	24.7	21.8	30.4
8	38	F	20.4	19.9	23.6
9	37	F	21.2	26.6	30.0
10	48	F	25.0	18.8	30.0
11	43	F	19.7	21.5	27.8
12	43	F	21.2	34.3	39.9
13	40	F	21.6	47.0	57.5
14	50	F	20.4	35.3	40.1
15	23	F	24.2	30.0	39.3
16	40	F	22.5	31.2	38.7
17	41	F	22.1	27.7	33.4
18	24	F	23.1	16.8	22.2
19	43	F	24.6	28.1	35.5
20	35	F	21.5	20.9	25.0
21	53	F	19.3	54.8	62.6
22	47	F	24.2	14.0	18.6
23	47	F	23.3	21.4	26.0
24	47	M	25.3	41.6	48.7
25	35	M	23.6	12.7	22.1
26	53	M	29.3	30.0	34.7
27	64	M	27.1	41.2	42.7
28	25	M	21.1	24.5	36.4
29	49	M	27.9	23.1	29.1
30	31	M	28.7	15.9	24.2
31	27	M	28.6	14.3	20.6
32	47	M	29.6	16.3	21.0
33	40	M	24.0	55.9	57.8
34	37	M	23.4	33.7	41.9
35	24	M	21.2	25.7	31.4
36	25	M	20.3	25.6	30.5
37	37	M	19.3	15.3	20.7
38	39	M	27.6	32.4	34.7
39	34	M	27.5	41.7	45.1
40	47	M	29.5	26.6	32.7
41	25	M	20.6	24.9	29.7
42	31	M	24.9	9.0	15.2
43	38	M	27.2	17.4	22.5
44	38	M	28.9	11.4	19.9
45	42	M	28.3	28.6	32.1

Table 1. VitDPAS study participants. Age, gender, BMI and 25(OH)D₃ serum levels (in ng/ml) at d0 and d1 are indicated.

Therefore, the dose range of 52,000–110,000 IU vitamin D₃ used in this study was within safe limits. Blood samples were collected immediately before supplementation (day 0 (d0)) and 24 h after dosing (day 1 (d1)) for serum and PBMC isolation. Based on data collected from previous studies, such as VitDbol^{24,25}, a 24-hour post-supplementation period is considered the optimal time point. Serum 25(OH)D₃ levels were assayed using the electrochemiluminescence binding assay on Cobas Pure Immunoassay Analyzer (Roche) in the analytical laboratory of the Municipal Polyclinic Hospital in Olsztyn, Poland. All participants provided written informed consent, and all experimental procedures adhered to relevant guidelines and regulations.

PBMC isolation

PBMCs from the 45 participants of the VitDPAS study were promptly processed without any *in vitro* culture. Following the collection of 8 ml of peripheral blood on d0 and d1, PBMCs were isolated within one hour using Vacutainer CPT Cell Preparation Tubes with sodium citrate (Becton Dickinson), following the manufacturer's protocol. After isolation, cells were washed with phosphate-buffered saline, aliquoted at a concentration of 4 million cells per ml, and preserved at -80 °C for future RNA isolation.

Transcriptome analysis

Total RNA was extracted from PBMCs using the High Pure RNA Isolation Kit (Roche) following the manufacturer's instructions. RNA quality was assessed using an Agilent TapeStation, and library preparation was conducted after rRNA depletion with kits and protocols from New England Biolabs. RNA-seq libraries were sequenced at a read length of 75 bp on a NextSeq2000 system (Illumina) according to standard protocols at the EMBL Gene Core (Heidelberg, Germany). Quality control of the sequencing files was performed using FastQC (version 0.11.9, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) (Supplementary Table S1 online). Reads were aligned to the GRCh38 reference genome (Ensembl version 111.38) using the STAR aligner²⁶ (version 2.7.2a), and quantification was carried out with FeatureCounts³⁰ (version 2.0.1) using default parameters. To standardize gene nomenclature across all datasets, Human Gene Nomenclature Committee (HGNC) symbols were updated using the R package HGNChelper (version 0.8.1, <https://CRAN.R-project.org/package=HGNChelper>). Annotation, including gene identifiers, descriptions, genomic locations, and biotypes, was added via the Ensembl database (release 100) using the R package BiomaRt⁴¹ (version 2.42.1). Entrez Gene identifiers were incorporated using the R package org.Hs.eg.db (version 3.10.0), and any incomplete mappings for target genes were manually verified and retrieved from NCBI (<http://www.ncbi.nlm.nih.gov/home/genes>). Genes lacking genomic position information or those mitochondrially encoded were excluded from the analysis. This study focused solely on protein-coding genes.

Differential gene expression analysis

Differential gene expression analysis was conducted in R (version 4.3.1) on MacOS 13 (Ventura) using the *EdgeR* package⁴² (version 4.0.16) for robust assessment. To reduce transcriptional noise associated with non-coding genes, the analysis focused on 19,287 protein-coding genes. Read counts were normalized to counts per million (CPM) to account for library size differences. Low-expression genes (CPM < 10) were filtered out using the *FilterByExpr*() function, minimizing the multiple testing burden and optimizing statistical accuracy within the *EdgeR* framework. Following filtering, library sizes were recalculated, and the trimmed mean of M-values normalization was applied. The transcriptome data structure was explored using multidimensional scaling (MDS) via *EdgeR*'s *plotMDS*() function, where distances approximate typical log₂ fold changes (FC) between samples. These distances were calculated as the root mean square deviation (Euclidean distance) of log₂ FC values for genes showing significant changes (p-value < 0.05 or Benjamini-Hochberg adjusted p-value = FDR (false discovery rate) < 0.05)) post-vitamin D₃ supplementation. Mean-Difference (MA) plots were generated with *vizzy* (version 1.0.0, <https://github.com/ATpoint/vizzy>). Differential expression for each gene was determined through the generalized linear model quasi-likelihood (GLM-QL) pipeline. A trended negative binomial dispersion estimate was calculated using the Cox-Reid profile-adjusted likelihood method. This estimate, along with empirical Bayes-moderated quasi-likelihood gene-wise dispersion estimates⁴³, was employed for GLM fitting, with empirical Bayes shrinkage robustified against outlier dispersions, as recommended. To identify differentially expressed genes with substantial changes, the *glmTreat* approach⁴⁴ was applied, setting a threshold for absolute log₂FC > 0.25 (Supplementary Table S2 online).

To integrate the gene expression data from the VitDPAS and VitDHiD cohorts, matched common gene identifiers across both datasets were identified, ensuring that only shared genes were included in the combined dataset. To account for differences in sequencing depth, the expression values were normalized using CPM, which adjusts raw counts accordingly. Additionally, a scaling step was applied to harmonize the data across time points (d0 and d1). This involved calculating scaling factors based on the average gene expression levels of both cohorts and adjusting the data to minimize technical discrepancies, thereby facilitating a more accurate comparison.

Classification of study participants

To classify participants as high, mid, or low responders, the average absolute log₂FC values of all *in vivo* vitamin D target genes (or relevant subsets) for each individual (Supplementary Table S2 online) was plotted against the changes in total 25(OH)D₃ concentrations (d1/d0) (Table 1). Based on prior experience with the vitamin D intervention studies VitDmet²⁴ and VitDbol^{24,25}, plotting the ratio of 25(OH)D₃ values proves to be more effective than using their delta. A trendline was drawn through the 1/0 point to divide the plot, and the orthogonal distance of each data point from this line was calculated. These distances were then ranked in descending order, and clustering of participants into three response groups was performed using the *K-means* algorithm (KMeans function from the sklearn.cluster library, available at <https://pypi.org/project/scikit-learn/>).

Functional characterization of vitamin D target genes

The STRING database (<http://string-db.org/>)³⁵ was used to confirm and visualize the physical and functional interactions among vitamin D target genes. In this network, each node represents a protein, while edges between nodes indicate associations inferred from a combination of experimental, computational, and literature-derived data. Functional annotation of the genes, with an emphasis on their primary roles, was derived from databases like Human Protein Atlas (<http://www.proteinatlas.org/>) and GeneCards (<http://www.genecards.org/>).

Results

The VitDPAS intervention trial

The VitDPAS study design is detailed in Supplementary Figure S1A online. The study began by administering a weight-adjusted vitamin D₃ bolus (1,000 IU/kg) to 45 healthy participants (23 females, 22 males; mean age: 39.6 years; mean BMI: 23.8, Supplementary Figure S1B online). Blood samples were collected immediately before supplementation (d0) and 24 h afterward (d1). Baseline serum 25(OH)D₃ levels at d0 ranged from 9.0 to 55.9 ng/ml, with an average of 26.5 ng/ml, while levels at d1 increased to a range of 18.6 to 62.6 ng/ml, with an average of 32.9 ng/ml (Table 1). Thus, the vitamin D₃ bolus resulted in a significant average increase in vitamin D status by 6.4 ng/ml, or 29% ($p = 3.8 \times 10^{-23}$, paired T-test, Supplementary Figure S1C online). In summary, a single, weight-adjusted vitamin D₃ bolus was sufficient to elevate the vast majority of VitDPAS participants to a level of vitamin D sufficiency.

Vitamin D-induced transcriptomic response in VitDPAS participants

Differential gene expression analysis of RNA-seq data at d1 compared to baseline (d0) revealed that vitamin D₃ bolus supplementation significantly impacted the transcriptome of the 45 VitDPAS participants. Principal component analysis, visualized in an MDS plot, showed a notable shift, with d1 data points averaging below the respective d0 points along dimension 2 (Fig. 1A). Out of 8,937 protein-coding genes expressed in PBMCs across all participants (defined as mean CPM > 10 for d0 and d1), 758 genes were significantly ($p < 0.05$) regulated by vitamin D (Figure S2). Of these, 369 were upregulated and 389 downregulated (Supplementary Table S2 online, Supplementary Figure S2 online). Applying an additional threshold of absolute $\log_2FC > 0.25$ narrowed this list to 94 vitamin D target genes, including 39 upregulated and 55 downregulated. A stricter criterion of FDR < 0.05 identified 70 genes, with 15 upregulated and 55 downregulated, and 39 of these also meeting the absolute $\log_2FC > 0.25$ threshold (5 upregulated and 34 downregulated) (Fig. 1B). Taken together, vitamin D₃ bolus supplementation significantly modulates the transcriptome of circulating immune cells in VitDPAS participants.

Segregation of VitDPAS study participants by vitamin D responsiveness

To evaluate the vitamin D responsiveness among the 45 VitDPAS participants, each individual's average absolute $\log_2FC$ of all 758 vitamin D target genes was plotted against the ratio of their vitamin D status at d1 relative to baseline (d0) (Fig. 2A). This analysis revealed a spectrum of responsiveness, with high responders clustering at the upper left of the plot (green), mid responders near the trendline (yellow), and low responders at the lower right (red). The distance from the trendline was proportional to each participant's vitamin D response index, which showed a strong negative correlation (-0.82 , Pearson correlation, $p = 3.7 \times 10^{-12}$, paired T-test) with the d1/d0 vitamin D status ratio. This response index correlated less strongly with baseline vitamin D status (0.61 , $p = 0.0000089$, Supplementary Figure S3A online), and showed no significant correlation with age (0.23 , $p = 0.12$, Supplementary Figure S3B online) or BMI (-0.08 , $p = 0.58$, Supplementary Figure S3C online) of the individuals. *K-means* clustering further classified the participants into 9 low responders, 19 mid responders, and 17 high responders (Supplementary Figure S4 online). When the analysis was limited to the 94 genes with an absolute $\log_2FC > 0.25$, the results were similar; however, participants #9, 14, 16, and 39 were reclassified as mid responders (Fig. 2B). Narrowing the gene list further to 70 targets with FDR < 0.05, individuals #1, 9, 14, 16, 19, and 39 were reclassified as mid responders, while participants #18 and #22 were newly classified as high responders, and #28 as a low responder. Consequently, participants #1 and #19 appear as unstable high responders, while #9, 14, 16, 18, 22, 28, and 39 are unstable mid responders. In summary, VitDPAS participants were stratified into high, mid, and low responders to vitamin D₃ supplementation.

Joint classification of VitDPAS and vitdhid participants

Comparison of the 758 *in vivo* vitamin D target genes from the VitDPAS study with the 1,654 targets identified in the VitDHID study ($p < 0.05$) revealed 232 shared genes (Fig. 3A). The two studies, which both investigated healthy individuals (VitDHID with 25 participants from Finland), were highly comparable and analyzed using identical methods. The combined transcriptome datasets from all 70 participants were normalized and analyzed using principal component analysis via MDS plots. Based on the 232 common target genes, on average, d1 datasets clustered to the left of d0 datasets along the first dimension, suggesting that vitamin D₃ supplementation consistently impacted the PBMC transcriptome across both cohorts (Fig. 3B). Segregation of the 70 individuals by vitamin D responsiveness classified them into 22 high responders, 34 mid responders, and 14 low responders (Fig. 3C). Four VitDPAS participants were reclassified based on this joint analysis: participant #19 shifted from unstable high to mid responder, #16 and #39 from unstable mid to mid responders, and #28 from unstable mid to low responder. Similarly, four VitDHID participants were reclassified (three from mid to high responders and one from mid to low responder)³⁷. For the remaining 41 VitDPAS and 21 VitDHID participants (88.6% overall), the joint analysis produced results consistent with each individual study's classification. In summary, this combined analysis of the VitDPAS and VitDHID vitamin D intervention studies identified 232 common target genes and yielded consistent response classifications for 62 of the 70 participants.

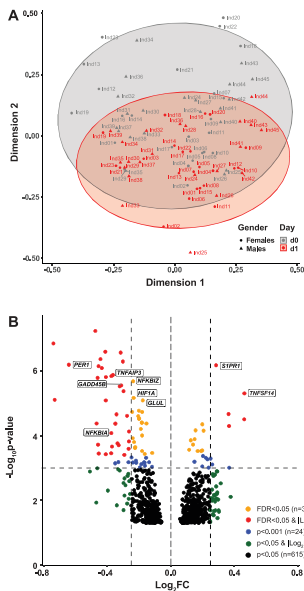


Fig. 1. Transcriptome-wide analysis of vitamin D₃ bolus supplementation in the VitDPAS cohort. **(A)** The quality of transcriptome datasets (of 758 target genes) from the 45 VitDPAS participants at d0 and d1 was assessed using dimensionality reduction techniques, visualized through a MDS plot. **(B)** A Volcano plot was employed to illustrate the impact of vitamin D₃ supplementation on the expression of 758 significantly regulated genes ($p < 0.05$). Selected target genes are highlighted.

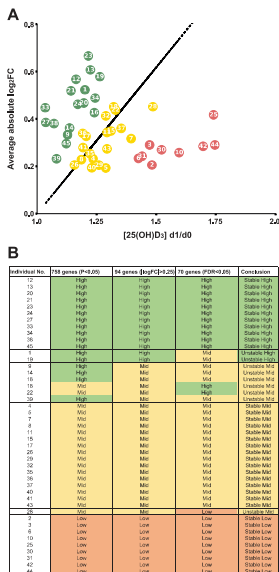


Fig. 2. Classification of VitDPAS study participants. (A) The 45 VitDPAS participants were classified into high (green), mid (yellow), and low (red) responders based on the expression profiles of all 758 *in vivo* vitamin D target genes. (B) Comparison of classification methods utilizing various subsets of vitamin D target genes to evaluate the robustness of the segregation approach.

Functional impact of joint VitDPAS and vitdhid target genes

A stringent comparison (FDR < 0.05) between 70 vitamin D target genes identified in the VitDPAS study and 452 targets from the VitDHid study²⁷ revealed 26 shared genes (Fig. 4A). Notably, of these common genes, only 4 were upregulated by vitamin D₃ supplementation, while 22 were downregulated (Supplementary Figure S5 online). The 26 genes were validated by two different approaches. A repeated vitamin D₃ bolus supplementation with the same study participants indicated for 80–90% of genes a comparable regulation on the RNA level (Supplementary Figure S6 online). Moreover, an ATAC-seq (assay for transposase accessible chromatin with high-throughput sequencing) analysis for accessible chromatin at TSS (transcription start site) regions confirmed for all downregulated genes a decrease in accessibility (Supplementary Figure S7 online).

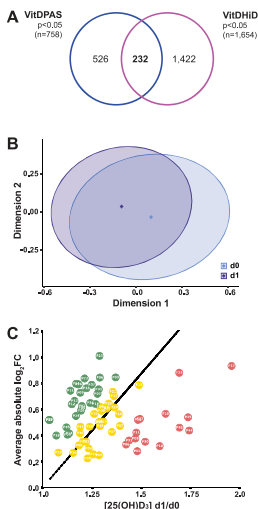


Fig. 3. Overlap and classification of vitamin D target genes between VitDPAS and VitDHID studies. **(A)** A Venn diagram highlights the intersection of 758 significant ($p < 0.05$) in vivo vitamin D target genes identified in the VitDPAS study and 1,654 target genes from the VitDHID study. **(B)** The quality of transcriptome datasets for the 232 shared target genes at d0 and d1 was assessed using dimensionality reduction, depicted in a MDS plot. **(C)** Based on the expression profiles of the 232 common target genes, the 70 participants across both studies were categorized into high (green), mid (yellow), and low (red) vitamin D responders.

The STRING database was utilized to explore potential functional relationships among the proteins encoded by the 26 shared genes (Supplementary Figure S8 online). This analysis, combined with a review of publicly available databases such as Human Protein Atlas and GeneCards, facilitated the classification of these genes/proteins into five distinct functional categories: (i) "inflammatory response and TNF (tumor necrosis factor)/NF κ B (nuclear factor κ B) signaling" includes *TNFAIP3* (TNF alpha-induced protein 3), *TNFSF14* (TNF superfamily member 14), *NFKBIA* (NF κ B inhibitor alpha), *NFKBIZ* (NF κ B inhibitor zeta), *PELI1* (pellino E3 ubiquitin protein ligase 1), and *ETS2* (ETS proto-oncogene 2, transcription factor); (ii) "immune cell migration and survival" comprises *S1PR1* (sphingosine-1-phosphate receptor 1), *CXCR4* (C-X-C motif chemokine receptor 4), *CX3CR1* (C-X3-C motif chemokine receptor 1), and *GIMAP6* (GTPase, IMAP family member 6); (iii) "transcriptional regulation" contains *PER1* (period circadian regulator 1), *CSRNPI* (cysteine and serine-rich nuclear protein 1), *KLF10* (KLF transcription factor 10), *KLF11*, and *HIF1A* (hypoxia-inducible factor 1 subunit alpha); (iv) "stress response and apoptosis" consists of *RBM7* (RNA binding motif protein 7), *TSC22D3* (TSC22 domain family member 3), *CDKN1A* (cyclin-dependent kinase inhibitor 1A), *HERPUD1* (homocysteine-inducible ER protein with ubiquitin-like domain 1), and *GADD45B* (growth arrest and DNA damage-inducible beta); (v)

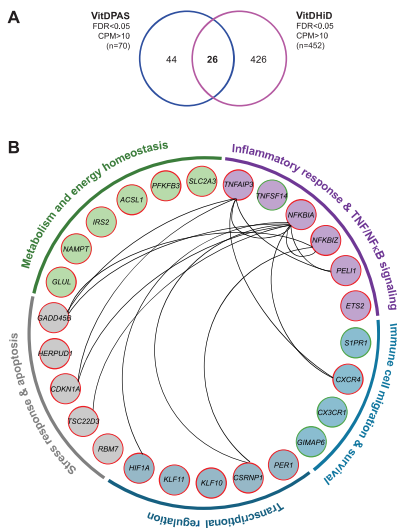


Fig. 4. Overlap and functional classification of vitamin D target genes between VitDPAS and VitDHID studies. (A) A Venn diagram illustrates the overlap between 70 highly significant (FDR < 0.05) in vivo vitamin D target genes identified in the VitDPAS study and 452 target genes from the VitDHID study. (B) The 26 shared vitamin D target genes are grouped into distinct functional categories. Gene interactions, derived from the STRING database (Supplementary Figure S8 online), are depicted with connecting lines. Upregulated genes are shown in green, while downregulated genes are represented in red.

"metabolism and energy homeostasis" embraces *GLUL* (glutamate-ammonia ligase), *NAMPT* (nicotinamide phosphoribosyltransferase), *IRS2* (insulin receptor substrate 2), *ACSL1* (acyl-CoA synthetase long-chain family member 1), *PFKFB3* (6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3), and *SLC2A3* (solute carrier family 2 member 3) (Fig. 4B). Interestingly, based on EnrichR³⁶ analysis of the 232 common target genes between VitDPAS and VitDHiD (Fig. 3A), TNF signaling is the top ranking pathway in the database KEGG (Kyoto Encyclopedia of Genes and Genomes)³⁷. Furthermore, according to STRING analysis (Supplementary Figure S8 online), NFKB1A, HIF1A, TNFAIP3, GADD45B, CXCR4, NFKB1Z, CDKN1A, and CSRN1 represent the most well-characterized core within this protein set. NFKB1A, NFKB1Z, and TNFAIP3 function downstream of stress sensors, including GADD45D (responsive to DNA damage), CDKN1A (cell cycle regulation), HIF1A (hypoxia response), and CXCR4 (immune activation) (Fig. 4B). Interestingly, in the majority of healthy individuals across both studies (84.3%), the genes encoding these proteins are downregulated in response to vitamin D₃ supplementation, a trend that also holds for the average expression of the full set of 26 genes (Supplementary Figure S5 online). However, in a subset of 7 individuals, this gene set is upregulated, indicating an opposite effect of vitamin D₃ for these persons. In summary, the functional classification of 26 common highly significant target genes emphasizes a range of biological roles influenced by vitamin D₃ supplementation. A central outcome of this regulation is the downregulation of vitamin D target genes involved in inflammatory responses, primarily through TNF and NFκB signaling pathways.

Discussion

The primary aim of this study was to validate and expand the concept of the vitamin D response index in a European cohort distinct from the one where it was originally developed³². Unlike other studies that examined gene expression changes following vitamin D₃ supplementation over an 8-week period³⁸, the VitDPAS study employed a different approach: administering a single vitamin D₃ bolus and taking a second measurement just one day later. This methodology facilitates direct comparison with *in vitro* studies, which traditionally use 24-hour stimulations with 1,25(OH)₂D₃³⁹. By bridging these experimental paradigms, the VitDPAS findings offer valuable insights into vitamin D's gene-regulatory effects on immune function in healthy individuals, potentially advancing our understanding in this area.

The VitDPAS study described here, conducted with a Polish cohort, represents an advanced extension of our previous VitDHiD study³⁸ performed in Kuopio, Finland. Both studies included healthy participants; however, the 45 VitDPAS participants were, on average, 12.1 years older and had a 1.1 kg/m² higher BMI compared to the 25 VitDHiD participants. Despite these differences, the distribution of high, mid, and low vitamin D responders was similar across both studies, with rates of 37.8%, 42.2%, and 20% in VitDPAS, closely aligning with 36%, 48%, and 16% in VitDHiD, respectively. Interestingly, when combining data from the 70 participants of both studies, the overall response rates were 40% high responders, 38.6% mid responders, and 21.4% low responders. For comparison, a similarly designed study conducted with 50 male and 50 female participants in Jeddah, Saudi Arabia (21°N)^{40,41}, identified 22% high responders, 39% mid responders, and 39% low responders⁴². This indicates that the proportion of low responders in a population living closer to the equator is nearly double that observed in Northern or Central European populations.

The regions of the two vitamin D intervention studies, Olsztyn (53°N) and Kuopio (63°N), experience extended "vitamin D winters"^{43,44} lasting approximately 4 and 5 months, respectively. These are periods during which endogenous vitamin D₃ production is not possible. This is particularly critical for individuals classified as low vitamin D responders, who require adequate vitamin D₃ supplementation during the winter months to prevent health issues associated with vitamin D deficiency⁴⁵. Although, daily vitamin D₃ supplementation of up to 4000 IU is considered safe⁴⁶, we recommend VitDPAS participants to adjust their daily intake during the winter months according to their body weight, ensuring it does not exceed 40 IU of vitamin D₃/kg. The VitDPAS study is currently ongoing, with participants being supplemented and monitored over a total duration of two years.

Previous reports have highlighted significant interindividual variability in the sets of *in vivo* vitamin D target genes³⁸. The transcriptome analysis of the 45 participants in the VitDPAS study supports this finding. Out of the 8,397 protein-coding genes expressed at a reasonable level (CPM > 10) in the PBMCs of all participants, 758 (9.0%) exhibited a significant ($p < 0.05$) response to vitamin D. However, only 232 of these vitamin D target genes (30.6%) overlapped with those identified in the VitDHiD study. While the overlap increases to 37.1% when using more stringent statistical criteria (FDR < 0.05), this still suggests that only a small subset of *in vivo* vitamin D targets (26 genes) is consistently identified across both studies with high confidence.

An intriguing observation from the VitDPAS study is that, although the initial list of 758 vitamin D target genes included an approximately equal number of upregulated and downregulated genes, only 4 of the 26 genes in the final shortlist were upregulated. This finding suggests that the most critical targets of vitamin D in healthy individuals are more likely to be downregulated than elevated. This aligns with previous findings from the VitDHiD study³⁸, which indicated that vitamin D-triggered pathways are predominantly downregulated or maintained in equilibrium, rather than being actively upregulated.

In a previous study utilizing VitDHiD data, the genes *NFKB1A*, *NFKB1Z*, and *TNFAIP3* were identified as top-ranking players in eight signaling pathways related to innate immunity⁴⁸. Interestingly, the downregulation of *NFKB1A* and *NFKB1Z*, both inhibitors of NFκB, promotes the translocation of the transcription factor NFκB from the cytoplasm to the nucleus, thereby activating multiple proinflammatory target genes. Moreover, downregulation of *TNFAIP3*, an upstream inhibitor of TNF-mediated inflammatory pathways, may further amplify the inflammatory response. These findings suggest that vitamin D enhances the sensitivity of inflammatory signaling pathways to external triggers, thereby facilitating a more efficient response when necessary.

In conclusion, the VitDPAS study validated and expanded upon the principles of the vitamin D response index. It underscored the need for special attention to low vitamin D responders within this Polish cohort,

ensuring adequate vitamin D₃ supplementation during the winter months. The combined analysis of the VitDPA and VitDHD studies identified 26 high-confidence vitamin D target genes, whose core function is the regulation of inflammatory responses in response to external stress sensors.

Data availability

Fastq files of the raw data can be found at Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo>) with accession number GSE278885.

Received: 20 November 2024; Accepted: 13 May 2025

Published online: 19 May 2025

References

- Holick, M. F., MacLaughlin, J. A. & Doppelt, S. H. Regulation of cutaneous previtamin D₃ photosynthesis in man: skin pigment is not an essential regulator. *Science* **211**, 590–593. <https://doi.org/10.1126/science.6256855> (1981).
- Martens, P. L., Gysmans, C., Verstuyf, A. & Mathieu, A. C. Vitamin D's effect on immune function. *Nutrients* **12** <https://doi.org/10.3390/nu12051248> (2020).
- Verway, M. et al. Vitamin D induces interleukin-1 β expression: paracrine macrophage epithelial signaling controls M. tuberculosis infection. *PLoS Pathog.* **9**, e1003407. <https://doi.org/10.1371/journal.ppat.1003407> (2013).
- Dimitrov, V. et al. Vitamin D-regulated gene expression profiles: species-specificity and cell-specific effects on metabolism and immunity. *Endocrinology* **162** <https://doi.org/10.1210/endo.bqa218> (2021).
- Bishop, E., Ismailova, A., Dimeloe, S. K., Hewison, M. & White, J. H. Vitamin D and immune regulation: antibacterial, antiviral, anti-inflammatory. *JBRMR* **5**, e10405. <https://doi.org/10.1002/jbrm.4.10405> (2021).
- Harrison, S. R. et al. Autoimmune disease and rheumatoid arthritis. *Calcif Tissue Int.* **106**, 58–75. <https://doi.org/10.1007/s00223-019-00577-2> (2020).
- Sunycz, J. A. The use of calcium and vitamin D in the management of osteoporosis. *Ther. Clin. Risk Manag.* **4**, 827–836. <https://doi.org/10.2167/term.v4i352> (2008).
- Uchimoto, R., Oyabu, M. & Kamet, Y. Vitamin D and sarcopenia: potential of vitamin D supplementation in sarcopenia prevention and treatment. *Nutrients* **12** <https://doi.org/10.3390/nu12103189> (2020).
- Feige, J. et al. Vitamin D supplementation in multiple sclerosis: a critical analysis of potentials and threats. *Nutrients* **12** <https://doi.org/10.3390/nu12030783> (2020).
- Vaughan-Shaw, P. G. et al. The effect of vitamin D supplementation on survival in patients with colorectal cancer: systematic review and meta-analysis of randomised controlled trials. *Br. J. Cancer*, **123**, 1705–1712. <https://doi.org/10.1038/s41416-020-01060-8> (2020).
- Bouillon, R. et al. The health effects of vitamin D supplementation: evidence from human studies. *Nat. Rev. Endocrinol.* **18**, 96–110. <https://doi.org/10.1038/s41574-021-00593-x> (2022).
- Gospodarska, E., Ghosh Dastidar, B. & Carlberg, C. Intervention approaches in studying the response to vitamin D₃ supplementation. *Nutrients* **15** <https://doi.org/10.3390/nu15133382> (2023).
- Whitfield, G. K. et al. Cloning of a functional vitamin D receptor from the lamprey (*Petromyzon marinus*), an ancient vertebrate lacking a calcified skeleton and teeth. *Endocrinology* **144**, 2704–2716. <https://doi.org/10.1210/en.2002-2211> (2003).
- Hausman, M. R. et al. Vitamin D receptor: molecular signaling and actions of nutritional ligands in disease prevention. *Nutr. Rev.* **66**, 98–112. <https://doi.org/10.1111/j.1753-4887.2008.00093.x> (2008).
- Chun, R. F., Liu, P. T., Modlin, R. L., Adams, J. S. & Hewison, M. Impact of vitamin D on immune function: lessons learned from genome-wide analysis. *Front. Physiol.* **5**, 151. <https://doi.org/10.3389/fphys.2014.00151> (2014).
- Hollis, B. W. Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: implications for Establishing a new effective dietary intake recommendation for vitamin D. *J. Nutr.* **135**, 317–322 (2005).
- Holick, M. F. et al. Evaluation, treatment, and prevention of vitamin D deficiency: an endocrine society clinical practice guideline. *J. Clin. Endocrinol. Metab.* **96**, 1911–1930. <https://doi.org/10.1210/nc.2011-0385> (2011).
- Pike, J. W., Meyer, M. B., Lee, S. M., Ornel, M. & Benkusky, N. A. The vitamin D receptor: contemporary genomic approaches reveal new basic and translational insights. *J. Clin. Invest.* **127**, 1146–1154. <https://doi.org/10.1172/JCI38887> (2017).
- GTEx Consortium. The Genotype-Tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585. <https://doi.org/10.1038/ng.2653> (2013).
- Campbell, M. J. Vitamin D and the RNA transcriptome: more than mRNA regulation. *Front. Physiol.* **5**, 181. <https://doi.org/10.3389/fphys.2014.00181> (2014).
- Seuter, S. et al. Molecular evaluation of vitamin D responsiveness of healthy young adults. *J. Steroid Biochem. Mol. Biol.* **174**, 314–321. <https://doi.org/10.1016/j.jsbmb.2016.06.003> (2017).
- Carlberg, C. & Haq, A. The concept of the personal vitamin D response index. *J. Steroid Biochem. Mol. Biol.* **175**, 12–17. <https://doi.org/10.1016/j.jsbmb.2016.12.011> (2018).
- Carlberg, C. & Mykko, M. P. Linking mechanisms of vitamin D signaling with multiple sclerosis. *Cells* **12** <https://doi.org/10.3390/cells12192391> (2023).
- Vukic, M. et al. Relevance of vitamin D receptor target genes for monitoring the vitamin D responsiveness of primary human cells. *PLoS ONE* **10**, e0124339. <https://doi.org/10.1371/journal.pone.0124339> (2015).
- Seuter, S. et al. Molecular evaluation of vitamin D responsiveness of healthy young adults. *J. Steroid Biochem. Mol. Biol.* **174**, 314–321 (2017).
- Hanel, A. et al. Common and personal target genes of the micronutrient vitamin D in primary immune cells from human peripheral blood. *Sci. Rep.* **10**, 21051. <https://doi.org/10.1038/s41598-020-78288-0> (2020).
- Ghosh Dastidar, B. et al. In vivo vitamin D targets reveal the upregulation of focal adhesion-related genes in primary immune cells of healthy individuals. *Sci. Rep.* **14**, 17552. <https://doi.org/10.1038/s41598-024-68741-9> (2024).
- Scragg, R. et al. Effect of monthly High-Dose vitamin D supplementation on cardiovascular disease in the vitamin D assessment study: A randomized clinical trial. *JAMA Cardiol.* **2**, 608–616. <https://doi.org/10.1001/jamacardio.2017.0175> (2017).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21. <https://doi.org/10.1093/bioinformatics/btt653> (2013).
- Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930. <https://doi.org/10.1093/bioinformatics/btt656> (2014).
- Durbin, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191. <https://doi.org/10.1038/nprot.2009.97> (2009).
- 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. <https://doi.org/10.1093/bioinformatics/btp616> (2010).

33. Chen, Y., Lun, A. T. & Smyth, G. K. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research* 5, 1438. <https://doi.org/10.12688/f1000research.8987.2> (2016).
34. Carlberg, C. et al. Primary vitamin D target genes allow a categorization of possible benefits of vitamin D₃ supplementation. *PLoS ONE* 8, e71042. <https://doi.org/10.1371/journal.pone.0071042> (2013).
35. Szklarczyk, D. et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 51, D638–D646. <https://doi.org/10.1093/nar/gkac1000> (2023).
36. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* 14 <https://doi.org/10.1186/1471-2105-14-128> (2013).
37. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. <https://doi.org/10.1093/nar/gkw1092> (2017).
38. Hossain-nezhad, A., Spira, A. & Holick, M. F. Influence of vitamin D status and vitamin D₃ supplementation on genome wide expression of white blood cells: a randomized double-blind clinical trial. *PLoS ONE* 8, e58725. <https://doi.org/10.1371/journal.pone.0058725> (2013).
39. Hanel, A. & Carlberg, C. Time-resolved gene expression analysis monitors the regulation of inflammatory mediators and Attenuation of adaptive immune response by vitamin D. *Int. J. Mol. Sci.* 23 <https://doi.org/10.3390/ijms23020911> (2022).
40. AlGhamdi, S. A. et al. A single oral vitamin D₃ bolus reduces inflammatory markers in healthy Saudi males. *Int. J. Mol. Sci.* 23 <https://doi.org/10.3390/ijms231911992> (2022).
41. Alsaifiani, H. M. et al. A single vitamin D₃ bolus supplementation improves vitamin D status and reduces Proinflammatory cytokines in healthy females. *Nutrients* 14 <https://doi.org/10.3390/nu14193963> (2022).
42. AlGhamdi, S. A. et al. Evaluation of the vitamin D response index in a Saudi cohort. *Saudi Pharm. J.* 32, 102137. <https://doi.org/10.1016/j.sjps.2024.102137> (2024).
43. Ohlund, I., Lind, T., Hernell, O. & Silfverdal, S. A. Karlstad Akesson, P. Increased vitamin D intake differentiated according to skin color is needed to Meet requirements in young Swedish children during winter: a double-blind randomized clinical trial. *Am. J. Clin. Nutr.* 106, 105–112. <https://doi.org/10.3945/ajcn.116.147108> (2017).
44. Khanna, T. et al. Comprehensive analysis of seasonal and geographical variation in UVB radiation relevant for vitamin D production in Europe. *Nutrients* 14 <https://doi.org/10.3390/nu14235189> (2022).
45. Holick, M. F. Vitamin D deficiency. *N Engl. J. Med.* 357, 266–281. <https://doi.org/10.1056/NEJMe070553> (2007).
46. Johnson, K. C. et al. Safety and tolerability of high-dose daily vitamin D₃ supplementation in the vitamin D and type 2 diabetes (D2d) study: a randomized trial in persons with prediabetes. *Eur. J. Clin. Nutr.* 76, 1117–1124. <https://doi.org/10.1038/s41430-022-01068-8> (2022).
47. Jaroslawska, J. & Carlberg, C. In vivo regulation of signal transduction pathways by vitamin D stabilizes homeostasis of human immune cells and counteracts molecular stress. *Int. J. Mol. Sci.* 24, 14632. <https://doi.org/10.3390/ijms241914632> (2023).
48. Jaroslawska, J., Ghosh Dastidar, R. & Carlberg, C. In vivo vitamin D target genes interconnect key signaling pathways of innate immunity. *PLoS ONE* 19, e0306426. <https://doi.org/10.1371/journal.pone.0306426> (2024).

Acknowledgements

Kind thanks to the Gene Core Facility at the EMBL in Heidelberg, Germany, for massively parallel sequencing services, the Centre of Informatics Tricity Academic Supercomputer and Network of the Technical University of Gdansk (<https://task.gda.pl/en>) for the use of their supercomputers for sequence alignment and Procter & Gamble Health Poland for donating Viganotellen Max 4000 to our study.

Author contributions

The VitDPAS study was designed and conducted by E.G., M.Ra., K.T.M., J.R., and C.C.. Data analysis and visualization were performed by R.G.D., M.Ry. and J.J., while RNA isolation and RNA-seq library preparation were carried out by E.G. and C.C.. C.C. refined the figures and drafted the manuscript, which was reviewed and approved by all authors.

Funding

This publication is part of a project that has received funding from the European Union's Horizon2020 research and innovation program under grant agreement no. 952601, from the David and Amy Fulton Foundation, Seattle, US and from the National Science Centre (Poland), project number 2023/49/B/NZ9/00402.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-02495-w>.

Correspondence and requests for materials should be addressed to C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

Dynamic system identification defines a minimal transcription factor control architecture of human endocrine gene regulation

Ranjini Ghosh Dastidar¹, Emilia Gospodarska¹, Kornelia Tokarczyk-Malesa², Jerzy Romaszko², Elisa Domínguez-Hüttinger³, and Carsten Carlberg^{1,4#}

¹ InLife Institute of Animal Reproduction and Food Research, Polish Academy of Sciences, Olsztyn, Poland

² Department of Family Medicine and Infectious Diseases, School of Medicine, Collegium Medicum, University of Warmia and Mazury, Olsztyn, Poland

³ Departamento de Biología Molecular y Biotecnología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Ciudad México, México

⁴ Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland

#Corresponding author: Prof. Carsten Carlberg

InLife Institute of Animal Reproduction and Food Research

Polish Academy of Sciences

ul. Trylińskiego 18, 10-683 Olsztyn, Poland

Tel.: +48-604-336-461

E-mail: c.carlberg@pan.olsztyn.pl

ABSTRACT

Endocrine signals orchestrate transcriptional programs across diverse physiological contexts, yet the system-level regulatory architecture governing these responses in humans remains poorly characterized. Here, we integrate longitudinal transcriptome profiling from controlled vitamin D₃ intervention studies with pharmacokinetic modeling and dynamic regulatory network inference to quantitatively investigate endocrine-responsive gene regulation *in vivo*. Applying ordinary differential equation–based system identification to dense temporal gene expression trajectories, we infer a compact transcription factor control architecture comprising 11 regulators and 117 interactions, centered on TP53, MYC, AP1 components, STAT3, and HIF1A, that reproduces observed transcriptional dynamics across supplementation cycles and interindividual response phenotypes. The inferred regulatory framework demonstrates predictive performance when evaluated in an independent intervention cohort, indicating robustness and transferability of the identified control structure. Rather than establishing direct causality for individual regulatory interactions, this study provides a quantitatively interpretable and experimentally testable model of transcription factor–mediated endocrine gene regulation. Conceptually, our work establishes the feasibility of integrating physiological perturbation, longitudinal molecular profiling, and dynamic systems modeling to infer regulatory network organization in humans. This framework offers a generalizable strategy for mechanistic investigation of gene regulatory systems beyond static or correlative omics analyses.

Keywords: Vitamin D; vitamin D₃ supplementation; transcriptomics; gene regulatory network; ordinary differential equations; systems biology; precision nutrition; interindividual variability.

Abbreviations

1,25(OH) ₂ D ₃	1 α ,25-dihydroxyvitamin D ₃
25(OH)D ₃ (VitD)	25-hydroxyvitamin D ₃
BMI	body mass index
CPM	counts per million
CSRNP1	cysteine and serine rich nuclear protein 1
EGR1	early growth response 1
ETS	ETS proto-oncogene, transcription factor
FC	fold change
FDR	false discovery rate
FOS	Fos proto-oncogene, AP1 transcription factor subunit
HGNC	human genome nomenclature committee
HIF1A	hypoxia inducible factor 1 subunit alpha
JUNB	JunB proto-oncogene, AP1 transcription factor subunit
KLF10	Krüppel like factor 10
MDS	multidimensionality scaling
MOR	mode of regulation
MYC	MYC proto-oncogene, bHLH transcription factor
NFKB1	nuclear factor kappa B subunit 1
NR4A2	nuclear receptor subfamily 4 group A member 2
ODE	ordinary differential equation
PBMC	peripheral blood mononuclear cell
RNA-seq	RNA sequencing

STAT	signal transducer and activator of transcription
TMM	trimmed mean of M-values
TP53	tumor protein P53
VDR	vitamin D receptor

Introduction

Vitamin D signaling represents one of the most extensively characterized endocrine transcriptional regulatory systems (1), and provides a paradigmatic model system for endocrine control of transcription factor network dynamics in human physiology (2). However, its quantitative system-level regulatory architectures in healthy humans remain largely unresolved. The biologically active metabolite of vitamin D₃, 1 α ,25-dihydroxyvitamin D₃, mediates genomic actions through the vitamin D receptor (VDR) (3,4), a ligand-activated transcription factor expressed in most human tissues (5). While the classical role of vitamin D in calcium homeostasis and skeletal health is well established (6), accumulating evidence indicates that vitamin D signaling also contributes to the regulation of immune function, cellular metabolism, and stress responses (7-9). However, despite extensive epidemiological and experimental research, the molecular mechanisms underlying vitamin D-dependent physiological effects in healthy individuals remain incompletely understood. In particular, it remains unclear whether endocrine transcriptional responses *in vivo* can be described by dynamically identifiable regulatory network structures capable of predictive modeling. Large randomized intervention trials have yielded heterogeneous and often inconclusive clinical outcomes of vitamin D₃ supplementation, particularly in populations without severe deficiency (10,11). A central limitation of large-scale vitamin D intervention trials is the implicit assumption of homogeneous molecular responsiveness across individuals (12). Emerging molecular data suggest instead that vitamin D signaling exhibits substantial interindividual variability, affecting both transcriptional regulation and downstream cellular functions (13). This variability has led to the proposal of the “vitamin D response index”, a concept describing the capacity of an individual to translate systemic vitamin D status into measurable molecular responses (14).

Human intervention studies provide a powerful experimental framework for investigating nutrient-dependent gene regulation under physiologically relevant

conditions (15). Previous transcriptome-based studies of vitamin D₃ supplementation have demonstrated that immune cells are particularly sensitive to vitamin D-mediated regulatory effects (16-18). These studies identified both common and individual-specific vitamin D target genes, many of which are associated with inflammatory signaling pathways (19). Quantitative inference of gene regulatory network dynamics from human intervention data remains challenging due to limited temporal resolution and high-dimensional regulatory complexity. While computational approaches have successfully reconstructed static transcriptional networks from population-scale datasets (20), dynamically identifiable regulatory architectures underlying physiological perturbation responses are rarely resolved *in vivo*. Establishing such mechanistic regulatory frameworks is essential for linking endocrine signaling to predictive transcriptional system behavior.

Unlike previous intervention studies that primarily described transcriptional responses (21,22), the present work integrates longitudinal perturbation data with dynamic regulatory modeling to identify mechanistic determinants of interindividual vitamin D responsiveness. Here we move beyond descriptive transcriptomics to intervention-driven dynamic system identification of endocrine transcriptional control. We aim to define the dynamic regulatory architecture underlying human molecular responsiveness to repeated vitamin D₃ supplementation. By integrating longitudinal immune transcriptomics with pharmacokinetic modeling and data-driven reconstruction of transcription factor networks, we sought to quantitatively connect systemic vitamin D exposure with regulatory system dynamics. We investigated (i) whether repeated bolus interventions induce temporal transcriptional reprogramming, (ii) whether response phenotypes reflect stable individual system properties, and (iii) which regulatory nodes control divergence in response trajectories. Using ordinary differential equation (ODE)-based modeling of a reduced regulatory kernel, defined as the minimal dynamically controllable core governing system behavior, we sought to translate intervention-derived transcriptomic observations into a predictive mechanistic

framework for nutrient-dependent gene regulation *in vivo*. In this way, we establish intervention-driven dynamic system identification as a generalizable strategy for quantitatively linking endocrine nutrient exposure to transcriptional regulatory control in humans.

Material and Methods

Sample collection

The VitDPAS intervention study (ClinicalTrials.gov identifier: NCT06104111) (23) started in November 2023 in Olsztyn, Poland. Eligible participants were 20-65 years of age with a body mass index (BMI) between 18.5 and 29.9 kg/m². Exclusion criteria included any self-reported disease or infection within two weeks prior to the planned blood draw; severe visual or hearing impairment; gastrointestinal disorders (e.g., celiac disease or irritable bowel syndrome); chronic kidney disease or kidney stones; chronic liver disease; pathological bone fractures or osteoporotic fractures within the previous five years; current cancer diagnosis; treatment affecting calcium metabolism; neurological disorders associated with recurrent symptoms (e.g., Parkinson's disease, post-stroke hemiplegia, epilepsy, recurrent dizziness, or syncope); diseases that may elevate serum calcium levels (e.g., sarcoidosis, lymphoma, or hyperparathyroidism); use of medications known to affect 25-hydroxyvitamin D₃ (25(OH)D₃) levels (e.g., phenobarbital, carbamazepine, or phenytoin); Paget's disease of bone; inflammatory joint diseases including rheumatoid arthritis, Reiter's syndrome, or psoriatic arthritis; participation in other clinical studies within the previous four months (except observational studies); known infection with HIV, hepatitis viruses, or other infectious diseases; substance abuse (drug dependence, alcoholism, or tobacco smoking); mental illness; homeostatic disorders that could complicate blood collection; and pregnancy, breastfeeding, or planned pregnancy. The study protocol was approved by the Ethics Committee of the Olsztyn Chamber of Doctors (approval number 31/2023/VIII). All participants provided written informed consent, and all procedures were conducted in accordance with relevant guidelines and regulations.

The VitDPAS study recruited 45 participants (23), but four individuals did not provide blood samples at all seven scheduled time points. Consequently, the final cohort comprised 41 participants (20 females and 21 males; age range 24-64 years; BMI 19.3-29.5 kg/m²; **Table S1**). Participants received a vitamin D₃ bolus of 1,000 IU/kg body weight with breakfast on study days 0, 28, and 56, corresponding to a monthly supplementation regimen over three months. Depending on body weight, this corresponded to a dose range of 52,000-110,000 IU vitamin D₃ per bolus, consistent with established safety profiles of intermittent high-dose vitamin D₃ supplementation. For comparison, other intervention studies, such as ViDA (24), administered monthly doses of 100,000 IU vitamin D₃ for periods exceeding two years.

Peripheral blood samples were collected immediately before supplementation and 24 hours after each vitamin D₃ administration (days 1, 29, and 57), as well as at day 84 as a reference time point. Samples were used for serum preparation and isolation of peripheral blood mononuclear cells (PBMCs). Based on previous vitamin D intervention studies, including VitDbol (NCT02063334) (25,26) and VitDHiD (NCT03537027) (18), the 24-hour post-supplementation time point represents an optimal window for detecting transcriptional responses to vitamin D. Serum 25(OH)D₃ concentrations were measured using an electrochemiluminescence binding assay on a Cobas Pure Immunoassay Analyzer (Roche) in the analytical laboratory of the Municipal Polyclinical Hospital in Olsztyn, Poland.

PBMC isolation

PBMCs from the 41 participants of the VitDPAS study were processed *ex vivo* without prior *in vitro* stimulation. Following the collection of 8 ml of peripheral blood on days 0, 1, 28, 29, 56, 57, and 84 PBMCs were isolated within one hour using Vacutainer CPT Cell Preparation Tubes with sodium citrate (Becton Dickinson) according to the manufacturer's protocol. After isolation, cells were washed with phosphate-buffered saline, aliquoted at a concentration of 4×10^6 cells per ml, and stored at -80 °C for subsequent RNA isolation.

Transcriptome analysis

Total RNA was extracted from PBMCs using the High Pure RNA Isolation Kit (Roche) according to the manufacturer's instructions. RNA quality was assessed with an Agilent TapeStation. RNA-seq libraries were prepared after rRNA depletion using kits and protocols from New England Biolabs and sequenced as 75-bp reads on a NextSeq 2000 system (Illumina) at the EMBL GeneCore facility (Heidelberg, Germany) following standard protocols. Sequencing quality was assessed with FastQC (v0.11.9; www.bioinformatics.babraham.ac.uk/projects/fastqc) (**Table S2**). Reads were aligned to the human reference genome GRCh38 (Ensembl v111.38) using the STAR aligner (27) (v2.7.2a), and gene-level counts were generated with FeatureCounts (28) (v2.0.1) using default parameters. To standardize gene nomenclature, Human Gene Nomenclature Committee (HGNC) symbols were updated using the R package HGNCHELPER (v0.8.1). Gene annotation, including identifiers, descriptions, genomic coordinates, and biotypes, was obtained from the Ensembl database (release 100) using the R package BiomaRt (29) (v2.42.1). Entrez Gene identifiers were added using the R package org.Hs.eg.db (v3.10.0), and incomplete mappings were manually verified using the NCBI Gene database. Genes lacking genomic position information or encoded by the mitochondrial genome were excluded from further analysis. Downstream analyses were restricted to protein-coding genes.

Differential gene expression analysis

Differential gene expression analysis was performed in R (v4.3.1) on macOS 13 (Ventura) using the edgeR package (30) (v4.0.16). To ensure robust variance estimation in longitudinal generalized linear modeling, analyses were restricted to protein-coding genes with reliable expression levels. Low-abundance transcripts were excluded using an average expression threshold of CPM > 10 implemented via the `filterByExpr()` function in edgeR. Genes with zero expression in any participant at any time point were excluded to ensure stable variance estimation. This filtering strategy reduces statistical noise arising from sparse count distributions while preserving genes

with sufficient information content for repeated-measures inference. Library sizes were subsequently normalized using the trimmed mean of M-values (TMM) method. Six pairwise comparisons were performed: three short-term comparisons (d1 versus d0, d29 versus d28, d57 versus d56) and three long-term comparisons relative to baseline (d28 versus d0, d56 versus d0, d84 versus d0). Differential expression was assessed using a paired generalized linear model quasi-likelihood (GLM-QL) design (31). Genes were considered significantly regulated at $p < 0.05$, with false discovery rate (FDR) control applied using the Benjamini-Hochberg method ($FDR < 0.05$) (**Table S3**). The glmTreat approach was applied using a minimum absolute $\log_2$ fold change ($\log_2FC$) threshold of $\log_2(1.1)$. Gene overlaps across comparisons were visualized using Venn and Euler diagrams implemented in the ggvenn and eulerr R packages.

Classification of study participants

Participants were classified as high, mid, or low vitamin D responders based on three independent supplementation cycles: R1 (d0-d1), R2 (d28-d29), and R3 (d56-d57). For each supplementation cycle, participants were embedded in a two-dimensional response space where the x-coordinate represented the post/pre-supplementation ratio of serum 25(OH)D₃ level (d1/d0, d29/d28, or d57/d56) and the y-coordinate represented the average absolute $\log_2FC$ of the 23 genes that were significantly regulated ($p < 0.05$) across all three short-term supplementation cycle comparisons. A trendline was constrained to pass through the epicenter ($x = 1$, $y = 0$), and the signed perpendicular distance of each participant from this line was calculated. These one-dimensional distances were then used for *k-means* clustering ($k = 3$, $n_init = 20$; scikit-learn, Python 3.9.13) to assign participants to a response group for each supplementation cycles. To assess the stability of these classifications, each participant's assigned group was compared against their previously reported classification from our earlier cohort study (23). Participants whose classification was in agreement across both analyses were designated as stable responders, while those whose classification differed were designated as unstable.

Pharmacokinetic modeling of 25(OH)D₃ kinetics

In a previous intervention study (32), serum 25(OH)D₃ concentrations were measured at 0, 24, and 48 hours following monthly vitamin D₃ bolus administration. These data were used to parameterize kinetic models for the three supplementation cycles. Inter-bolus decay kinetics was modeled as exponential decay, $C(t) = C_0 \cdot \exp(-d \cdot t)$, where d represents the decay rate derived analytically. The supplementation phase was described by the differential equation $dC/dt = k - d \cdot C$, where k denotes the absorption rate constant. Parameter k was estimated by minimizing the residual sum of squares across the observed time points. Serum 25(OH)D₃ measurements from the VitDPAS cohort were overlaid onto the fitted curves to verify that the observed values fell within the modeled range. Kinetic modeling and visualization were performed in R (v4.4.2) and independently verified in Julia (v1.11.4) using the DifferentialEquations.jl package.

Gene regulatory network construction and *-reduction*

A gene regulatory network (33) was constructed using the human regulon database DoRothEA (34) via the dorothea R package. To ensure biological interpretability and dynamic identifiability, regulatory interactions were restricted to high-confidence curated transcription factor-target relationships. Structural identifiability of the dynamic regulatory system was enforced through iterative *kernel reduction* (35). Seven common primary genes, *CSRNP1* (cysteine and serine rich nuclear protein 1), *ETS2* (ETS proto-oncogene 2, transcription factor), *FOS* (Fos proto-oncogene, AP1 transcription factor subunit), *HIF1A* (hypoxia inducible factor 1 subunit alpha), *JUNB* (JunB proto-oncogene, AP1 transcription factor subunit), *KLF10* (Krüppel like factor 10), and *NR4A2* (nuclear receptor subfamily 4 group A member 2) were selected as transcriptional regulators that were significantly regulated (FDR < 0.05) in all of the comparisons d1 versus d0, d28 versus d0, d56 versus d0, and d84 versus d0. Direct interactions involving these genes, together with the top 50 secondary interactions ranked by absolute mode of regulation (MOR) score, were compiled into a directed igraph network with activating (MOR > 0) and inhibitory (MOR < 0) edges.

Structural identifiability was enhanced by *kernel reduction* (**Figure S1**) and by fitting across repeated perturbation cycles, thereby constraining parameter degeneracy. In each iteration, nodes with in- or out-degree ≤ 2 were substituted by their regulatory influence, by introducing direct connections between upstream and downstream nodes. For each indirect path ($A \rightarrow B \rightarrow C$), a new edge ($A \rightarrow C$) was created with sign and magnitude determined by the product of the corresponding MOR values, preserving the regulatory logic of the original interactions. Vitamin D-mediated effects were propagated analogously. After each removal step, node degrees were recalculated, and the procedure was repeated until all remaining nodes exceeded the degree threshold. The resulting network kernel represents the minimal set of densely interconnected regulatory elements that cannot be further reduced under these criteria. This reduction approach preserves network structure while minimizing dimensionality for dynamic parameter estimation.

ODE-based modeling and parametric analysis

Kernel reduction yielded a minimal regulatory kernel composed of 11 genes connected by 117 regulatory interactions (71 activations, 46 repressions). For each gene, Hill-type ODEs were generated automatically: $d[\text{Gene}]/dt = (\alpha + \sum \text{activation terms}) \times (\prod \text{repression terms}) - \delta \cdot [\text{Gene}]$, where activation and repression term followed $\gamma \cdot X^n / (K^n + X^n)$ and $\gamma_r / (1 + (X/K_r)^{n_r})$, respectively (**Figure S2**). This network comprised 373 free parameters per supplementation cycle, including basal production (α), degradation (δ), activation and repression coefficients (γ , γ_r), half-saturation constants (K , K_r), and Hill exponents (n , n_r) (**Table S4**). Parameter values were obtained by simultaneous fitting across multiple time points and supplementation cycles. Serum 25(OH)D₃ (VitD) was modeled as a dynamical forcing term ($d\text{VitD}/dt = k - d \cdot \text{VitD}$), with synthesis (k) and decay (d) rates fixed to pharmacokinetic estimates and excluded from optimization.

Four independent models were trained for distinct participant groups: the full cohort ($n = 41$), stable high responders ($n = 8$), stable mid responders ($n = 8$), and low responders ($n = 6$) (**Table S1**). Mean normalized CPM values of the 11 genes across

six time points (d0, d1, d28, d29, d56 and d57) served as fitting targets. Parameters were optimized by differential evolution (BlackBoxOptim.jl), minimizing the sum of squared residuals between simulated ODE trajectories and observed mean expression profiles. Each model was fitted across three supplementation cycles using cycle-specific vitamin D pharmacokinetics. Optimized parameters were validated by simulating the inter-bolus decay periods, using d1 concentrations as initial conditions to predict d28 values for each supplementation cycle. Results were visualized with Plots.jl and StatsPlots.jl.

To identify determinants of differential response, optimized parameter sets were analyzed in R (v4.4.2). Gene importance was quantified using a composite score integrating (i) degree centrality, (ii) the sum of absolute outgoing regulatory coefficients, and (iii) vitamin D-weighted contributions for directly regulated genes. Genes were ranked and classified as hub, key, moderate, or peripheral regulators. Additionally, pairwise parameter ratios between high- and low-responder models were computed for each interaction to identify edges with the strongest divergence (**Table S4**).

Parameter sensitivity analysis and out-of-sample validation

To evaluate the robustness and effective dimensionality of the ODE framework, local sensitivity analysis was conducted by computing the relative change in model output following $\pm 5\%$ perturbations of each of the 373 optimizable parameters. Sensitivity measures were averaged across genes and supplementation cycles to obtain a global estimate of parameter influence on system dynamics. For out-of-sample validation, the VitDPAS-trained model was used to predict mean gene expression in PBMCs 24 hours after bolus administration in an independent cohort of 25 participants from the VitDHiD study (NCT03537027) (18), who received once a fixed vitamin D₃ dose of 80,000 IU. Cohort-level baseline gene expression and serum 25(OH)D₃ concentrations were used as initial conditions, while vitamin D pharmacokinetics were specified using the 48-

hour absorption model fitted in the VitDPAS dataset (32). Predictions were generated independently using the parameter sets of cycles R1, R2, and R3.

Software and code availability

All analyses were performed using R (v4.4.2, Pile of Leaves; macOS, aarch64-apple-darwin20), Python (v3.9.13), and Julia (v1.11.4). Key R packages included edgeR (v4.0.16), ggplot2, limma, dorothea, igraph, openxlsx, eulerr, and ggvenn. Python analyses used scikit-learn, pandas, numpy, matplotlib, and scipy. Julia-based modeling employed DifferentialEquations.jl, BlackBoxOptim.jl, Plots.jl, and StatsPlots.jl. Complete analysis scripts and reproducible pipelines are publicly available in three GitHub repositories. Scripts for differential gene expression analysis and data integration are available at https://github.com/ranjinighoshdastidar/VitDPAS-DGE_RNAseq (<https://doi.org/10.5281/zenodo.18929051>). Participant classification scripts are available at <https://github.com/ranjinighoshdastidar/VitDPAS-Classification> (<https://doi.org/10.5281/zenodo.18929070>). Scripts for pharmacokinetic modeling, gene regulatory network construction, ODE-based modeling, and parametric analysis are available at <https://github.com/ranjinighoshdastidar/VitDPAS-Modeling> (<https://doi.org/10.5281/zenodo.18929074>).

Results

Interindividual dynamics of serum 25(OH)D₃ responses to repeated weight-adjusted vitamin D₃ bolus supplementation

To establish a longitudinal physiological perturbation framework suitable for quantitative system identification of endocrine transcriptional regulation, the VitDPAS intervention study was conducted in 41 healthy participants receiving three cycles of weight-adjusted vitamin D₃ bolus supplementation (1,000 IU per kg body weight). The cohort comprised 20 females and 21 males with a mean age of 39.7 years and a mean body mass index (BMI) of 23.8 kg/m² (**Table S1**). Supplementation was administered on study days 0, 28, and 56, and peripheral blood samples were collected immediately

before and 24 hours after each intervention, as well as at day 84 to capture cumulative long-term effects (**Figure 1A**).

Baseline serum concentrations of 25(OH)D₃ exhibited pronounced interindividual variability, ranging from 9.0 to 54.8 ng/mL, with a cohort mean of 26.1 ng/mL. Following the first bolus administration, circulating levels increased to a range of 15.2-62.6 ng/mL (mean 32.9 ng/mL), indicating efficient systemic absorption of the administered dose. However, subsequent supplementation cycles revealed heterogeneous temporal response trajectories (**Figure 1B**). While the majority of participants exhibited cumulative increases in circulating 25(OH)D₃ concentrations, individuals with relatively high baseline levels frequently showed paradoxical reductions in vitamin D status over the 12-week intervention period. At the final time point (day 84), serum concentrations ranged from 18.3 to 51.0 ng/mL (mean 29.6 ng/mL), reflecting the interplay between exogenous supplementation and endogenous homeostatic regulation of vitamin D metabolism. Across the cohort, 26 participants exhibited a net increase in vitamin D status over the study period, eight showed minimal changes (<2 ng/mL), and seven displayed overall decreases in circulating concentrations. Each bolus administration resulted in an average increase of 5.9 ng/mL, whereas the mean cumulative increase after three supplementation cycles was only 3.5 ng/mL (**Table S1**), indicating that the systemic response to repeated perturbation is non-linear and subject to regulatory feedback.

To quantitatively characterize vitamin D kinetics and provide a dynamic input function for subsequent transcriptional modeling, pharmacokinetic analysis was performed using independent longitudinal data (32) that included additional sampling 48 hours after bolus administration (**Figure S3A**). These data enabled estimation of absorption rate constants ranging from 0.196 to 0.368 h⁻¹ (**Figure S3B**). Incorporation of these parameters into a simple exponential production-decay model demonstrated that the heterogeneous serum vitamin D trajectories observed in the VitDPAS cohort can be quantitatively reproduced using physiologically plausible kinetic assumptions.

In summary, these findings establish repeated vitamin D₃ supplementation as a controlled endocrine perturbation that generates individualized systemic exposure profiles, thereby providing a suitable experimental framework for analyzing transcriptional regulatory system dynamics in humans.

Progressive, temporally integrated transcriptome reorganization during sustained endocrine stimulation

To investigate the transcriptional consequences of repeated endocrine perturbation, RNA-seq was performed on PBMC samples collected at all seven study time points. After quality filtering and expression thresholding, 8,396 protein-coding genes were identified as consistently expressed across the cohort (**Table S3**). Multidimensional scaling (MDS) analysis revealed that acute transcriptional responses measured 24 hours after individual supplementation events were relatively modest, with minimal global transcriptome displacement compared to baseline samples (**Figure S4**, upper panels). In contrast, cumulative endocrine exposure induced progressively larger transcriptome-wide reorganization, as reflected by increasing separation of time points along principal axes of variation (**Figure S4**, lower panels).

Differential expression analysis confirmed this temporal structure. Acute responses following the first bolus were characterized by 757 regulated genes, whereas subsequent boluses elicited 153 and 873 regulated genes, respectively (**Figure S5**). Importantly, only 23 genes were consistently responsive across all acute comparisons, indicating limited overlap of immediate transcriptional responses between supplementation cycles. Application of multiple-testing correction (FDR < 0.05) further reduced the number of significantly regulated genes, with 70 genes remaining significant exclusively in the d1 versus d0 comparison, and no genes meeting this threshold after the second or third bolus (**Figure 2**, upper panel). By contrast, cumulative supplementation resulted in progressively stronger transcriptional responses. At weeks 4, 8, and 12, the numbers of significantly regulated genes increased to 698, 2,137, and 2,619, respectively (**Figure 2**, lower panel), indicating

that repeated endocrine stimulation leads to amplification and temporal integration of transcriptional responses.

Taken together, these observations demonstrate that transcriptional responses to vitamin D₃ supplementation are not dominated by immediate acute effects but instead emerge through cumulative regulatory processes that progressively reshape the transcriptional landscape. This temporal pattern is consistent with models in which endocrine signaling induces stepwise reconfiguration of transcription factor networks and chromatin states, thereby enabling integrative transcriptional adaptation to sustained physiological stimulation.

Quantitative classification of interindividual transcriptional responsiveness

The limited overlap of acutely regulated genes across supplementation cycles (**Figure S5**) suggested substantial interindividual variability in transcriptional responsiveness to vitamin D. To quantify this heterogeneity in a systematic and reproducible manner, we developed an integrated response metric that combined transcriptional and systemic exposure parameters. Specifically, for each participant and supplementation cycle, the mean absolute log₂FC across the 23 consistently responsive genes was calculated and combined with the ratio of post- to pre-supplementation serum 25(OH)D₃ concentrations. Projection of individuals into this two-dimensional response space revealed structured but dynamic response trajectories across the intervention period (**Figure S6**). These trajectories reflected both transcriptional response magnitude and endocrine exposure efficiency and showed that response phenotypes were not static but evolved across supplementation cycles.

To derive a consolidated measure of individual responsiveness, response coordinates were aggregated across all cycles to generate a composite vitamin D response index. Unsupervised clustering based on this index identified three response phenotypes: high responders ($n = 19$), mid responders ($n = 16$), and low responders ($n = 6$) (**Figure 3**; **Table S1**). Baseline vitamin D status differed significantly between

response groups, with low responders generally exhibiting lower initial serum 25(OH)D₃ concentrations. In contrast, demographic parameters including age, sex, and BMI showed no significant association with response classification. These findings indicate that transcriptional responsiveness to vitamin D represents an intrinsic dynamical property of the transcriptional regulatory system that cannot be explained solely by differences in systemic exposure or demographic characteristics.

Furthermore, longitudinal tracking of individual response trajectories demonstrated that most participants maintained consistent relative positioning within the response space, supporting the notion that vitamin D responsiveness reflects a stable yet plastic regulatory phenotype. This classification provided the basis for subsequent mechanistic modeling aimed at identifying regulatory determinants of response heterogeneity.

Reconstruction of a vitamin D-responsive transcription factor regulatory network

To elucidate regulatory mechanisms underlying cumulative transcriptional responses, we focused on genes that were significantly regulated at all long-term comparison time points. Intersection analysis across the d1, d28, d56, and d84 comparisons yielded a set of 32 persistently regulated genes (**Figure S7A**). Given that secondary transcriptional effects are typically mediated through transcription factor cascades (36), subsequent analysis was restricted to transcription factor-encoding genes within this set. This filtering identified seven candidate regulators: *CSRNP1*, *ETS2*, *FOS*, *HIF1A*, *JUNB*, *KLF10*, and *NR4A2* (**Figure S7B**).

These transcription factors were used as seed nodes for construction of a regulatory interaction network based on curated transcription factor-target relationships obtained from the DoRothEA database (34) (**Figure 4A**). Initial network expansion incorporating first-order regulatory interactions resulted in a system comprising the six primary vitamin D-responsive transcription factors (37,38) and 34 downstream interacting regulators (**Figure S8A; Figure 4B**). Network refinement involved removal of nodes

lacking supported regulatory connections within the reconstructed system. In particular, *CSRNP1* (39) and *KLF10* were excluded due to absence of validated interactions, resulting in a connected transcription factor network structure (**Figure 4C**).

Kernel reduction identifies a dynamically controllable regulatory core

To enable mechanistic modeling of transcriptional dynamics while preserving essential regulatory structure, the expanded transcription factor network was subjected to iterative *kernel reduction*. This procedure systematically removed peripheral nodes with limited connectivity while propagating their regulatory influence to adjacent interactions to maintain overall network logic. Application of kernel reduction criteria resulted in a minimal regulatory module comprising 11 transcription factor genes connected by 117 directed interactions (**Figure 4D**). Topology analysis revealed a densely interconnected regulatory architecture characterized by multiple feedback loops and hub-centered interaction motifs. In addition to direct vitamin D targets such as *ETS2*, *FOS*, and *HIF1A*, the network included several major signaling regulators, including *EGR1* (early growth response 1), *ETS1*, *JUN*, *MYC* (*MYC* proto-oncogene, bHLH transcription factor), *NFKB1* (nuclear factor kappa B subunit 1), *STAT1* (signal transducer and activator of transcription 1), *STAT3*, and *TP53* (tumor protein P53) (**Figure S8**). The presence of these hubs suggested that endocrine transcriptional responses are mediated through hierarchical coordination of central regulatory nodes rather than independent gene-level effects.

The regulatory kernel exhibited a high degree of internal connectivity and included multiple reciprocal regulatory relationships, consistent with models of transcription factor network control characterized by feedback-dominated dynamics. This compact architecture provided a tractable yet biologically interpretable framework for quantitative modeling of transcriptional system behavior in response to repeated endocrine perturbation.

Mechanistic modeling demonstrates dynamic sufficiency of the reduced regulatory kernel

To quantitatively characterize transcription factor dynamics in response to repeated endocrine perturbation, the reduced regulatory network was formalized as a system of ODEs incorporating Hill-type regulatory functions to describe both activating and inhibitory transcriptional interactions (**Figure S2**). In this framework, transcription factor expression levels were modeled as continuous dynamic variables influenced by upstream regulatory inputs, basal transcriptional activity, and degradation kinetics. Circulating serum 25(OH)D₃ concentrations were incorporated as an external forcing variable to represent endocrine input into the regulatory system. Time-dependent vitamin D dynamics were parameterized using pharmacokinetic estimates derived from the exponential production-decay model described above, thereby linking systemic exposure to intracellular regulatory responses.

Model calibration involved simultaneous optimization of interaction strengths, Hill coefficients, and decay parameters across all supplementation cycles. This resulted in a total of 373 optimizable parameters per cycle (**Table S4**), reflecting the complexity of the regulatory system while maintaining the high-dimensional regulatory system dynamically identifiable through *kernel reduction*. The fitted model achieved high-fidelity reconstruction of observed transcription factor expression trajectories across all study time points (**Figure 5**). Simulated expression profiles closely matched experimentally measured RNA-seq data for each regulatory node, capturing both the magnitude and temporal structure of transcriptional responses. In particular, the model reproduced progressive amplification of regulatory activity observed during cumulative supplementation as well as transient fluctuations following individual bolus administrations.

In summary, the reduced transcription factor kernel is dynamically sufficient to reproduce longitudinal endocrine transcriptional responses, indicating that the

essential regulatory logic governing system behavior is captured within this compact network architecture.

Sensitivity analysis reveals hierarchical regulatory control structure

To assess the effective dimensionality and dynamical controllability of the inferred regulatory system, local sensitivity analysis was performed by quantifying relative changes in model output following perturbation of each optimizable parameter. Sensitivity coefficients were averaged across all transcription factor expression trajectories and supplementation cycles to obtain a global measure of parameter influence. This analysis revealed that only 11 of the 373 parameters exhibited moderate sensitivity, whereas the vast majority exerted negligible effects on system behavior (**Figure S11A**). These results indicate substantial functional redundancy and distributed parameter redundancy characteristic of densely coupled regulatory systems.

The most influential parameters were associated with STAT3-mediated repression of FOS, EGR1, TP53 and MYC (**Figure S11B**), identifying this regulatory axis as a key determinant of transcriptional dynamics. At the gene level, *MYC*, *FOS*, and *TP53* displayed the highest dynamical sensitivity, independently confirming their classification as dominant regulatory hubs based on network topology analysis.

Taken together, the transcriptional system behavior is governed by a limited subset of strongly connected control motifs, resulting in a reduced effective control dimension despite the nominal complexity of the model parameter space.

Hub-centered regulatory interactions distinguish response phenotypes

To investigate the regulatory basis of interindividual variability in transcriptional responsiveness, subgroup-specific parameter sets derived from stable responder models were compared. Although the underlying network topology remained conserved across response phenotypes, systematic differences in regulatory interaction strengths were observed. In particular, TP53- and MYC-centered regulatory

modules exhibited distinct activation and repression patterns between high- and low-responder models. These differences affected both response amplitude and temporal kinetics of transcription factor expression, indicating coordinated reconfiguration of regulatory control within the same network architecture. Quantitative parameter ratio analysis identified TP53, JUN, MYC, HIF1A, and FOS as principal determinants of response stratification (**Figure 6A**). Mapping of these determinants onto the integrated regulatory network highlighted their central positioning within hub-centered interaction motifs (**Figure 6B**).

In summary, the interindividual variability in endocrine transcriptional responsiveness arises from quantitative redistribution of control weights within a conserved regulatory topology.

Generalization of inferred regulatory dynamics across independent perturbation cohorts

To evaluate the transferability of the inferred regulatory architecture, the optimized model was applied to predict transcription factor expression responses in an independent vitamin D intervention cohort (VitDHiD) (18). Using cohort-mean baseline expression levels and serum 25(OH)D₃ concentrations as initial conditions, model simulations were performed to predict gene expression 24 hours after bolus supplementation. Predicted expression profiles showed strong agreement with experimentally measured data across regulatory nodes (**Figure S12**). All independently derived parameter sets achieved high predictive accuracy ($R^2 > 0.94$), demonstrating that the regulatory logic inferred from the VitDPAS cohort generalizes beyond the training dataset and supplementation regimen. Prediction errors were largest for EGR1, consistent with its high interindividual variability and position as a dynamically sensitive node within the regulatory network.

Taken together, these results establish that longitudinal physiological perturbation enables quantitative inference of dynamically identifiable transcriptional regulatory

systems in humans and demonstrate that endocrine transcriptional responses are governed by hierarchical hub-centered control motifs.

Discussion

This study establishes an intervention-driven systems biology framework for quantitatively characterizing endocrine transcriptional regulation in humans through the integration of longitudinal transcriptomics, pharmacokinetic modeling, and mechanistic regulatory network inference. By applying repeated weight-adjusted vitamin D₃ bolus interventions, we demonstrate that endocrine signaling induces temporally integrated transcriptional reprogramming in immune cells, accompanied by pronounced interindividual variability in molecular responsiveness. Importantly, the combination of network reduction and dynamic system identification enabled the derivation of a compact transcription factor regulatory kernel that is both structurally minimal yet dynamically sufficient to explain longitudinal gene expression trajectories *in vivo*.

A central finding of this work is the progressive nature of transcriptome remodeling during sustained endocrine exposure. While individual vitamin D₃ boluses elicited relatively modest short-term transcriptional perturbations, cumulative supplementation resulted in a marked increase in both the number and magnitude of regulated genes. This temporal amplification supports a model in which endocrine signaling does not primarily act through acute transcriptional activation but instead drives gradual reconfiguration of regulatory states. Such dynamics are consistent with the concept of progressive stabilization of regulatory network states, potentially mediated by chromatin-dependent transcriptional memory (40) and stepwise stabilization of regulatory network activity. However, the causal epigenetic mechanisms underlying this integration remain to be directly demonstrated. The present findings therefore extend previous short-term intervention studies (18,25) by providing systems-level evidence that endocrine micronutrient signaling can induce progressive transcriptional adaptation under physiological conditions.

The observed heterogeneity in both systemic vitamin D kinetics and transcriptional responses reinforces the concept of the personal vitamin D response index (14). By integrating transcriptional sensitivity with endocrine exposure efficiency, the present classification framework captured reproducible yet dynamic response phenotypes across repeated perturbation cycles. Importantly, response stratification was only partially explained by baseline vitamin D status and showed no significant association with demographic parameters such as age, sex, or BMI. These findings suggest that interindividual variability in vitamin D responsiveness reflects intrinsic properties of transcriptional regulatory systems rather than simple differences in hormone availability (41,42). This systems-level interpretation provides a mechanistic perspective on the heterogeneous outcomes observed in large-scale vitamin D₃ supplementation trials (43), which traditionally assume uniform biological responsiveness (44).

A major conceptual advance of the present study lies in the identification of a dynamically identifiable transcription factor control architecture underlying vitamin D-dependent gene regulation. Although longitudinal transcriptome analysis initially highlighted primary endocrine-responsive transcriptional regulators such as ETS2, FOS, and HIF1A, integrative network analysis and dynamic modeling revealed that endocrine responses are organized within a broader hub-centered regulatory structure dominated by TP53 and MYC and functionally connected to STAT3, NFKB1, and AP1 components. This hierarchical organization suggests that vitamin D signaling converges on conserved regulatory modules that integrate stress sensing, immune activation, and metabolic adaptation (45). The identification of TP53-MYC antagonistic control motifs is particularly notable, as this axis represents a fundamental organizing principle of cellular decision-making under stress conditions.

Dynamic modeling further demonstrated that the reduced regulatory kernel is not only structurally parsimonious but also mechanistically sufficient to reproduce temporal transcription factor dynamics across supplementation cycles and responder

phenotypes. Sensitivity analysis revealed that the effective dimensionality of the regulatory system is substantially lower than the nominal parameter space, indicating that transcriptional system behavior is governed by a limited subset of strongly connected control motifs. Out-of-sample prediction in an independent intervention cohort confirmed that the inferred regulatory logic generalizes beyond the training dataset, supporting the concept that endocrine transcriptional responses are constrained by conserved network architectures rather than cohort-specific regulatory configurations.

Subgroup-specific parameterization revealed that interindividual differences in vitamin D responsiveness arise primarily from quantitative redistribution of regulatory influence within a conserved network topology. In particular, differential weighting of TP53- and MYC-centered regulatory motifs, modulated by AP1 components and HIF1A, emerged as a principal determinant of response phenotype divergence (46,47). This observation supports a regulatory model in which endocrine signals shape transcriptional outputs through state-dependent reconfiguration of pre-existing control circuits rather than through activation of distinct pathway modules. Such a mechanism provides a systems-level explanation for the pleiotropic yet context-dependent physiological effects attributed to vitamin D signaling (48).

Beyond the specific biological insights into vitamin D action, the present work establishes a generalizable methodological paradigm for mechanistic inference of human regulatory network dynamics under physiological perturbation conditions. Intervention-driven system identification represents a conceptual shift from descriptive transcriptomics toward predictive regulatory modeling, aligning with emerging systems medicine approaches that aim to define individualized regulatory states. The convergence of vitamin D signaling with core stress-adaptive transcriptional networks further suggests that endocrine micronutrients may modulate fundamental control architectures governing immune resilience and metabolic homeostasis.

Several limitations of the present study should be acknowledged. The cohort size, while sufficient for dynamic modeling, restricts statistical power for detecting subgroup-specific gene regulation and precludes direct linkage of molecular response phenotypes to long-term clinical outcomes. Moreover, bulk PBMC transcriptomics does not resolve cell-type-specific regulatory responses, which may contribute to the observed variability. The modeling framework assumes deterministic regulatory interactions and does not explicitly incorporate stochastic gene expression dynamics, post-transcriptional regulation, or epigenetic state transitions. Additionally, reliance on curated regulatory interaction databases may limit the detection of context-specific transcriptional interactions. Future studies integrating single-cell multi-omics, chromatin accessibility profiling, functional immune phenotyping, and longitudinal clinical follow-up will be essential to validate and extend the proposed regulatory model.

In conclusion, this study defines a dynamic systems framework linking endocrine vitamin D exposure to transcription factor network activity and downstream gene expression responses in humans. By integrating longitudinal intervention data with mechanistic modeling, we demonstrate that interindividual nutrient responsiveness can be quantitatively described in terms of regulatory system dynamics rather than static biomarker levels. More broadly, the findings highlight controlled nutritional perturbation as a powerful experimental strategy for identifying human regulatory network behavior under physiological conditions (49). Intervention-driven dynamic system identification therefore represents a promising approach for mechanistic prediction of individualized nutrient responses and for advancing predictive regulatory systems biology of human endocrine signaling (50-53).

Funding

This publication is part of a project that has received funding from the European Union's Horizon2020 research and innovation program under grant agreement no. 952601, from the David and Amy Fulton Foundation, Seattle, US and from the National

Science Centre (Poland), project number 2023/49/B/NZ9/00402 (to CC). EDH acknowledges support by UNAM-PAPIIT IN216526.

Disclosure

The authors have nothing to disclose.

Acknowledgements

Kind thanks to the Gene Core Facility at the EMBL in Heidelberg, Germany, for massively parallel sequencing services, the Centre of Informatics Tricity Academic Supercomputer and Network of the Technical University of Gdansk (<https://task.gda.pl/en>) for the use of their the supercomputers for sequence alignment and Procter & Gamble Health Poland for donating Vigantoletten Max 4000 to our study.

Author contribution

The VitDPAS study was designed by CC and conducted by EG, KTM and JR. Data analysis and visualization were performed by RGD, while RNA isolation and RNA-seq library preparation were carried out by EG. RGD performed mechanistic modeling under the supervision of EDH. CC refined the figures and drafted the manuscript, which was reviewed and approved by all authors.

Data Availability

Fastq files of the raw data can be found at Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) under accession numbers GSE278885 (days 0 and 1) and GSE319989 (days 28, 29, 56, 57 and 84).

References

1. Carlberg, C., Raczyk, M. and Zawrotna, N. (2023) Vitamin D: a master example of nutrigenomics. *Redox Biol*, **62**, 102695.
2. Kitano, H. (2002) Systems biology: a brief overview. *Science*, **295**, 1662-1664.
3. Haussler, M.R., Jurutka, P.W., Mizwicki, M. and Norman, A.W. (2011) Vitamin D receptor (VDR)-mediated actions of 1 α ,25(OH)₂vitamin D₃: genomic and non-genomic mechanisms. *Best practice & research. Clinical endocrinology & metabolism*, **25**, 543-559.
4. Carlberg, C. (2023) Genomic signaling of vitamin D. *Steroids*, **198**, 109271.
5. Pike, J.W., Meyer, M.B., Lee, S.M., Onal, M. and Benkusky, N.A. (2017) The vitamin D receptor: contemporary genomic approaches reveal new basic and translational insights. *J Clin Invest*, **127**, 1146-1154.
6. Christakos, S. (2021) Vitamin D: a critical regulator of intestinal physiology. *JBMR Plus*, **5**, e10554.

7. Martens, P.J., Gysemans, C., Verstuyf, A. and Mathieu, A.C. (2020) Vitamin D's effect on immune function. *Nutrients*, **12**, 1248.
8. Bishop, E., Ismailova, A., Dimeloe, S.K., Hewison, M. and White, J.H. (2021) Vitamin D and immune regulation: antibacterial, antiviral, anti-inflammatory. *JBMR Plus*, **5**, e10405.
9. Bouillon, R., Marocchi, C., Carmeliet, G., Bikle, D., White, J.H., Dawson-Hughes, B., Lips, P., Munns, C.F., Lazaretti-Castro, M., Giustina, A. *et al.* (2019) Skeletal and extra-skeletal actions of vitamin D: Current evidence and outstanding questions. *Endocr Rev*, **40**, 1109-1151.
10. Manson, J.E., Cook, N.R., Lee, I.M., Christen, W., Bassuk, S.S., Mora, S., Gibson, H., Gordon, D., Copeland, T., D'Agostino, D. *et al.* (2019) Vitamin D supplements and prevention of cancer and cardiovascular disease. *N Engl J Med*, **380**, 33-44.
11. Virtanen, J.K., Nurmi, T., Aro, A., Bertone-Johnson, E.R., Hypponen, E., Kroger, H., Lamberg-Allardt, C., Manson, J.E., Mursu, J., Mantyselka, P. *et al.* (2022) Vitamin D supplementation and prevention of cardiovascular disease and cancer in the Finnish Vitamin D Trial: a randomized controlled trial. *Am J Clin Nutr*, **115**, 1300-1310.
12. Pilz, S., Trummer, C., Theiler-Schwetz, V., Grubler, M.R., Verheyen, N.D., Odler, B., Karras, S.N., Zittermann, A. and März, W. (2022) Critical appraisal of large vitamin D randomized controlled Trials. *Nutrients*, **14**, 303.
13. Jaroslawska, J. and Carlberg, C. (2023) In vivo regulation of signal transduction pathways by vitamin D stabilizes homeostasis of human immune cells and counteracts molecular stress. *Int J Mol Sci*, **24**, 14632.
14. Carlberg, C. and Haq, A. (2018) The concept of the personal vitamin D response index. *J Steroid Biochem Mol Biol*, **175**, 12-17.
15. Nourazarain, A. and Vaziri, Y. (2025) Nutrigenomics meets multi-omics: integrating genetic, metabolic, and microbiome data for personalized nutrition strategies. *Genes Nutr*, **20**, 30.
16. Tripathi, T. and Carlberg, C. (2025) Early in vivo target genes in human immune cells highlight vitamin D's role in antioxidant defense. *Frontiers in immunology*, **16**, 1559486.
17. Hanel, A., Neme, A., Malinen, M., Hamalainen, E., Malmberg, H.R., Etheve, S., Tuomainen, T.P., Virtanen, J.K., Bendik, I. and Carlberg, C. (2020) Common and personal target genes of the micronutrient vitamin D in primary immune cells from human peripheral blood. *Scientific reports*, **10**, 21051.
18. Ghosh Dastidar, R., Jaroslawska, J., Malinen, M., Tuomainen, T.P., Virtanen, J.K., Bendik, I. and Carlberg, C. (2024) In vivo vitamin D targets reveal the upregulation of focal adhesion-related genes in primary immune cells of healthy individuals. *Scientific reports*, **14**, 17552.
19. Jaroslawska, J., Ghosh Dastidar, R. and Carlberg, C. (2024) In vivo vitamin D target genes interconnect key signaling pathways of innate immunity. *PLoS ONE*, **19**, e0306426.
20. Moeckel, C., Mouratidis, I., Chantzi, N., Uzun, Y. and Georgakopoulos-Soares, I. (2024) Advances in computational and experimental approaches for deciphering transcriptional regulatory networks: Understanding the roles of cis-regulatory elements is essential, and recent research utilizing MPRA, STARR-seq, CRISPR-Cas9, and machine learning has yielded valuable insights. *Bioessays*, **46**, e2300210.
21. Hossein-nezhad, A., Spira, A. and Holick, M.F. (2013) Influence of vitamin D status and vitamin D₃ supplementation on genome wide expression of white blood cells: a randomized double-blind clinical trial. *PLoS ONE*, **8**, e58725.
22. Carlberg, C., Seuter, S., de Mello, V.D., Schwab, U., Voutilainen, S., Pulkki, K., Nurmi, T., Virtanen, J., Tuomainen, T.P. and Uusitupa, M. (2013) Primary vitamin D target genes allow a categorization of possible benefits of vitamin D₃ supplementation. *PLoS ONE*, **8**, e71042.
23. Gospodarska, E., Ghosh Dastidar, R., Jaroslawska, J., Rybinski, M., Raczyk, M., Tokarczyk-Malesa, K., Romaszko, J. and Carlberg, C. (2025) Transcriptomic profiling of immune modulation induced by vitamin D₃ in the VitDPAS and VitDHiD cohort studies. *Scientific reports*, **15**, 17334.
24. Scragg, R., Stewart, A.W., Waayer, D., Lawes, C.M.M., Toop, L., Sluyter, J., Murphy, J., Khaw, K.T. and Camargo, C.A., Jr. (2017) Effect of monthly high-dose vitamin D supplementation on cardiovascular disease in the vitamin D assessment study: a randomized clinical trial. *JAMA cardiology*, **2**, 608-616.
25. Vukic, M., Neme, A., Seuter, S., Saksa, N., de Mello, V.D., Nurmi, T., Uusitupa, M., Tuomainen, T.P., Virtanen, J.K. and Carlberg, C. (2015) Relevance of vitamin D receptor target genes for monitoring the vitamin D responsiveness of primary human cells. *PLoS ONE*, **10**, e0124339.
26. Seuter, S., Virtanen, J.K., Nurmi, T., Pihlajamäki, J., Mursu, J., Voutilainen, S., Tuomainen, T.P., Neme, A. and Carlberg, C. (2017) Molecular evaluation of vitamin D responsiveness of healthy young adults. *J Steroid Biochem Mol Biol*, **174**, 314-321.

27. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.
28. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923-930.
29. Durinck, S., Spellman, P.T., Birney, E. and Huber, W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols*, **4**, 1184-1191.
30. Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139-140.
31. Chen, Y., Lun, A.T. and Smyth, G.K. (2016) From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research*, **5**, 1438.
32. Rybinski, M., Ghosh Dastidar, R., Zawrotna, N., Gospodarska, E. and Carlberg, C. (2025) Chromatin and transcriptional dynamics underlying the immune-modulatory effects of vitamin D₃ in vivo. *Scientific reports*, **16**, 2997.
33. Karlebach, G. and Shamir, R. (2008) Modelling and analysis of gene regulatory networks. *Nat Rev Mol Cell Biol*, **9**, 770-780.
34. Garcia-Alonso, L., Holland, C.H., Ibrahim, M.M., Turei, D. and Saez-Rodriguez, J. (2019) Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome research*, **29**, 1363-1375.
35. Kim, J.R., Kim, J., Kwon, Y.K., Lee, H.Y., Heslop-Harrison, P. and Cho, K.H. (2011) Reduction of complex signaling networks to a representative kernel. *Science signaling*, **4**, ra35.
36. Nurminen, V., Seuter, S. and Carlberg, C. (2019) Primary vitamin D target genes of human monocytes. *Front Physiol*, **10**, 194.
37. Tuoresmäki, P., Väisänen, S., Neme, A., Heikkinen, S. and Carlberg, C. (2014) Patterns of genome-wide VDR locations. *PLoS ONE*, **9**, e96105.
38. Neme, A., Seuter, S. and Carlberg, C. (2017) Selective regulation of biological processes by vitamin D based on the spatio-temporal cistrome of its receptor. *Biochim Biophys Acta*, **1860**, 952-961.
39. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R. and Weirauch, M.T. (2018) The human transcription factors. *Cell*, **172**, 650-665.
40. Hanel, A., Malmberg, H.R. and Carlberg, C. (2020) Genome-wide effects of chromatin on vitamin D signaling. *J Mol Endocrinol*, **64**, R45-R56.
41. Mavar, M., Soric, T., Bagaric, E., Saric, A. and Matek Saric, M. (2024) The power of vitamin D: is the future in precision nutrition through personalized supplementation plans? *Nutrients*, **16**, 1176.
42. Qahtan, F., Abu-Qiyas, S. and Papandreou, D. (2026) Genetic and epigenetic determinants of vitamin D metabolism: nutrigenomic insights for precision nutrition. *Frontiers in Nutrition*, **13**, 1772849.
43. Hands, J.M., Corr, P.G. and Frame, L.A. (2023) Clarifying the heterogeneity in response to Vitamin D in the development, prevention, and treatment of type 2 diabetes mellitus: a narrative review. *International journal of environmental research and public health*, **20**, 6187.
44. Fernandes, T.H. and Bell, V. (2024) The imprecision of micronutrient requirement values: the example of vitamin D. *J Food Sci*, **89**, 51-63.
45. Evan, G.I. and Vousden, K.H. (2001) Proliferation, cell cycle and apoptosis in cancer. *Nature*, **411**, 342-348.
46. Shaulian, E. and Karin, M. (2002) AP-1 as a regulator of cell life and death. *Nat Cell Biol*, **4**, E131-136.
47. Palazon, A., Goldrath, A.W., Nizet, V. and Johnson, R.S. (2014) HIF transcription factors, inflammation, and immunity. *Immunity*, **41**, 518-528.
48. Deritei, D., Aird, W.C., Ercsey-Ravasz, M. and Regan, E.R. (2016) Principles of dynamical modularity in biological regulatory networks. *Scientific reports*, **6**, 21957.
49. De Domenico, M., Allegri, L., Caldarelli, G., d'Andrea, V., Di Camillo, B., Rocha, L.M., Rozum, J., Sbarbati, R. and Zambelli, F. (2025) Challenges and opportunities for digital twins in precision medicine from a complex systems perspective. *NPJ Digit Med*, **8**, 37.
50. Carlberg, C. and Campbell, M.J. (2013) Vitamin D receptor signaling mechanisms: Integrated actions of a well-defined transcription factor. *Steroids*, **78**, 127-136.
51. Ordovas, J.M., Ferguson, L.R., Tai, E.S. and Mathers, J.C. (2018) Personalised nutrition and health. *Bmj*, **361**, bmj k2173.
52. Artusa, P. and White, J.H. (2025) Vitamin D and its analogs in immune system regulation. *Pharmacol Rev*, **77**, 100032.

53. Bjornsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D.R., Gustafsson, M., Jornsten, R., Lee, E.J., Li, X., Lilja, S. *et al.* (2019) Digital twins to personalize medicine. *Genome medicine*, **12**, 4.

Figure legends

Figure 1. Study design and serum 25(OH)D₃ response to repeated vitamin D₃ supplementation. Overview of the longitudinal study design (A). A cohort of 41 participants received vitamin D₃ supplementation in three cycles: cycle 1 (days 0-28), cycle 2 (days 28-56), and cycle 3 (days 56-84). Blood samples were collected at days 0, 1, 28, 29, 56, 57, and 84 for measurement of serum 25(OH)D₃ levels and transcriptome profiling by RNA-seq. Serum 25(OH)D₃ concentrations over the study period (B). Boxplots show the distribution of 25(OH)D₃ levels (ng/mL) at each sampling time point, with individual measurements indicated by dots and median values connected by a line. Shaded areas denote post-supplementation measurements (days 0-1, 28-29, and 56-57). Vitamin D₃ supplementation significantly increased 25(OH)D₃ levels after each cycle ($p < 0.0001$).

Figure 2. Temporal differential gene expression analysis during the VitDPAS study. Volcano plots showing differential gene expression between consecutive sampling time points and cumulative comparisons across the study period. Each point represents a gene. The x-axis shows the log₂FC in expression between the indicated time points, and the y-axis shows the log₁₀p-value. Vertical dashed lines indicate log₂FC thresholds, and the horizontal dashed line indicates the statistical significance threshold. Colors denote significance categories: light grey, not significant; black, nominally significant ($p < 0.05$); green, nominally significant with small FC; orange, significant after multiple-testing correction ($FDR < 0.05$); red, significant after FDR correction with substantial FC. Selected genes having a central role in later parts of the study are highlighted.

Figure 3. Classification of VitDPAS study participants according to transcriptional responsiveness to vitamin D₃. The 41 participants of the VitDPAS study were classified into high (green), mid (yellow), and low (red) responders based on the expression profiles of 23 *in vivo* vitamin D target genes that were consistently regulated across the three independent cycles of the vitamin D₃ bolus supplementation

experiment (**Figure S5**). Each numbered point represents an individual participant. Participant classification was derived from *k-means* clustering. The epicenter of the classification, reflecting the centroid of the response distribution across supplementation cycles, is indicated in the plot.

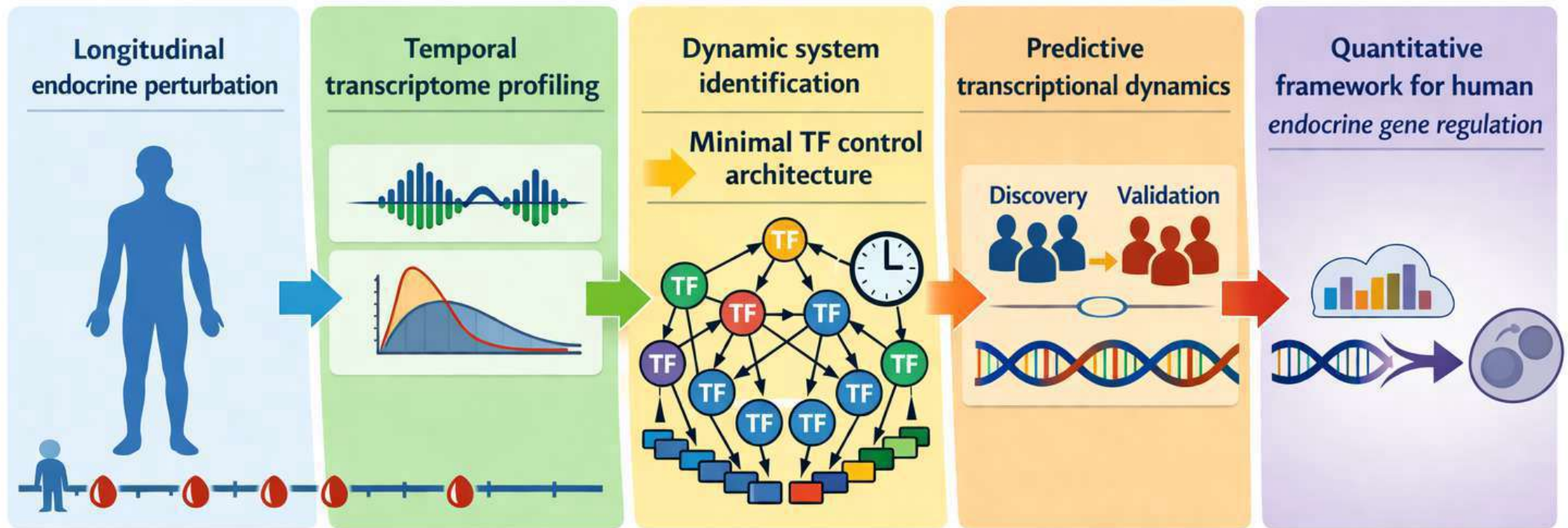
Figure 4. Construction of the vitamin D-responsive gene regulatory network. The initial network (Stage 1) was generated from seven transcription factor genes identified among persistently regulated vitamin D target genes (**A**). An expanded network (Stage 2) was subsequently constructed by integrating curated transcriptional regulatory interactions from the DoRothEA resource (**B**), thereby increasing network complexity through the inclusion of interacting transcription factors and downstream target genes connected to the initial input nodes. Further refinement involved sequential filtering steps: removal of the upstream genes *CSRNP1* and *KLF10* (Stage 3, **C**), and application of *kernel reduction* to identify the core set of interacting genes (Stage 4, **D**). Asterisks denote genes classified as vitamin D target genes, and node shapes distinguish transcription factor genes from non-transcription factor genes. Green edges indicate activating interactions, whereas red edges indicate repressive interactions. Network statistics (number of nodes, edges, transcription factors, and vitamin D target genes) are provided within each panel.

Figure 5. Model-based reconstruction of temporal transcription factor expression dynamics during repeated vitamin D₃ supplementation. Time-resolved expression profiles of 11 key transcription factor genes predicted by the regulatory network model. Panels display RNA-seq expression levels across all seven study time points. Red squares represent experimentally measured mean expression values ($\pm$ SD) across the cohort, whereas black lines with circles denote model-predicted trajectories obtained from the Hill-function-based ODE framework derived from the regulatory network shown in **Figure 4**. Shaded background regions indicate the three vitamin D₃ supplementation cycles, each comprising a baseline and a 24-hour post-supplementation measurement (d0-d1, d28-d29, and d56-d57). The close

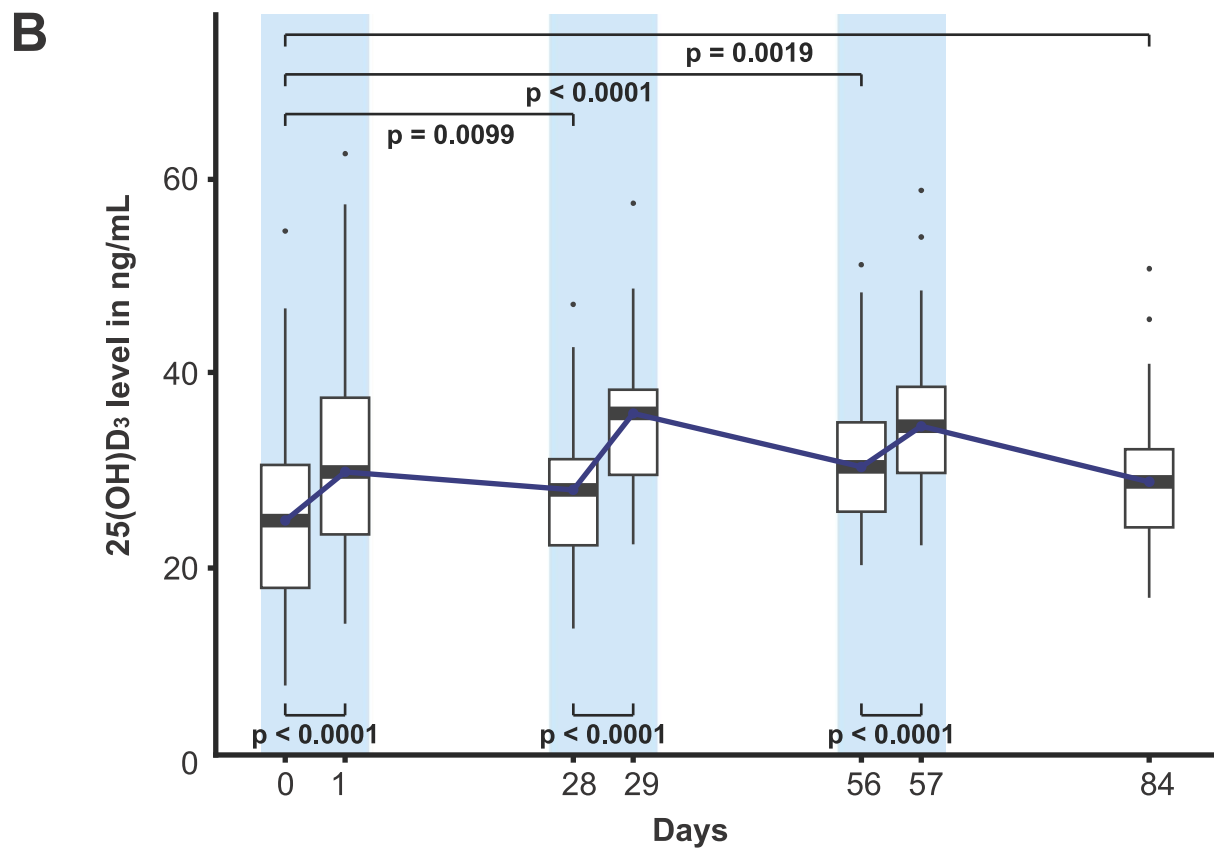
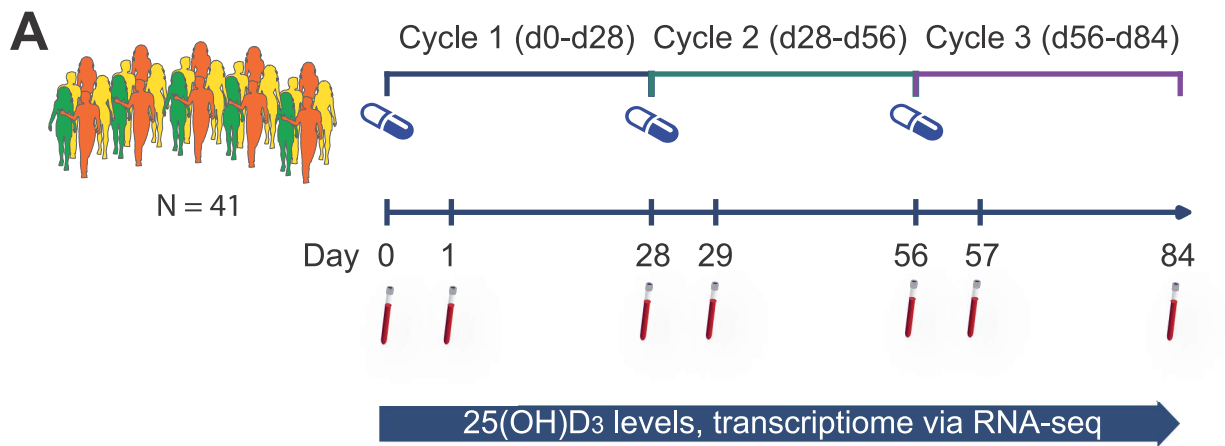
agreement between predicted and observed trajectories demonstrates that the inferred regulatory system quantitatively captures the temporal transcriptional responses induced by vitamin D₃ intervention.

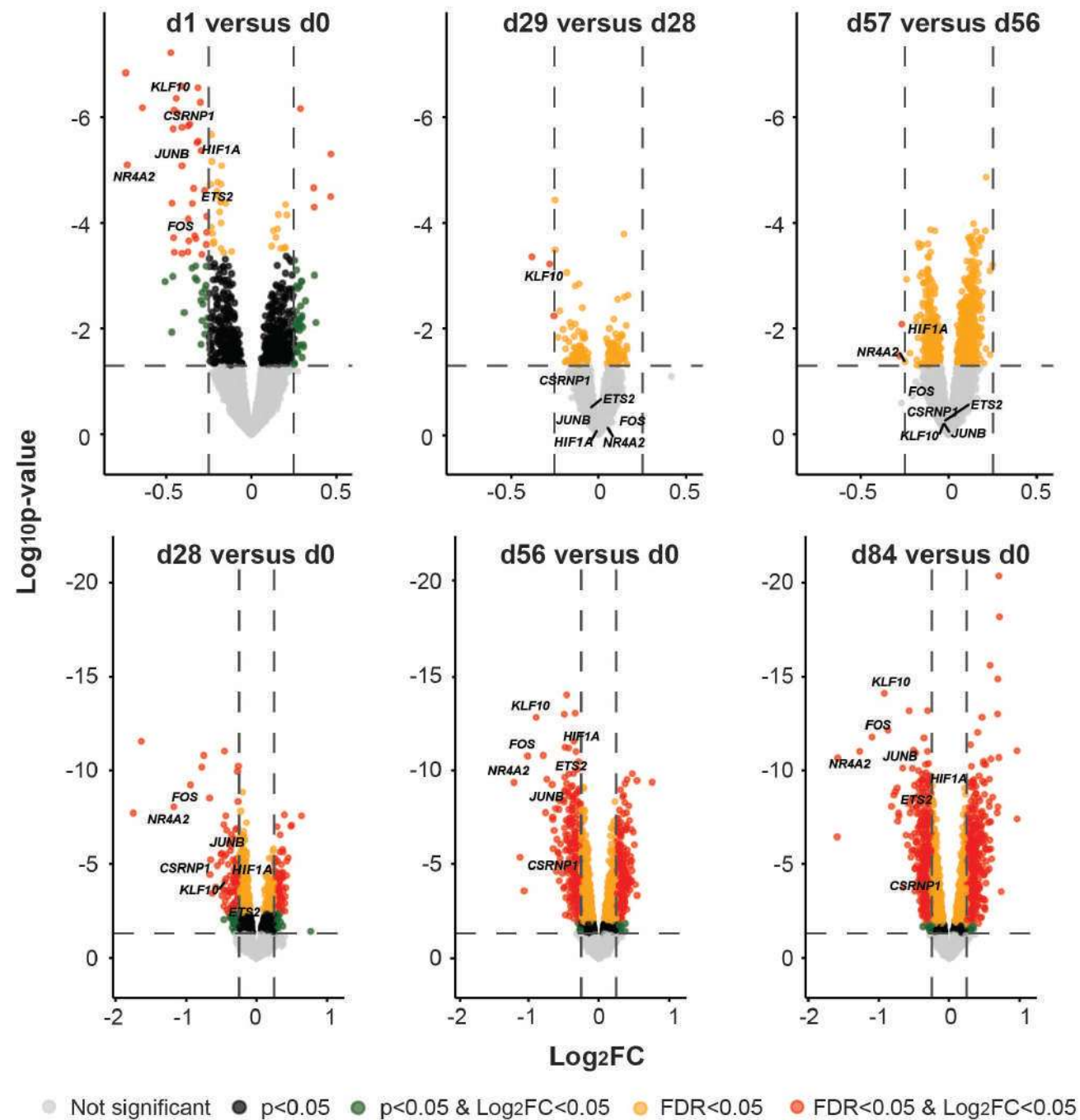
Figure 6. Regulatory interactions distinguishing high and low vitamin D responders and integrative network representation. Ranked regulatory interactions contributing to differential transcriptional responses between high and low vitamin D responders (**A**). Horizontal bars indicate the normalized difference in interaction strength derived from optimized model parameters, with positive values (green) denoting interactions stronger in high responders and negative values (red) indicating interactions stronger in low responders. Regulatory edges are annotated using activation ($\rightarrow$) or repression ($\neg$) symbols. The analysis identifies TP53-, MYC-, STAT3-, AP1 (JUN/FOS)-, and HIF1A-centered interactions as principal determinants of divergence in response phenotypes. Integrated transcription factor regulatory network summarizing differential control architecture linking circulating 25(OH)D₃ levels to downstream transcriptional responses (**B**). Nodes represent transcription factors and vitamin D status, while directed edges denote activating (green) or inhibitory (red) regulatory relationships. Solid and dashed edges reflect differential interaction strengths between responder classes, illustrating hub-centered control motifs underlying interindividual variability in vitamin D responsiveness.

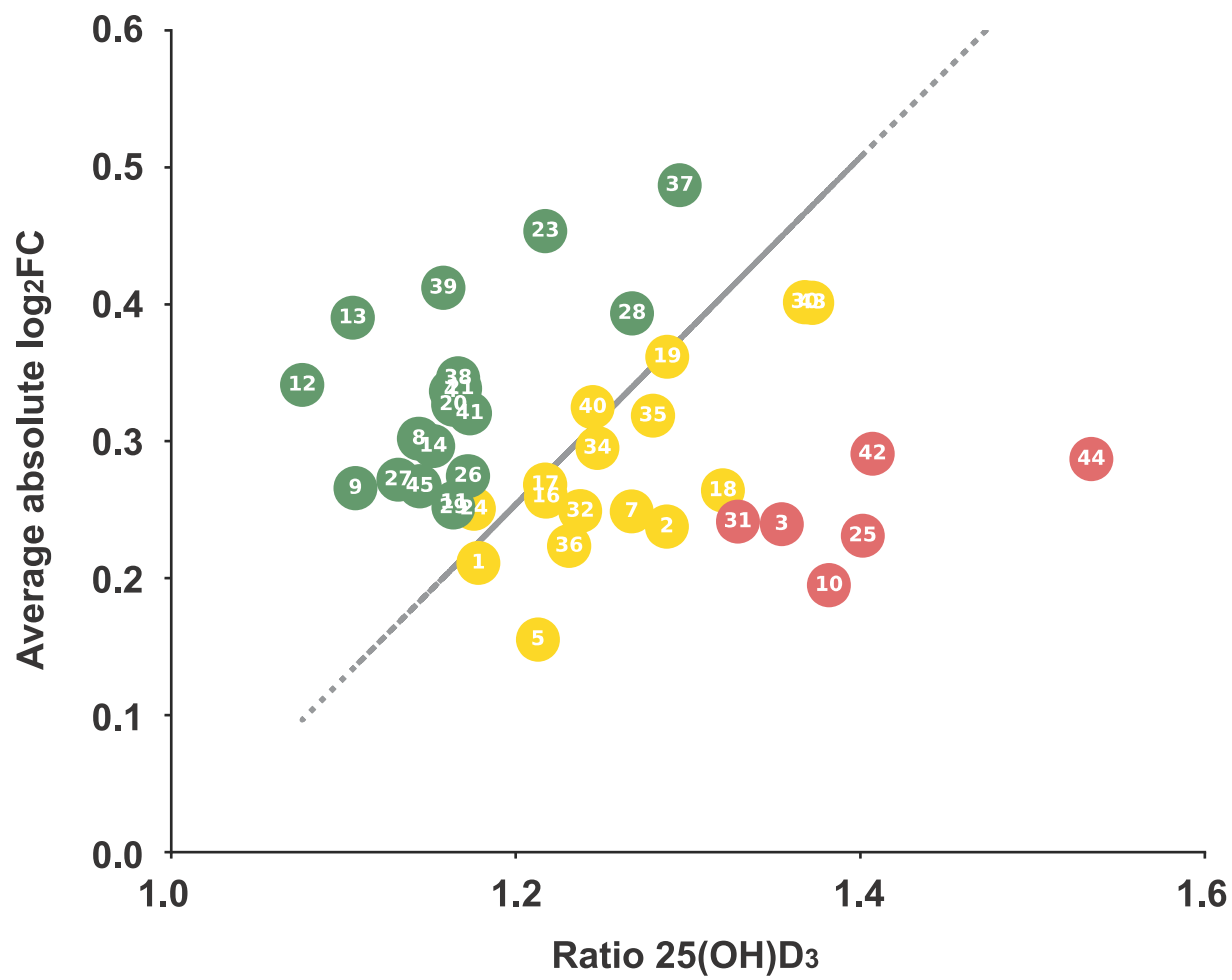
Dynamic system identification of endocrine gene regulation in humans

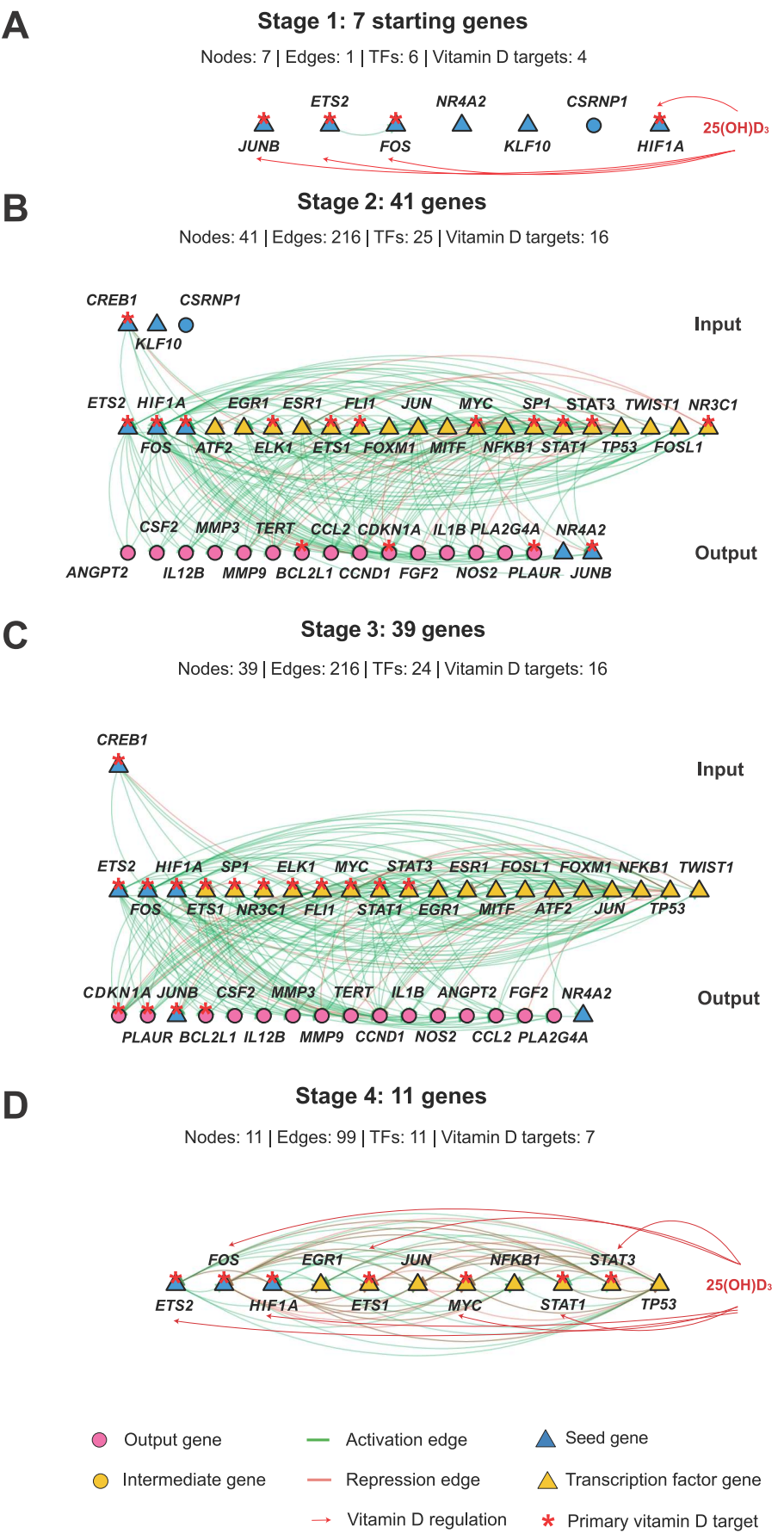


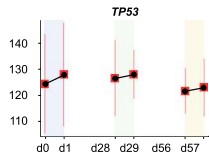
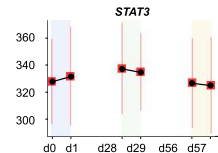
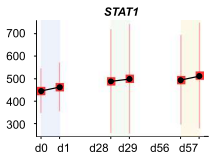
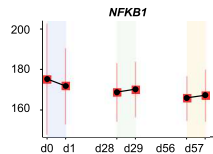
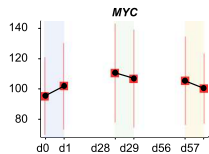
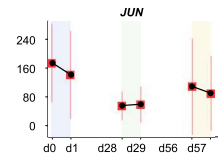
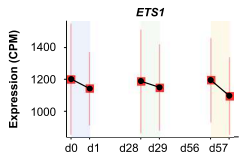
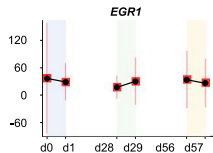
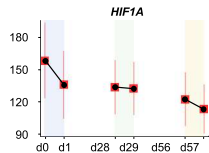
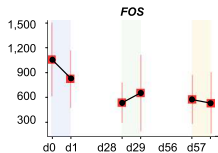
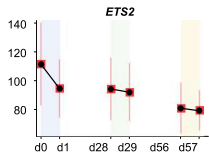
Longitudinal in vivo perturbation → quantitative regulatory inference





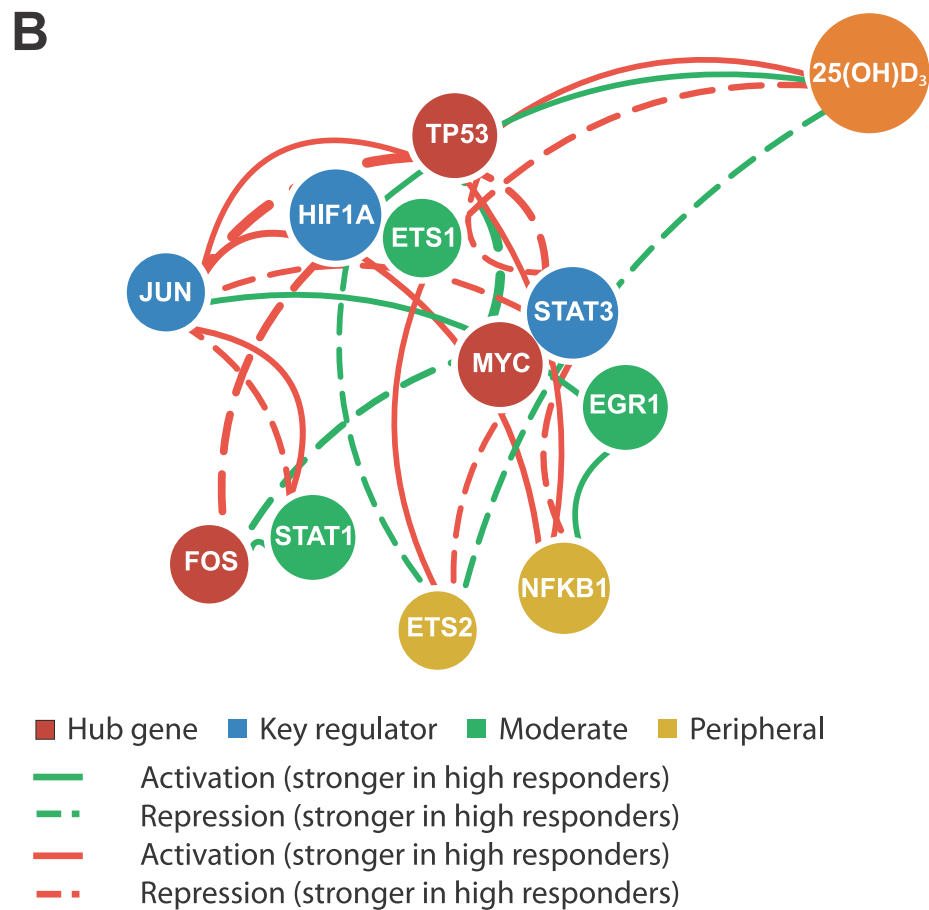
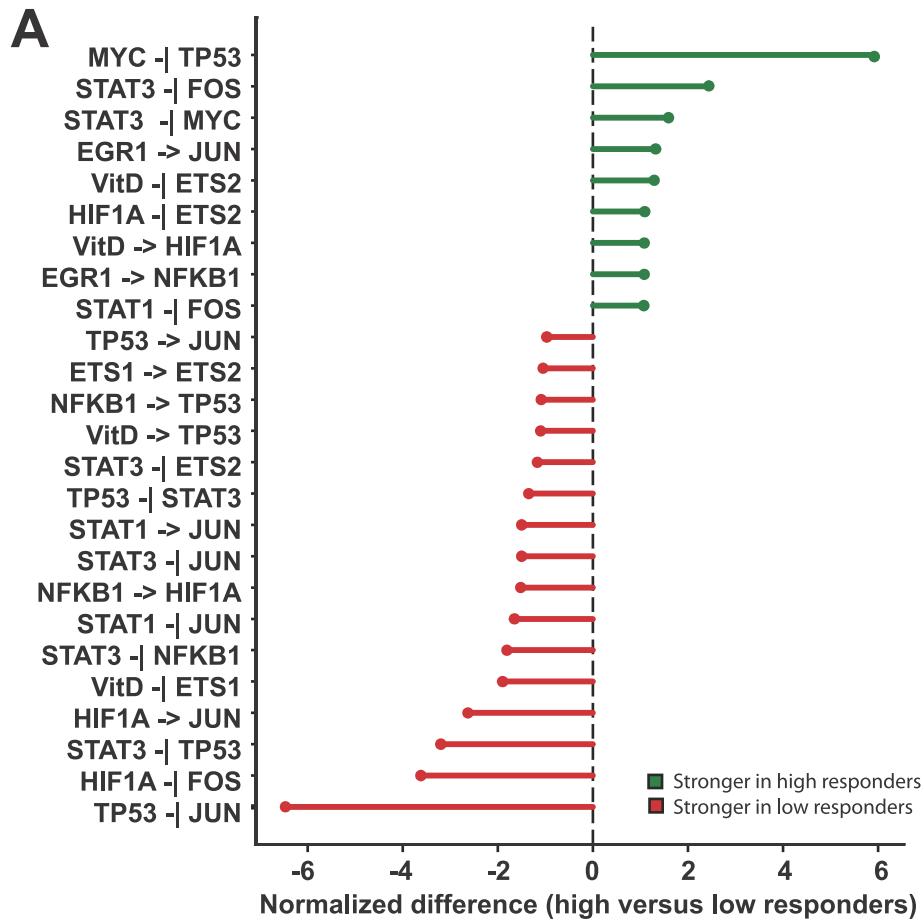






■ Experimental (mean±SD)

◆ Predicted



Supplementary Material

Supplementary Tables

Table S1. Participant characteristics and individual serum 25(OH)D₃ responses in the VitDPAS study. Age, sex, BMI, and serum 25(OH)D₃ concentrations of the 41 participants measured before (d0, d28, d56) and 24 hours after each vitamin D₃ bolus (d1, d29, d57), as well as at the final follow-up (d84). The table also shows the overall change in vitamin D status during the 12-week intervention (d84-d0), the average increase after each weight-adjusted vitamin D₃ bolus (1,000 IU/kg), and the classification of participants according to their vitamin D response index (**Figure 3**). Rows are colored according to response index classification in low (red), mid (yellow) and high (green). Participants selected for modeling of high, mid, and low responders are highlighted in bold.

Table S2. Sequencing depth of RNA-seq samples from the VitDPAS study. Summary of RNA sequencing coverage for all samples analyzed in the VitDPAS intervention study. The table lists the sample name, the sampling day, the participant identifier, and the number of assigned sequencing reads (in millions) obtained after read mapping and quality filtering.

Table S3. Differentially expressed genes in response to vitamin D₃ supplementation. List of genes significantly differentially expressed between the indicated sampling time points of the VitDPAS study. The table reports gene identifiers together with statistical metrics derived from the differential expression analysis, including log₂FC, p-value, and FDR. Comparisons correspond to the different supplementation cycles (d1 versus d0, d29 versus d28, d57 versus d56) and longer-term comparisons across the intervention period (d28 versus d0, d56 versus d0, d84 versus d0). Positive log₂FC values indicate upregulated genes, whereas negative values indicate downregulated genes following vitamin D₃ supplementation. All 19,295 protein-coding genes are listed, but statistical analysis focused on 8,378 genes with

average CPM > 10 and are expressed in all individuals (allowing all six time point comparisons; highlighted in green).

Table S4. Network-based prioritization and parametric characterization of vitamin D-responsive regulatory interactions. This multi-sheet table summarizes the quantitative analysis of the vitamin D-responsive gene regulatory network. **Sheet 1:** Ranking of network genes based on composite importance scores integrating normalized connectivity and normalized interaction strength. Additional annotations indicate whether genes are classified as vitamin D-regulated targets. **Sheet 2:** Identification of regulatory interactions that most strongly distinguish high from low vitamin D responders. Interactions are ranked based on normalized differential contributions across response groups, with annotation of vitamin D-related regulatory links. **Sheets 3-5:** Parameter values for all modeled regulatory interactions across three independent model runs (R1-R3). Values are provided for combined data as well as for high, mid, and low responder submodels. Interactions directly involving vitamin D signaling are indicated. **Sheet 6:** Methodological description of the parametric analysis workflow, including definitions of network metrics, normalization procedures, interaction filtering criteria, and model interpretation guidelines.

Supplementary Figures

Figure S1. Principle of *kernel reduction* for refinement of the vitamin D-responsive gene regulatory network. Schematic illustration of the *kernel reduction* procedure applied to simplify the transcriptional regulatory network. Starting from the expanded interaction network, genes and regulatory connections are iteratively filtered based on predefined criteria, such as interaction support and connectivity, to identify a core regulatory module (kernel). Nodes represent genes, and edges indicate predicted regulatory interactions. The reduction process retains genes that form strongly interconnected regulatory relationships while removing peripheral or weakly supported interactions.

Figure S2. Mathematical model of the core regulatory network. ODEs used for modeling of the 11 key transcription factor genes and 25(OH)D₃ levels. Values of the respective parameters can be found in **Table S4**.

Figure S3. Pharmacokinetic modeling of serum 25(OH)D₃ responses to repeated vitamin D₃ bolus supplementation. Comparison of observed changes in serum 25(OH)D₃ concentrations following weight-adjusted vitamin D₃ bolus administration in the VitDPAS cohort (n = 41) and in a previously reported single-participant study using the same triple-bolus regimen (32) (**A**). Individual measurements at baseline and 24 hours after supplementation are shown. Pharmacokinetic model-based estimation of 25(OH)D₃ production derived from the single-participant dataset, including the additional 48-hour post-bolus sampling point (**B**). Model fitting enabled calculation of absorption rate constants and predicted steady-state serum 25(OH)D₃ concentrations, providing quantitative parameters for describing interindividual variability in vitamin D status responses.

Figure S4. Principal component analysis of individual transcriptional responses across vitamin D₃ supplementation cycles. MDS plots derived from genome-wide log₂FC values illustrate similarities and differences in transcriptional responses among study participants. Each point represents an individual participant (Ind01-Ind45),

positioned according to the first two MDS dimensions, which capture the major sources of variation in gene expression responses. Panels correspond to the six pairwise comparisons between study time points. Points are colored according to the respective comparison, and ellipses indicate the distribution of individuals within each condition, highlighting the overall structure of transcriptional variation across the cohort. Symbols denote participant sex.

Figure S5. Overlap of vitamin D-responsive genes across supplementation cycles. Venn diagram illustrating the overlap of genes significantly regulated ($p < 0.05$) in the three 24-hour transcriptional response comparisons following each vitamin D₃ bolus supplementation. The 23 genes consistently responsive across all three cycles are listed. Genes that additionally met the multiple-testing-adjusted significance threshold ($FDR < 0.05$) after the first bolus are highlighted in bold, indicating the most robust primary vitamin D targets.

Figure S6. Classification of VitDPAS study participants based on repeated vitamin D₃ bolus experiments. The 41 participants of the VitDPAS study were classified as high (green), mid (yellow), or low (red) responders according to the expression profiles of 23 *in vivo* vitamin D target genes that were consistently regulated across three independent cycles of the vitamin D₃ bolus supplementation experiment (**Figure S5**). Each individual is represented by three nodes corresponding to supplementation cycles 1, 2, and 3, which are connected by two edges. The average response patterns of high, mid, and low responder groups are shown as solid nodes and edges in the respective colors. The overall average behavior of all participants is summarized by grey solid nodes and edges.

Figure S7. Overlap of vitamin D₃-regulated genes across supplementation time points. Venn diagrams showing the overlap of significantly regulated vitamin D target genes ($FDR < 0.05$) identified at four time points following vitamin D₃ supplementation (d1, d28, d56, and d84 versus baseline d0). Panel **A** includes all regulated target genes, whereas panel **B** is restricted to the subset encoding transcription factors. The

total number of regulated genes at each time point is indicated in parentheses. The intersection across all four comparisons represents a core set of genes that are persistently regulated during both short- and long-term supplementation.

Figure S8. Functional characterization of the vitamin D-responsive regulatory gene module. Interaction matrix of the 40 genes included in the gene regulatory network, indicating the predicted regulatory effects of each gene on the others (**A**). Green squares denote activating interactions, red squares indicate repressive interactions, and white cells represent the absence of a regulatory effect. Interaction matrix of the 11 remaining genes after *kernel reduction* (**B**). The number of activating (green) and repressing (red) interactions are given.

Figure S9. ODE-based regulatory network model accurately reproduces transcription factor expression dynamics for all VitDPAS participants. Boxplots show experimentally measured gene expression levels of all 41 VitDPAS study participants at baseline (0 hours) and 24 hours after vitamin D₃ bolus administration across three experimental cycles. Colored boxes represent cycle-specific measurements before (light shades) and after (dark shades) supplementation, while diamonds indicate simulated expression values generated by the ODE-based Hill-function model.

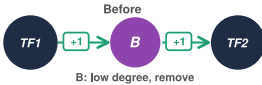
Figure S10. Subgroup-specific ODE models reproduce transcription factor expression dynamics in high, mid, and low vitamin D responders. Boxplots show experimentally measured expression levels of the 11 transcription factor genes at baseline (0 hours) and 24 hours after vitamin D₃ bolus administration across three experimental cycles. Panels represent responder subgroups: high responders (**A**), mid responders (**B**), and low responders (**C**). Colored boxes indicate supplementation cycle-specific measurements before (lighter shades) and after (darker shades) supplementation, while diamonds denote simulated expression values generated by subgroup-specific ODE models based on optimized parameters.

Figure S11. Parameter sensitivity analysis. Ranked sensitivity scores for all 373 parameters; dashed line indicates moderate threshold ($S = 0.01$). The blue vertical line marks the top 50 parameters shown in panel B (**A**). Heatmap of top 50 parameters across 11 genes, annotated with regulatory interactions. STAT3-, MYC-, and TP53-centered repressive interactions dominate (**B**).

Figure S12. Out-of-sample validation against the independent VitDHiD cohort. Boxplots show the distribution of expression levels across 25 VitDHiD participants at baseline (0 h, light boxes) and 24 hours after vitamin D₃ bolus administration (24 h, dark boxes) for each of the 11 transcription factor genes. Cohort-mean baseline expression served as the initial condition for the ODE model (black circles). Colored diamonds connected by lines indicate predicted 24-hour expression using R1 (red), R2 (blue), and R3 (green) parameter sets from the VitDPAS-trained model with VitDPAS-derived 48-hour pharmacokinetics. The model quantitatively predicts post-bolus expression levels across an independent cohort ($R^2 = 0.978$ with R2 parameters).

Remove node if: in-degree ≤ 2 OR out-degree ≤ 2

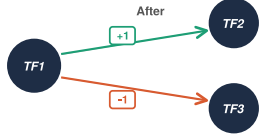
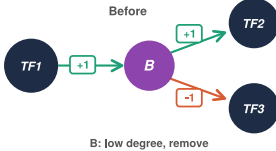
Case 1: $+1 \times +1 = +1$ (activation preserved)



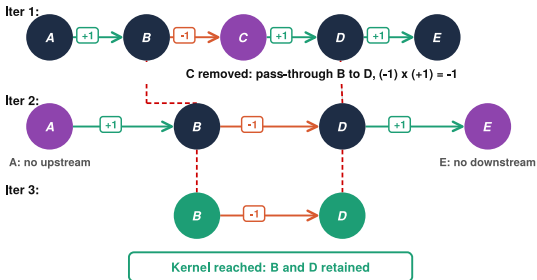
Case 2: $+1 \times -1 = -1$ (sign flips)



Case 3: multiple targets, all paths computed independently



Full iterative example



$\text{effect}(A \rightarrow C) = \text{effect}(A \rightarrow B) \times \text{effect}(B \rightarrow C)$

→ Activation (+1)

→ Inhibition (-1)

--- Node persistence

Removed node

Kernel node

7 activators, 3 repressors

ETS2

$$\begin{aligned} \frac{dETS2}{dt} = & \left(\alpha_1 + \gamma_1 \frac{HIF1A^{n_1}}{K_1^{n_1} + HIF1A^{n_1}} + \gamma_2 \frac{ETS1^{n_2}}{K_2^{n_2} + ETS1^{n_2}} + \gamma_3 \frac{JUN^{n_3}}{K_3^{n_3} + JUN^{n_3}} + \gamma_4 \frac{MYC^{n_4}}{K_4^{n_4} + MYC^{n_4}} + \gamma_5 \frac{STAT1^{n_5}}{K_5^{n_5} + STAT1^{n_5}} + \gamma_6 \frac{STAT3^{n_6}}{K_6^{n_6} + STAT3^{n_6}} + \gamma_7 \frac{TP53^{n_7}}{K_7^{n_7} + TP53^{n_7}} \right) \\ & \times \frac{\gamma_{r1}}{1 + \left(\frac{HIF1A}{K_{r1}} \right)^{n_{r1}}} \cdot \frac{\gamma_{r2}}{1 + \left(\frac{STAT3}{K_{r2}} \right)^{n_{r2}}} \cdot \frac{\gamma_{r3}}{1 + \left(\frac{VitD}{K_{r3}} \right)^{n_{r3}}} - \delta_1 ETS2 \end{aligned}$$

10 activators, 7 repressors

FOS

$$\begin{aligned} \frac{dFOS}{dt} = & \left(\alpha_2 + \gamma_8 \frac{ETS2^{n_8}}{K_8^{n_8} + ETS2^{n_8}} + \gamma_9 \frac{HIF1A^{n_9}}{K_9^{n_9} + HIF1A^{n_9}} + \gamma_{10} \frac{EGR1^{n_{10}}}{K_{10}^{n_{10}} + EGR1^{n_{10}}} + \gamma_{11} \frac{ETS1^{n_{11}}}{K_{11}^{n_{11}} + ETS1^{n_{11}}} + \gamma_{12} \frac{JUN^{n_{12}}}{K_{12}^{n_{12}} + JUN^{n_{12}}} + \gamma_{13} \frac{MYC^{n_{13}}}{K_{13}^{n_{13}} + MYC^{n_{13}}} \right. \\ & \left. + \gamma_{14} \frac{STAT1^{n_{14}}}{K_{14}^{n_{14}} + STAT1^{n_{14}}} + \gamma_{15} \frac{STAT3^{n_{15}}}{K_{15}^{n_{15}} + STAT3^{n_{15}}} + \gamma_{16} \frac{TP53^{n_{16}}}{K_{16}^{n_{16}} + TP53^{n_{16}}} + \gamma_{17} \frac{VitD^{n_{17}}}{K_{17}^{n_{17}} + VitD^{n_{17}}} \right) \\ & \times \frac{\gamma_{r4}}{1 + \left(\frac{HIF1A}{K_{r4}} \right)^{n_{r4}}} \cdot \frac{\gamma_{r5}}{1 + \left(\frac{JUN}{K_{r5}} \right)^{n_{r5}}} \cdot \frac{\gamma_{r6}}{1 + \left(\frac{MYC}{K_{r6}} \right)^{n_{r6}}} \cdot \frac{\gamma_{r7}}{1 + \left(\frac{STAT1}{K_{r7}} \right)^{n_{r7}}} \cdot \frac{\gamma_{r8}}{1 + \left(\frac{STAT3}{K_{r8}} \right)^{n_{r8}}} \cdot \frac{\gamma_{r9}}{1 + \left(\frac{TP53}{K_{r9}} \right)^{n_{r9}}} \cdot \frac{\gamma_{r10}}{1 + \left(\frac{VitD}{K_{r10}} \right)^{n_{r10}}} - \delta_2 FOS \end{aligned}$$

5 activators, 1 repressor

HIF1A

$$\frac{dHIF1A}{dt} = \left(\alpha_3 + \gamma_{18} \frac{EGR1^{n_{18}}}{K_{18}^{n_{18}} + EGR1^{n_{18}}} + \gamma_{19} \frac{MYC^{n_{19}}}{K_{19}^{n_{19}} + MYC^{n_{19}}} + \gamma_{20} \frac{NFKB1^{n_{20}}}{K_{20}^{n_{20}} + NFKB1^{n_{20}}} + \gamma_{21} \frac{STAT3^{n_{21}}}{K_{21}^{n_{21}} + STAT3^{n_{21}}} + \gamma_{22} \frac{VitD^{n_{22}}}{K_{22}^{n_{22}} + VitD^{n_{22}}} \right) \cdot \frac{\gamma_{r11}}{1 + \left(\frac{VitD}{K_{r11}} \right)^{n_{r11}}} - \delta_3 HIF1A$$

8 activators, 3 repressors

EGR1

$$\begin{aligned} \frac{dEGR1}{dt} = & \left(\alpha_4 + \gamma_{23} \frac{ETS2^{n_{23}}}{K_{23}^{n_{23}} + ETS2^{n_{23}}} + \gamma_{24} \frac{HIF1A^{n_{24}}}{K_{24}^{n_{24}} + HIF1A^{n_{24}}} + \gamma_{25} \frac{ETS1^{n_{25}}}{K_{25}^{n_{25}} + ETS1^{n_{25}}} + \gamma_{26} \frac{MYC^{n_{26}}}{K_{26}^{n_{26}} + MYC^{n_{26}}} + \gamma_{27} \frac{STAT1^{n_{27}}}{K_{27}^{n_{27}} + STAT1^{n_{27}}} \right. \\ & \left. + \gamma_{28} \frac{STAT3^{n_{28}}}{K_{28}^{n_{28}} + STAT3^{n_{28}}} + \gamma_{29} \frac{TP53^{n_{29}}}{K_{29}^{n_{29}} + TP53^{n_{29}}} + \gamma_{30} \frac{VitD^{n_{30}}}{K_{30}^{n_{30}} + VitD^{n_{30}}} \right) \\ & \times \frac{\gamma_{r12}}{1 + \left(\frac{HIF1A}{K_{r12}} \right)^{n_{r12}}} \cdot \frac{\gamma_{r13}}{1 + \left(\frac{STAT3}{K_{r13}} \right)^{n_{r13}}} \cdot \frac{\gamma_{r14}}{1 + \left(\frac{VitD}{K_{r14}} \right)^{n_{r14}}} - \delta_4 EGR1 \end{aligned}$$

4 activators, 8 repressors

ETS1

$$\begin{aligned} \frac{dETS1}{dt} = & \left(\alpha_5 + \gamma_{31} \frac{HIF1A^{n_{31}}}{K_{31}^{n_{31}} + HIF1A^{n_{31}}} + \gamma_{32} \frac{JUN^{n_{32}}}{K_{32}^{n_{32}} + JUN^{n_{32}}} + \gamma_{33} \frac{STAT3^{n_{33}}}{K_{33}^{n_{33}} + STAT3^{n_{33}}} + \gamma_{34} \frac{VitD^{n_{34}}}{K_{34}^{n_{34}} + VitD^{n_{34}}} \right) \\ & \times \frac{\gamma_{r15}}{1 + \left(\frac{FOS}{K_{r15}} \right)^{n_{r15}}} \cdot \frac{\gamma_{r16}}{1 + \left(\frac{HIF1A}{K_{r16}} \right)^{n_{r16}}} \cdot \frac{\gamma_{r17}}{1 + \left(\frac{JUN}{K_{r17}} \right)^{n_{r17}}} \cdot \frac{\gamma_{r18}}{1 + \left(\frac{MYC}{K_{r18}} \right)^{n_{r18}}} \cdot \frac{\gamma_{r19}}{1 + \left(\frac{STAT1}{K_{r19}} \right)^{n_{r19}}} \cdot \frac{\gamma_{r20}}{1 + \left(\frac{STAT3}{K_{r20}} \right)^{n_{r20}}} \cdot \frac{\gamma_{r21}}{1 + \left(\frac{TP53}{K_{r21}} \right)^{n_{r21}}} \cdot \frac{\gamma_{r22}}{1 + \left(\frac{VitD}{K_{r22}} \right)^{n_{r22}}} - \delta_5 ETS1 \end{aligned}$$

7 activators, 7 repressors

JUN

$$\begin{aligned} \frac{dJUN}{dt} = & \left(\alpha_6 + \gamma_{35} \frac{HIF1A^{n_{35}}}{K_{35}^{n_{35}} + HIF1A^{n_{35}}} + \gamma_{36} \frac{EGR1^{n_{36}}}{K_{36}^{n_{36}} + EGR1^{n_{36}}} + \gamma_{37} \frac{MYC^{n_{37}}}{K_{37}^{n_{37}} + MYC^{n_{37}}} + \gamma_{38} \frac{STAT1^{n_{38}}}{K_{38}^{n_{38}} + STAT1^{n_{38}}} + \gamma_{39} \frac{STAT3^{n_{39}}}{K_{39}^{n_{39}} + STAT3^{n_{39}}} + \gamma_{40} \frac{TP53^{n_{40}}}{K_{40}^{n_{40}} + TP53^{n_{40}}} + \gamma_{41} \frac{VitD^{n_{41}}}{K_{41}^{n_{41}} + VitD^{n_{41}}} \right) \\ & \times \frac{\gamma_{r23}}{1 + \left(\frac{FOS}{K_{r23}} \right)^{n_{r23}}} \cdot \frac{\gamma_{r24}}{1 + \left(\frac{HIF1A}{K_{r24}} \right)^{n_{r24}}} \cdot \frac{\gamma_{r25}}{1 + \left(\frac{MYC}{K_{r25}} \right)^{n_{r25}}} \cdot \frac{\gamma_{r26}}{1 + \left(\frac{STAT1}{K_{r26}} \right)^{n_{r26}}} \cdot \frac{\gamma_{r27}}{1 + \left(\frac{STAT3}{K_{r27}} \right)^{n_{r27}}} \cdot \frac{\gamma_{r28}}{1 + \left(\frac{TP53}{K_{r28}} \right)^{n_{r28}}} \cdot \frac{\gamma_{r29}}{1 + \left(\frac{VitD}{K_{r29}} \right)^{n_{r29}}} - \delta_6 JUN \end{aligned}$$

8 activators, 6 repressors

MYC

$$\begin{aligned} \frac{dMYC}{dt} = & \left(\alpha_7 + \gamma_{42} \frac{ETS2^{n_{42}}}{K_{42}^{n_{42}} + ETS2^{n_{42}}} + \gamma_{43} \frac{HIF1A^{n_{43}}}{K_{43}^{n_{43}} + HIF1A^{n_{43}}} + \gamma_{44} \frac{EGR1^{n_{44}}}{K_{44}^{n_{44}} + EGR1^{n_{44}}} + \gamma_{45} \frac{NFKB1^{n_{45}}}{K_{45}^{n_{45}} + NFKB1^{n_{45}}} + \gamma_{46} \frac{STAT1^{n_{46}}}{K_{46}^{n_{46}} + STAT1^{n_{46}}} \right. \\ & \left. + \gamma_{47} \frac{STAT3^{n_{47}}}{K_{47}^{n_{47}} + STAT3^{n_{47}}} + \gamma_{48} \frac{TP53^{n_{48}}}{K_{48}^{n_{48}} + TP53^{n_{48}}} + \gamma_{49} \frac{VitD^{n_{49}}}{K_{49}^{n_{49}} + VitD^{n_{49}}} \right) \\ & \times \frac{\gamma_{r30}}{1 + \left(\frac{FOS}{K_{r30}} \right)^{n_{r30}}} \cdot \frac{\gamma_{r31}}{1 + \left(\frac{HIF1A}{K_{r31}} \right)^{n_{r31}}} \cdot \frac{\gamma_{r32}}{1 + \left(\frac{JUN}{K_{r32}} \right)^{n_{r32}}} \cdot \frac{\gamma_{r33}}{1 + \left(\frac{STAT3}{K_{r33}} \right)^{n_{r33}}} \cdot \frac{\gamma_{r34}}{1 + \left(\frac{TP53}{K_{r34}} \right)^{n_{r34}}} \cdot \frac{\gamma_{r35}}{1 + \left(\frac{VitD}{K_{r35}} \right)^{n_{r35}}} - \delta_7 MYC \end{aligned}$$

7 activators, 4 repressors

NFKB1

$$\begin{aligned} \frac{dNFKB1}{dt} = & \left(\alpha_8 + \gamma_{50} \frac{HIF1A^{n_{50}}}{K_{50}^{n_{50}} + HIF1A^{n_{50}}} + \gamma_{51} \frac{EGR1^{n_{51}}}{K_{51}^{n_{51}} + EGR1^{n_{51}}} + \gamma_{52} \frac{ETS1^{n_{52}}}{K_{52}^{n_{52}} + ETS1^{n_{52}}} + \gamma_{53} \frac{MYC^{n_{53}}}{K_{53}^{n_{53}} + MYC^{n_{53}}} + \gamma_{54} \frac{STAT1^{n_{54}}}{K_{54}^{n_{54}} + STAT1^{n_{54}}} + \gamma_{55} \frac{STAT3^{n_{55}}}{K_{55}^{n_{55}} + STAT3^{n_{55}}} + \gamma_{56} \frac{TP53^{n_{56}}}{K_{56}^{n_{56}} + TP53^{n_{56}}} \right) \\ & \times \frac{\gamma_{r36}}{1 + \left(\frac{FOS}{K_{r36}} \right)^{n_{r36}}} \cdot \frac{\gamma_{r37}}{1 + \left(\frac{HIF1A}{K_{r37}} \right)^{n_{r37}}} \cdot \frac{\gamma_{r38}}{1 + \left(\frac{STAT3}{K_{r38}} \right)^{n_{r38}}} \cdot \frac{\gamma_{r39}}{1 + \left(\frac{VitD}{K_{r39}} \right)^{n_{r39}}} - \delta_8 NFKB1 \end{aligned}$$

3 activators, 0 repressors

STAT1

$$\frac{dSTAT1}{dt} = \left(\alpha_9 + \gamma_{57} \frac{STAT3^{n_{57}}}{K_{57}^{n_{57}} + STAT3^{n_{57}}} + \gamma_{58} \frac{TP53^{n_{58}}}{K_{58}^{n_{58}} + TP53^{n_{58}}} + \gamma_{59} \frac{VitD^{n_{59}}}{K_{59}^{n_{59}} + VitD^{n_{59}}} \right) - \delta_9 STAT1$$

2 activators, 1 repressor

STAT3

$$\frac{dSTAT3}{dt} = \left(\alpha_{10} + \gamma_{60} \frac{STAT1^{n_{60}}}{K_{60}^{n_{60}} + STAT1^{n_{60}}} + \gamma_{61} \frac{VitD^{n_{61}}}{K_{61}^{n_{61}} + VitD^{n_{61}}} \right) \cdot \frac{\gamma_{r40}}{1 + \left(\frac{TP53}{K_{r40}} \right)^{n_{r40}}} - \delta_{10} STAT3$$

10 activators, 6 repressors

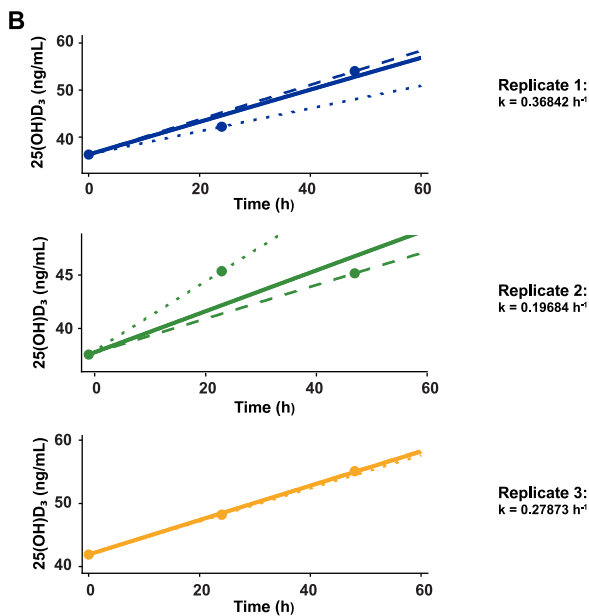
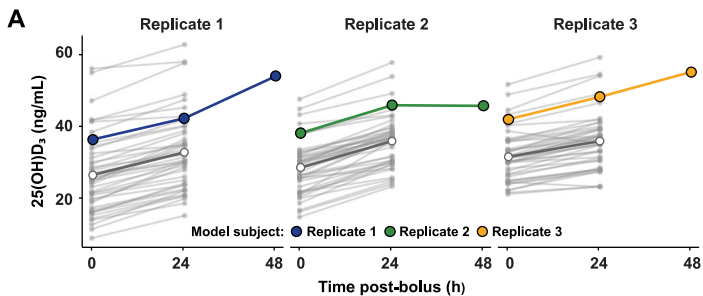
TP53

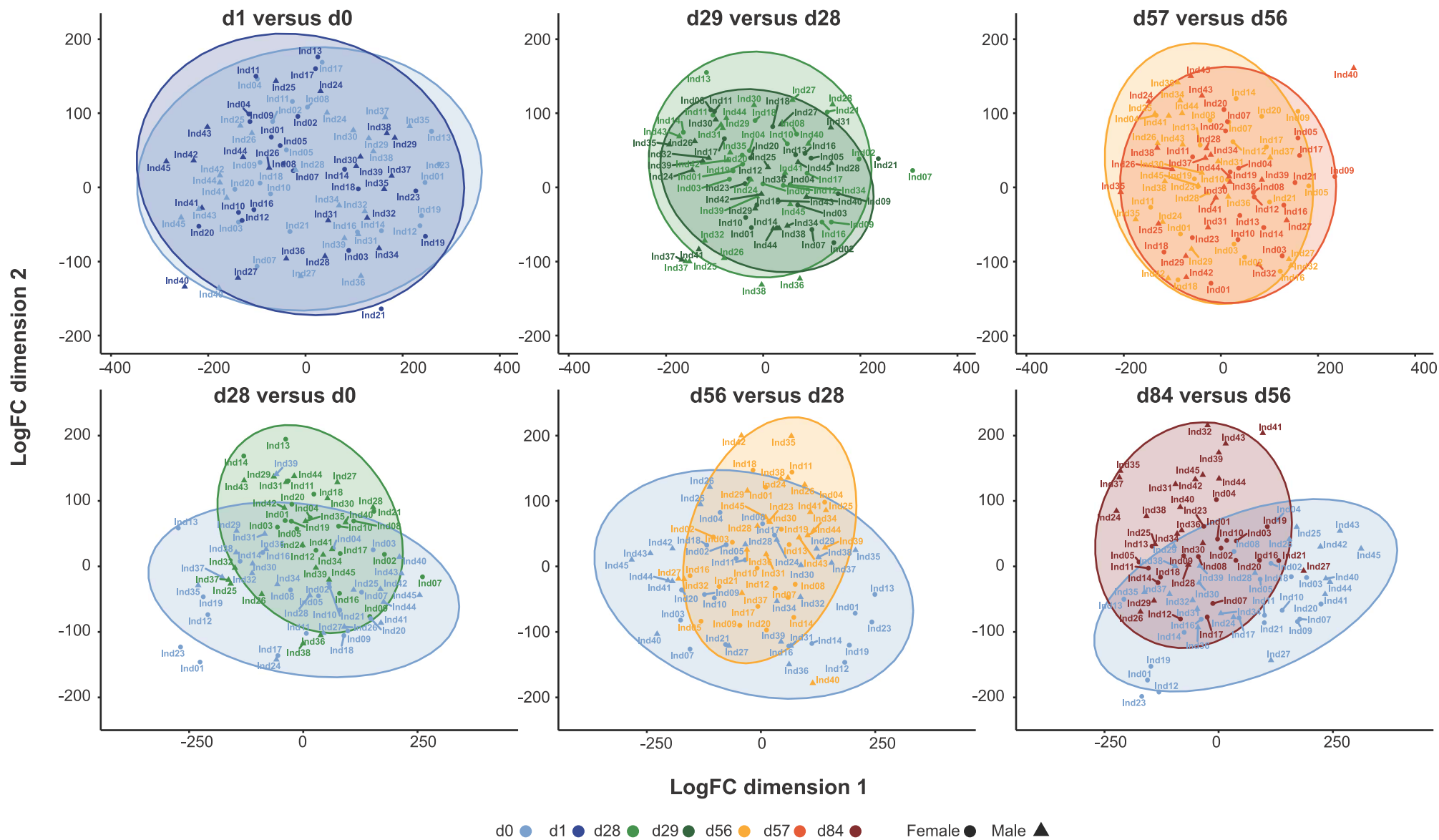
$$\begin{aligned} \frac{dTP53}{dt} = & \left(\alpha_{11} + \gamma_{62} \frac{ETS2^{n_{62}}}{K_{62}^{n_{62}} + ETS2^{n_{62}}} + \gamma_{63} \frac{FOS^{n_{63}}}{K_{63}^{n_{63}} + FOS^{n_{63}}} + \gamma_{64} \frac{HIF1A^{n_{64}}}{K_{64}^{n_{64}} + HIF1A^{n_{64}}} + \gamma_{65} \frac{EGR1^{n_{65}}}{K_{65}^{n_{65}} + EGR1^{n_{65}}} + \gamma_{66} \frac{ETS1^{n_{66}}}{K_{66}^{n_{66}} + ETS1^{n_{66}}} \right. \\ & \left. + \gamma_{67} \frac{MYC^{n_{67}}}{K_{67}^{n_{67}} + MYC^{n_{67}}} + \gamma_{68} \frac{NFKB1^{n_{68}}}{K_{68}^{n_{68}} + NFKB1^{n_{68}}} + \gamma_{69} \frac{STAT1^{n_{69}}}{K_{69}^{n_{69}} + STAT1^{n_{69}}} + \gamma_{70} \frac{STAT3^{n_{70}}}{K_{70}^{n_{70}} + STAT3^{n_{70}}} + \gamma_{71} \frac{VitD^{n_{71}}}{K_{71}^{n_{71}} + VitD^{n_{71}}} \right) \\ & \times \frac{\gamma_{r41}}{1 + \left(\frac{HIF1A}{K_{r41}} \right)^{n_{r41}}} \cdot \frac{\gamma_{r42}}{1 + \left(\frac{JUN}{K_{r42}} \right)^{n_{r42}}} \cdot \frac{\gamma_{r43}}{1 + \left(\frac{MYC}{K_{r43}} \right)^{n_{r43}}} \cdot \frac{\gamma_{r44}}{1 + \left(\frac{STAT1}{K_{r44}} \right)^{n_{r44}}} \cdot \frac{\gamma_{r45}}{1 + \left(\frac{STAT3}{K_{r45}} \right)^{n_{r45}}} \cdot \frac{\gamma_{r46}}{1 + \left(\frac{VitD}{K_{r46}} \right)^{n_{r46}}} - \delta_{11} TP53 \end{aligned}$$

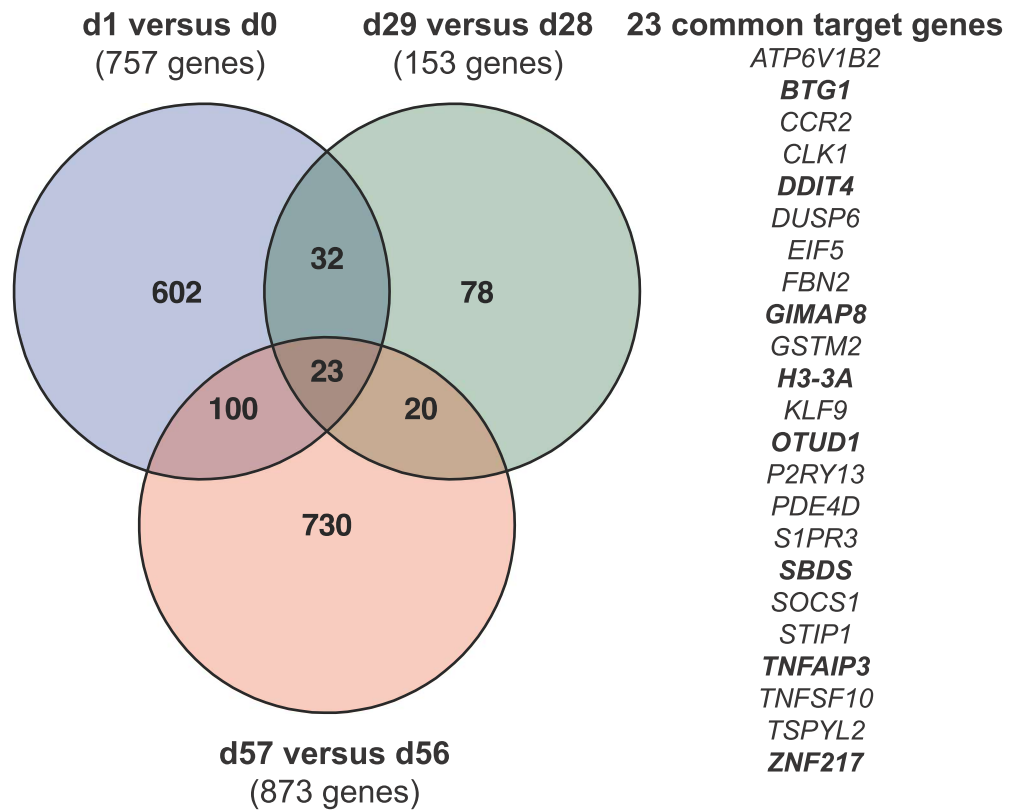
Fixed parameters, not optimized

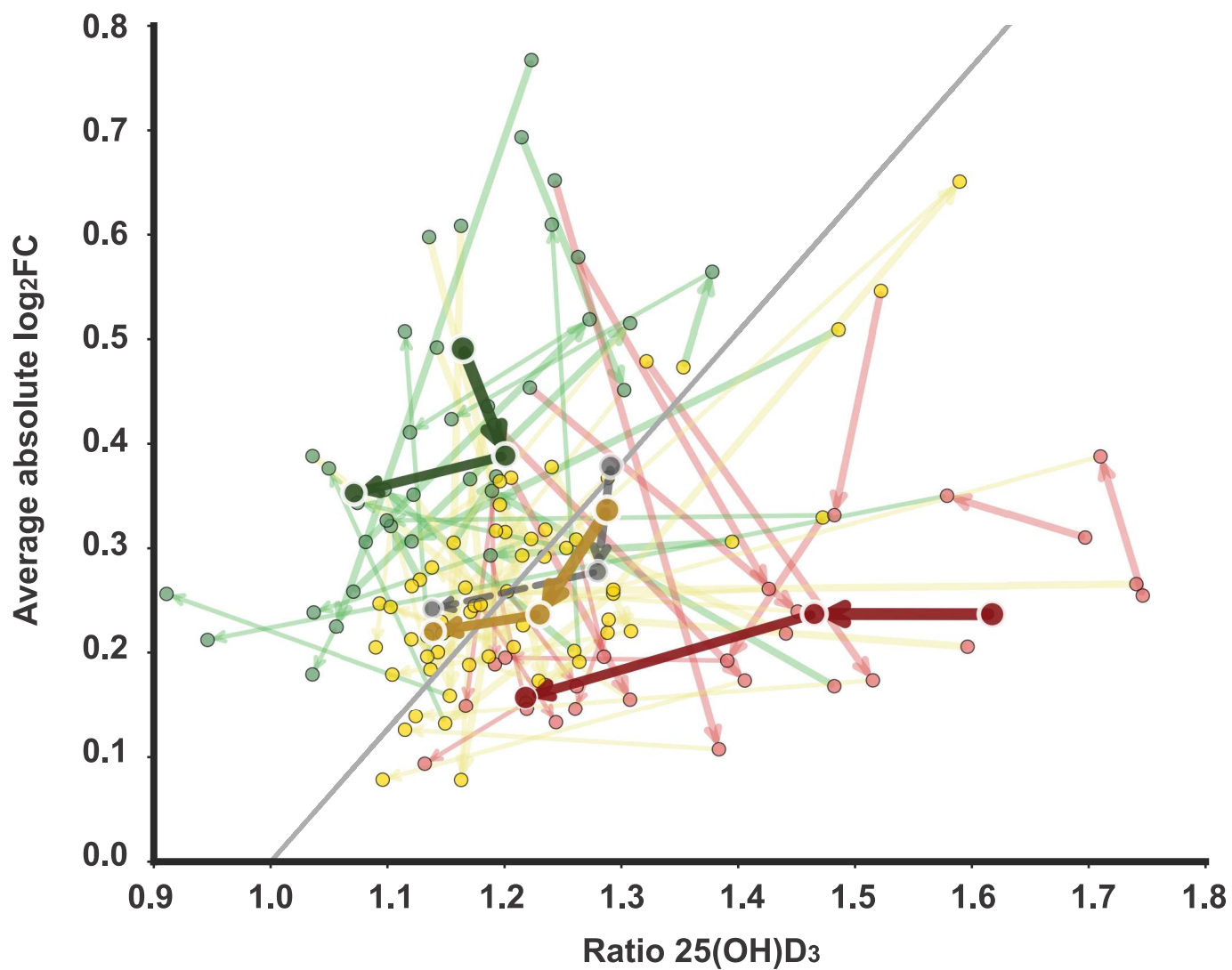
25(OH)D₃

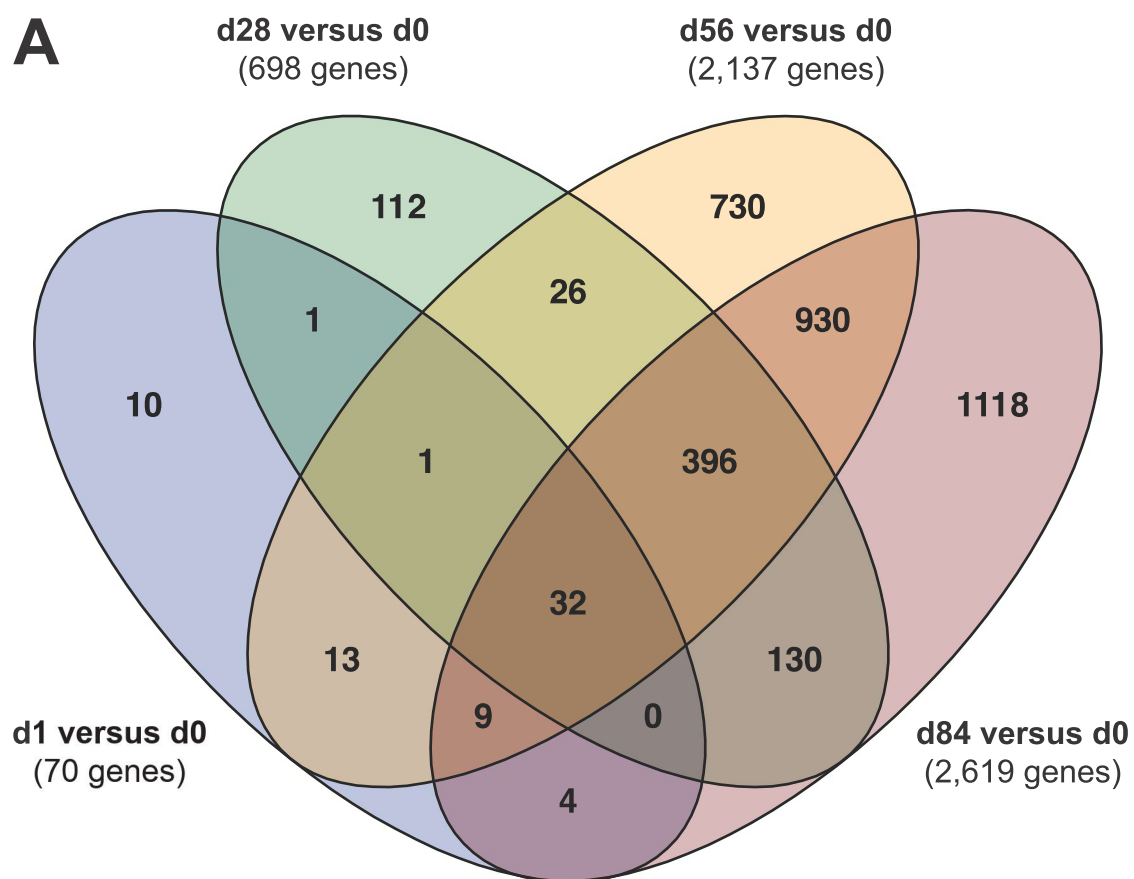
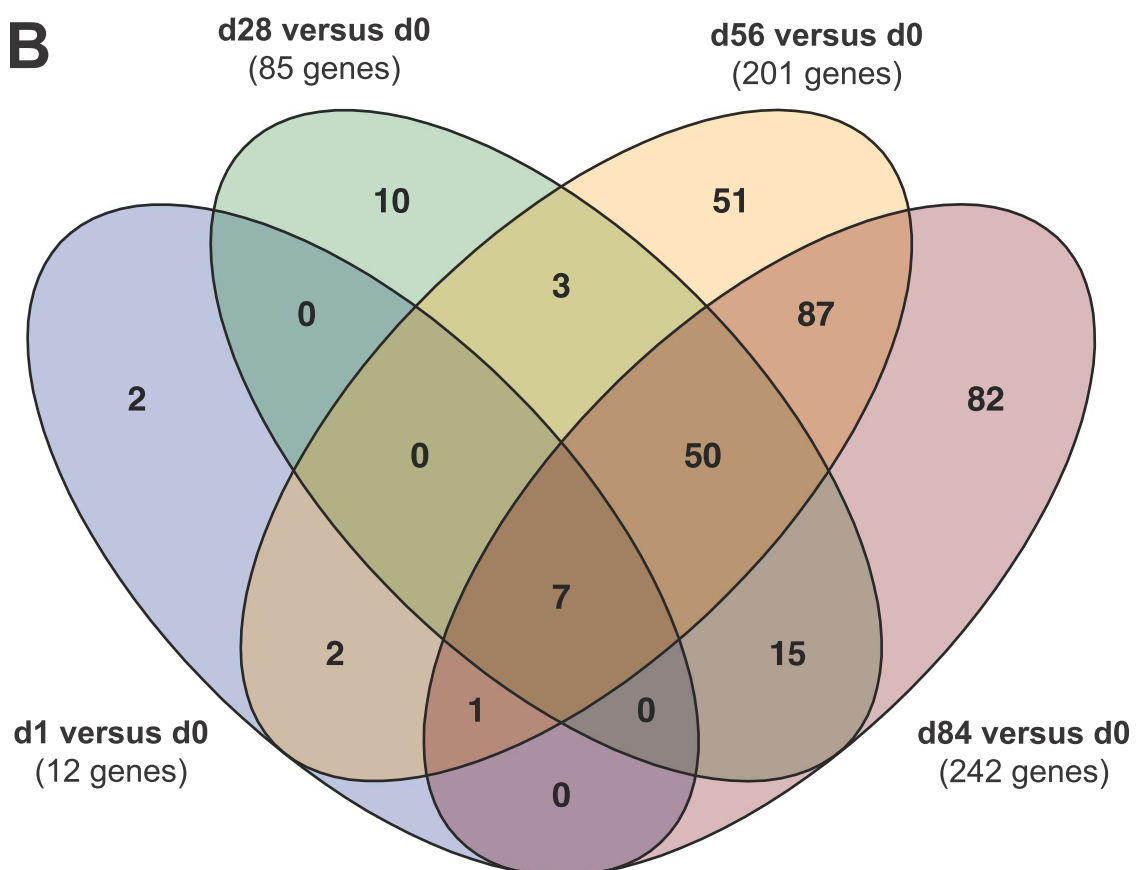
$$\frac{dVitD}{dt} = k_1 - d_1 \cdot VitD$$



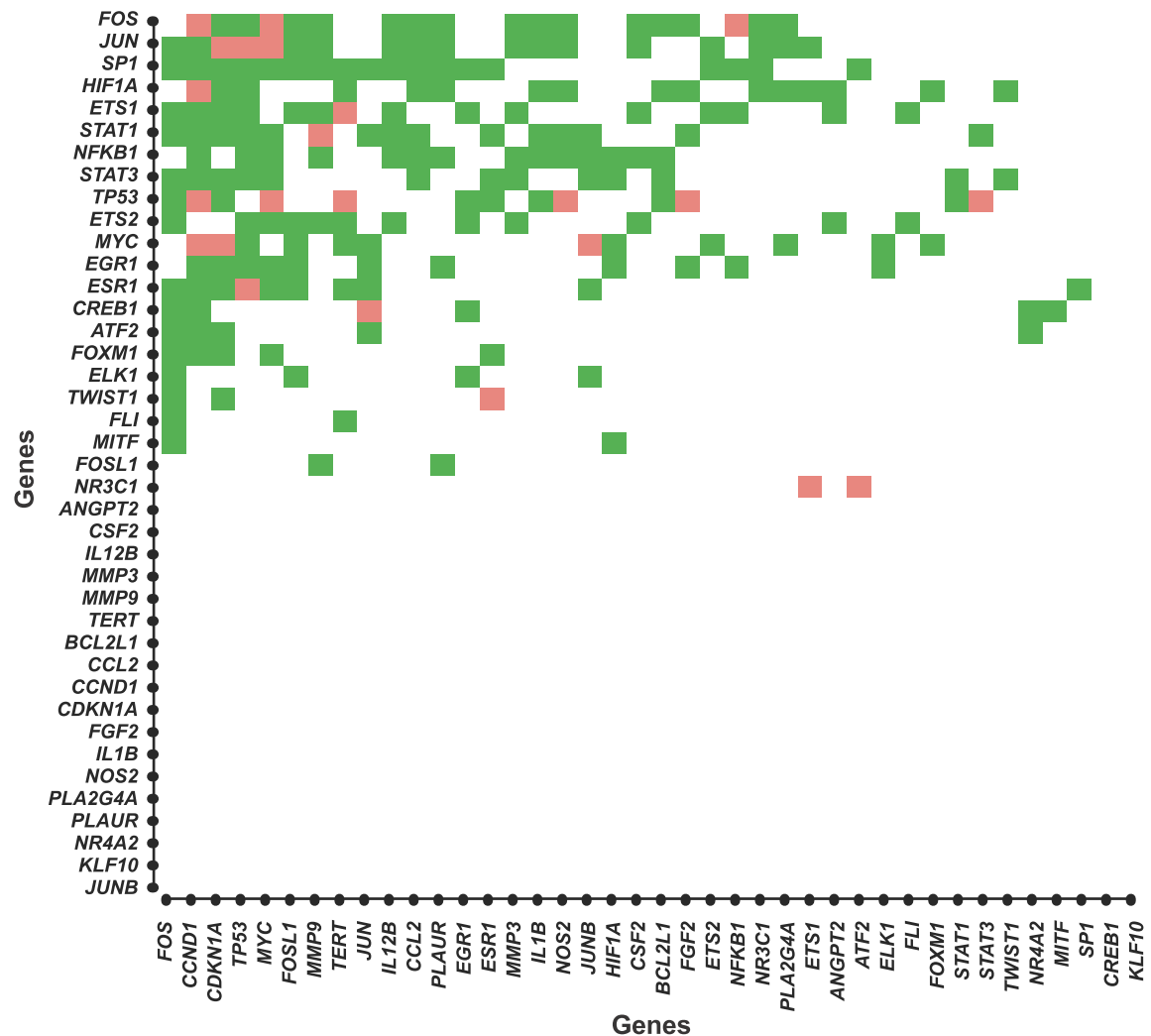






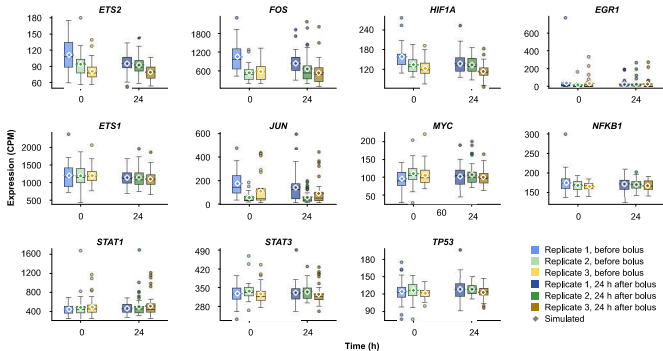
A**B**

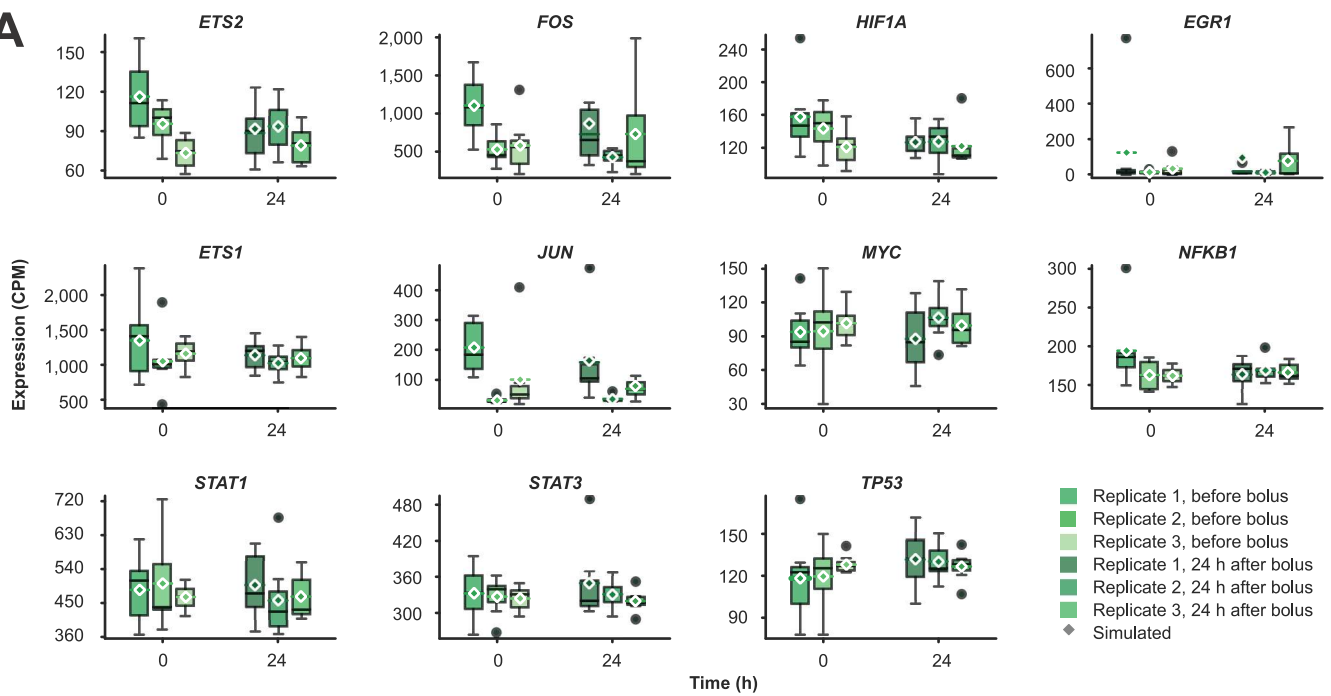
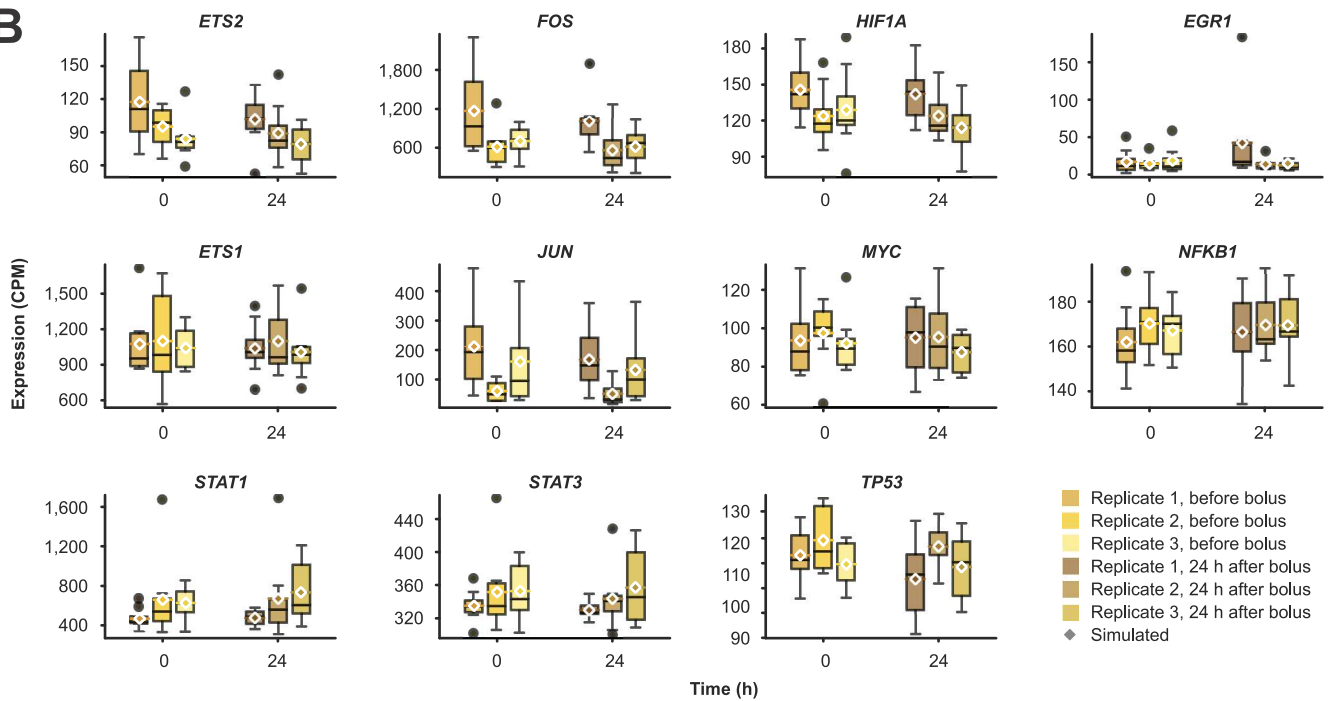
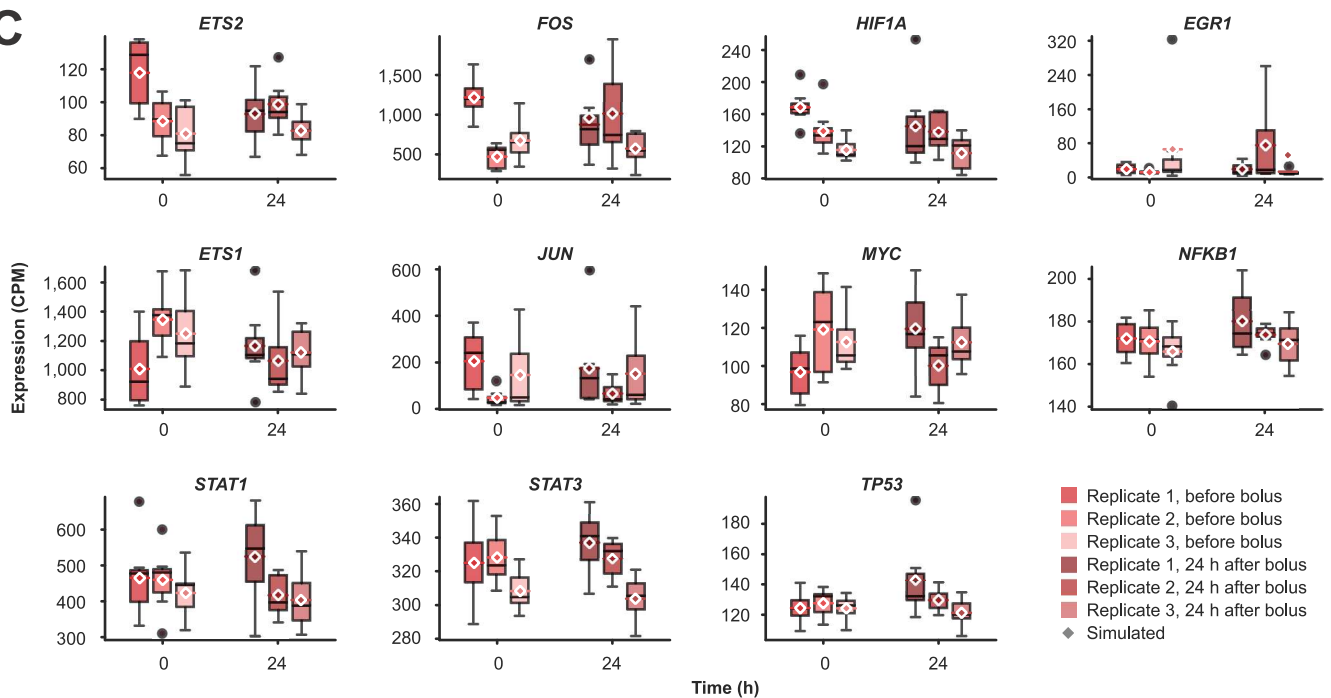
A

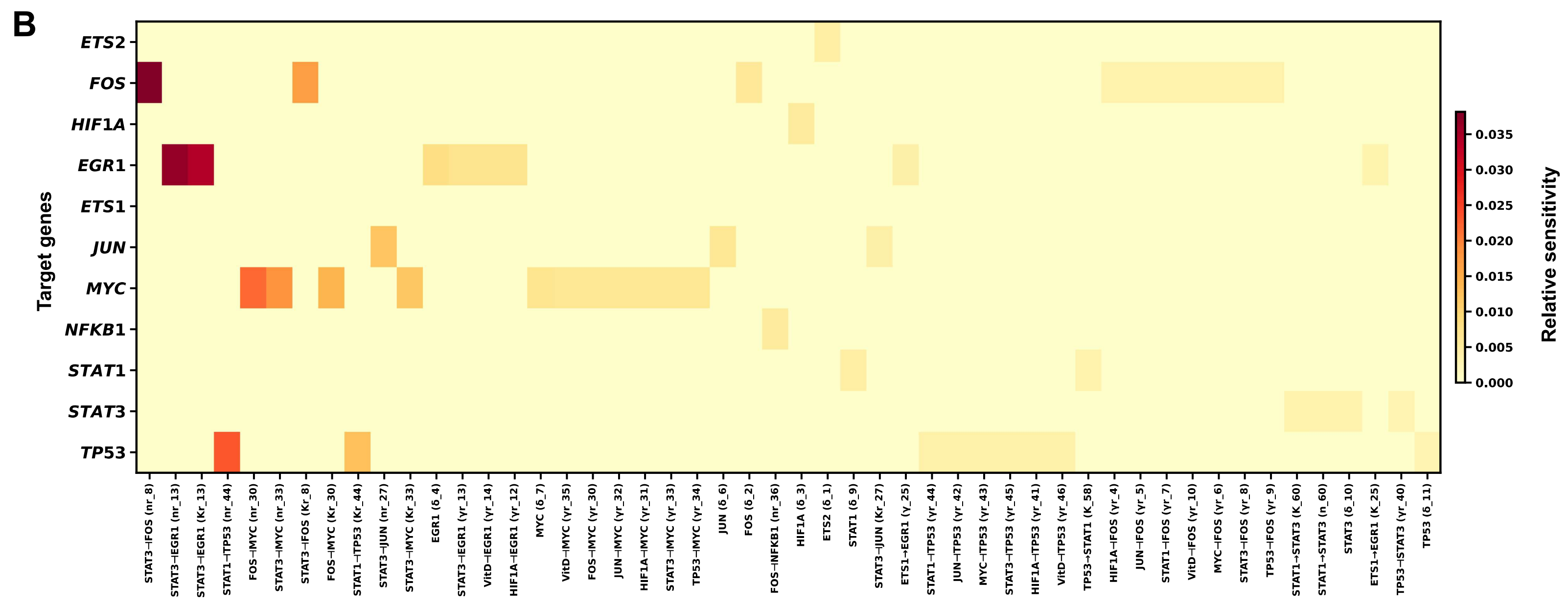
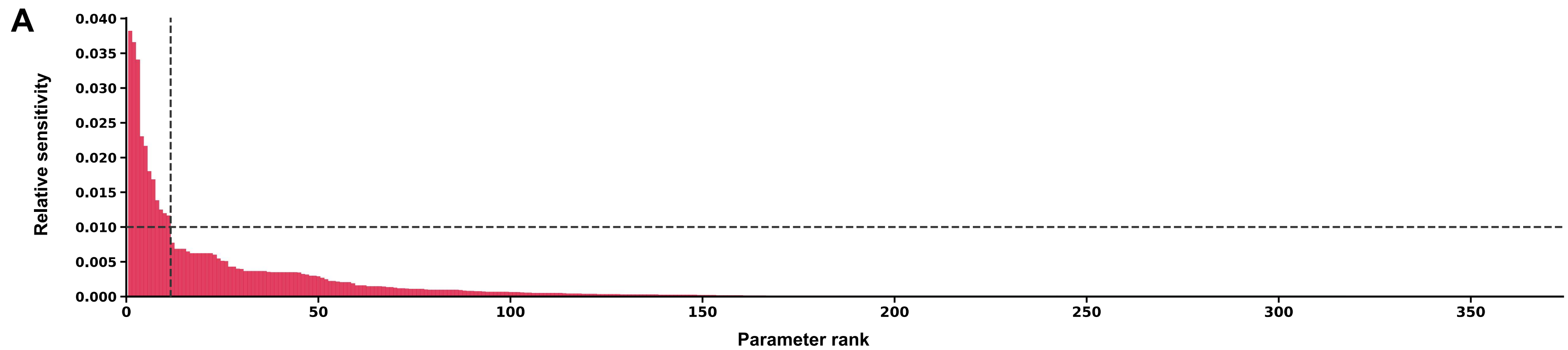


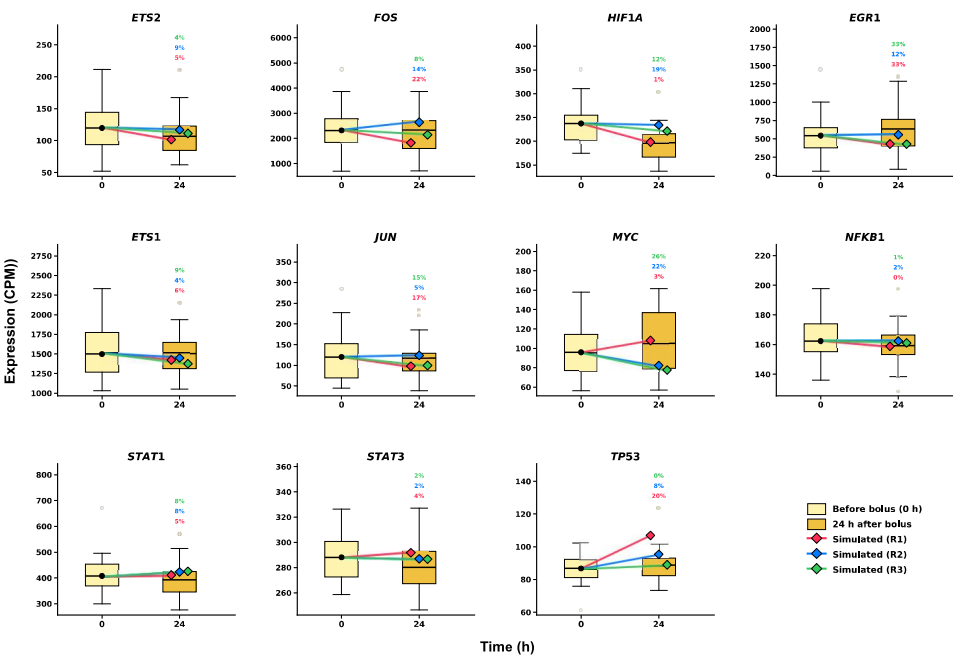
B

Genes	ETS2	FOS	HIF1A	EGR1	ETS1	JUN	MYC	NFKB1	STAT1	STAT3	TP53	25(OH)D ₃
ETS2	0	0	1, -1	0	1	1	2	0	1	1, -1	1	-3
FOS	2	0	6, -5	1	2	1, -1	5, -1	0	4, -1	6, -4	4, -1	7, -6
HIF1A	0	0	0	1	0	0	1	1	0	1	0	1, -1
EGR1	1	0	1, -1	0	1	0	2	0	1	1, -1	2	2, -2
ETS1	0	-1	2, -2	0	0	1, -1	-1	0	-1	1, -1	-1	3, -1
JUN	0	-1	4, -5	1	0	0	4, -1	0	4, -1	4, -4	3, -1	3, -6
MYC	1	-1	3, -2	1	0	-1	0	1	3	3, -2	2, -1	1, -3
NFKB1	0	-1	1, -1	1	1	0	1	0	1	1, -1	1	-2
STAT1	0	0	0	0	0	0	0	0	0	1	1	1
STAT3	0	0	0	0	0	0	0	0	1	0	-1	1
TP53	1	1	3, -2	1	1	-1	2, -1	1	2, -1	3, -2	0	1, -2



A**B****C**





Supplementary Material

Supplementary Tables

Table S1. Participant characteristics and individual serum 25(OH)D₃ responses in the VitDPAS study. Age, sex, BMI, and serum 25(OH)D₃ concentrations of the 41 participants measured before (d0, d28, d56) and 24 hours after each vitamin D₃ bolus (d1, d29, d57), as well as at the final follow-up (d84). The table also shows the overall change in vitamin D status during the 12-week intervention (d84-d0), the average increase after each weight-adjusted vitamin D₃ bolus (1,000 IU/kg), and the classification of participants according to their vitamin D response index (**Figure 3**). Rows are colored according to response index classification in low (red), mid (yellow) and high (green). Participants selected for modeling of high, mid, and low responders are highlighted in bold.

Table S2. Sequencing depth of RNA-seq samples from the VitDPAS study. Summary of RNA sequencing coverage for all samples analyzed in the VitDPAS intervention study. The table lists the sample name, the sampling day, the participant identifier, and the number of assigned sequencing reads (in millions) obtained after read mapping and quality filtering.

Table S3. Differentially expressed genes in response to vitamin D₃ supplementation. List of genes significantly differentially expressed between the indicated sampling time points of the VitDPAS study. The table reports gene identifiers together with statistical metrics derived from the differential expression analysis, including log₂FC, p-value, and FDR. Comparisons correspond to the different supplementation cycles (d1 versus d0, d29 versus d28, d57 versus d56) and longer-term comparisons across the intervention period (d28 versus d0, d56 versus d0, d84 versus d0). Positive log₂FC values indicate upregulated genes, whereas negative values indicate downregulated genes following vitamin D₃ supplementation. All 19,295 protein-coding genes are listed, but statistical analysis focused on 8,378 genes with

average CPM > 10 and are expressed in all individuals (allowing all six time point comparisons; highlighted in green).

Table S4. Network-based prioritization and parametric characterization of vitamin D-responsive regulatory interactions. This multi-sheet table summarizes the quantitative analysis of the vitamin D-responsive gene regulatory network. **Sheet 1:** Ranking of network genes based on composite importance scores integrating normalized connectivity and normalized interaction strength. Additional annotations indicate whether genes are classified as vitamin D-regulated targets. **Sheet 2:** Identification of regulatory interactions that most strongly distinguish high from low vitamin D responders. Interactions are ranked based on normalized differential contributions across response groups, with annotation of vitamin D-related regulatory links. **Sheets 3-5:** Parameter values for all modeled regulatory interactions across three independent model runs (R1-R3). Values are provided for combined data as well as for high, mid, and low responder submodels. Interactions directly involving vitamin D signaling are indicated. **Sheet 6:** Methodological description of the parametric analysis workflow, including definitions of network metrics, normalization procedures, interaction filtering criteria, and model interpretation guidelines.

Supplementary Figures

Figure S1. Principle of *kernel reduction* for refinement of the vitamin D-responsive gene regulatory network. Schematic illustration of the *kernel reduction* procedure applied to simplify the transcriptional regulatory network. Starting from the expanded interaction network, genes and regulatory connections are iteratively filtered based on predefined criteria, such as interaction support and connectivity, to identify a core regulatory module (kernel). Nodes represent genes, and edges indicate predicted regulatory interactions. The reduction process retains genes that form strongly interconnected regulatory relationships while removing peripheral or weakly supported interactions.

Figure S2. Mathematical model of the core regulatory network. ODEs used for modeling of the 11 key transcription factor genes and 25(OH)D₃ levels. Values of the respective parameters can be found in **Table S4**.

Figure S3. Pharmacokinetic modeling of serum 25(OH)D₃ responses to repeated vitamin D₃ bolus supplementation. Comparison of observed changes in serum 25(OH)D₃ concentrations following weight-adjusted vitamin D₃ bolus administration in the VitDPAS cohort (n = 41) and in a previously reported single-participant study using the same triple-bolus regimen (32) (**A**). Individual measurements at baseline and 24 hours after supplementation are shown. Pharmacokinetic model-based estimation of 25(OH)D₃ production derived from the single-participant dataset, including the additional 48-hour post-bolus sampling point (**B**). Model fitting enabled calculation of absorption rate constants and predicted steady-state serum 25(OH)D₃ concentrations, providing quantitative parameters for describing interindividual variability in vitamin D status responses.

Figure S4. Principal component analysis of individual transcriptional responses across vitamin D₃ supplementation cycles. MDS plots derived from genome-wide log₂FC values illustrate similarities and differences in transcriptional responses among study participants. Each point represents an individual participant (Ind01-Ind45),

positioned according to the first two MDS dimensions, which capture the major sources of variation in gene expression responses. Panels correspond to the six pairwise comparisons between study time points. Points are colored according to the respective comparison, and ellipses indicate the distribution of individuals within each condition, highlighting the overall structure of transcriptional variation across the cohort. Symbols denote participant sex.

Figure S5. Overlap of vitamin D-responsive genes across supplementation cycles. Venn diagram illustrating the overlap of genes significantly regulated ($p < 0.05$) in the three 24-hour transcriptional response comparisons following each vitamin D₃ bolus supplementation. The 23 genes consistently responsive across all three cycles are listed. Genes that additionally met the multiple-testing-adjusted significance threshold ($FDR < 0.05$) after the first bolus are highlighted in bold, indicating the most robust primary vitamin D targets.

Figure S6. Classification of VitDPAS study participants based on repeated vitamin D₃ bolus experiments. The 41 participants of the VitDPAS study were classified as high (green), mid (yellow), or low (red) responders according to the expression profiles of 23 *in vivo* vitamin D target genes that were consistently regulated across three independent cycles of the vitamin D₃ bolus supplementation experiment (**Figure S5**). Each individual is represented by three nodes corresponding to supplementation cycles 1, 2, and 3, which are connected by two edges. The average response patterns of high, mid, and low responder groups are shown as solid nodes and edges in the respective colors. The overall average behavior of all participants is summarized by grey solid nodes and edges.

Figure S7. Overlap of vitamin D₃-regulated genes across supplementation time points. Venn diagrams showing the overlap of significantly regulated vitamin D target genes ($FDR < 0.05$) identified at four time points following vitamin D₃ supplementation (d1, d28, d56, and d84 versus baseline d0). Panel **A** includes all regulated target genes, whereas panel **B** is restricted to the subset encoding transcription factors. The

total number of regulated genes at each time point is indicated in parentheses. The intersection across all four comparisons represents a core set of genes that are persistently regulated during both short- and long-term supplementation.

Figure S8. Functional characterization of the vitamin D-responsive regulatory gene module. Interaction matrix of the 40 genes included in the gene regulatory network, indicating the predicted regulatory effects of each gene on the others (**A**). Green squares denote activating interactions, red squares indicate repressive interactions, and white cells represent the absence of a regulatory effect. Interaction matrix of the 11 remaining genes after *kernel reduction* (**B**). The number of activating (green) and repressing (red) interactions are given.

Figure S9. ODE-based regulatory network model accurately reproduces transcription factor expression dynamics for all VitDPAS participants. Boxplots show experimentally measured gene expression levels of all 41 VitDPAS study participants at baseline (0 hours) and 24 hours after vitamin D₃ bolus administration across three experimental cycles. Colored boxes represent cycle-specific measurements before (light shades) and after (dark shades) supplementation, while diamonds indicate simulated expression values generated by the ODE-based Hill-function model.

Figure S10. Subgroup-specific ODE models reproduce transcription factor expression dynamics in high, mid, and low vitamin D responders. Boxplots show experimentally measured expression levels of the 11 transcription factor genes at baseline (0 hours) and 24 hours after vitamin D₃ bolus administration across three experimental cycles. Panels represent responder subgroups: high responders (**A**), mid responders (**B**), and low responders (**C**). Colored boxes indicate supplementation cycle-specific measurements before (lighter shades) and after (darker shades) supplementation, while diamonds denote simulated expression values generated by subgroup-specific ODE models based on optimized parameters.

Figure S11. Parameter sensitivity analysis. Ranked sensitivity scores for all 373 parameters; dashed line indicates moderate threshold ($S = 0.01$). The blue vertical line marks the top 50 parameters shown in panel B (**A**). Heatmap of top 50 parameters across 11 genes, annotated with regulatory interactions. STAT3-, MYC-, and TP53-centered repressive interactions dominate (**B**).

Figure S12. Out-of-sample validation against the independent VitDHiD cohort. Boxplots show the distribution of expression levels across 25 VitDHiD participants at baseline (0 h, light boxes) and 24 hours after vitamin D₃ bolus administration (24 h, dark boxes) for each of the 11 transcription factor genes. Cohort-mean baseline expression served as the initial condition for the ODE model (black circles). Colored diamonds connected by lines indicate predicted 24-hour expression using R1 (red), R2 (blue), and R3 (green) parameter sets from the VitDPAS-trained model with VitDPAS-derived 48-hour pharmacokinetics. The model quantitatively predicts post-bolus expression levels across an independent cohort ($R^2 = 0.978$ with R2 parameters).