

Dynamic Epigenetic Changes during a Relapse and Recovery Cycle in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome

Helliwell A.M.¹, Stockwell P.A.², Edgar C.D.¹, Chatterjee A.^{2*}, & Tate W.P.^{1*#}.
Departments of Biochemistry¹ and Pathology², University of Otago, Dunedin 9016, New Zealand

*joint senior authors

Corresponding author.(warren.tate@otago.ac.nz)

Abstract (242)

Background

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a complex disease with variable severity throughout the ongoing illness. Patients experience relapses where symptoms increase in severity, leaving them with a marked reduction in quality of life. Previous work has investigated molecular differences between ME/CFS patients and healthy controls, but the dynamic changes specific to each individual patient are unknown. Precision medicine can determine how each patient responds individually during variations in their long-term illness. We apply precision medicine here to map genomic changes in two selected ME/CFS patients through a relapse recovery cycle.

Results

DNA was isolated from Peripheral Blood Mononuclear Cells (PBMCs) from two patients and a healthy age/gender matched control in a longitudinal study to capture a patient relapse. Reduced representation DNA methylation sequencing profiles were obtained from each time point spanning the relapse recovery cycle. Both patients throughout the time course showed a significantly larger methylome variability (10-20 fold) compared with the control. During the relapse changes in the methylome profiles of the two patients were detected in regulatory-active regions of the genome that were associated respectively with 157 and 127 downstream genes, indicating disturbed metabolic, immune and inflammatory functions occurring during the relapse.

Conclusions

Severe health relapses in ME/CFS patients result in functionally important changes in their DNA methylomes that, while differing among patients, lead to similar compromised physiology. DNA methylation that is a signature of disease variability in ongoing ME/CFS may have practical applications for strategies to decrease relapse frequency.

32

Keywords: ME/CFS, DNA Methylation, RRBS, DMAP, Epigenetics

34 Introduction

35

36 Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a lifelong severely debilitating disease
 37 from which only a small proportion of individuals eventually fully recover (<5%) (1). While currently not well
 38 understood, it is estimated to have a global prevalence of ~1% (2) and to have a higher overall disease burden
 39 than conditions like multiple sclerosis, autism or HIV/AIDS (3). Patients experience a wide variety of
 40 debilitating symptoms including severe fatigue, post exertional malaise, and cognitive, sleep and orthostatic
 41 dysfunctions (4). These symptoms vary in severity such that ~25% of patients are house or bedbound
 42 throughout the illness. The remaining 75% of those affected transition to a life-long chronic phase where they
 43 may be able to participate in work and hobbies, albeit with a reduced capacity. However, they are vulnerable to
 44 frequent debilitating “relapses”, particularly after even minor stress.

45

46 The disease presentation and key research indicate that there is a complex pathophysiology affecting ME/CFS
 47 patients, with biological functions reduced in a number of systems including immune/inflammatory, and
 48 neurological as well as in metabolism. For example, a recent analysis found 80% of 612 metabolites analysed in
 49 plasma of ME/CFS patients were significantly decreased, indicating that there was an overall reduction in
 50 metabolic activity in patients. This has been compared to the ‘dauer effect’, a shut down like hibernation in
 51 animals (5). Additionally, patient mitochondria have a dysfunction in the mitochondrial complex V where the
 52 energy molecule ATP is synthesised, resulting in proteins in the upstream mitochondrial complexes, particularly
 53 complex I, being up regulated as an apparent attempt to compensate (6), as well the metabolic pathways and
 54 mechanisms to control regulation of reactive oxygen species (7). This could explain one component of why
 55 ME/CFS patients are unable to respond biologically to day-to-day stresses, let alone high-level stress events.
 56 Recently, we have proposed a theory based on fluctuating neuroinflammation to explain the sustained chronic
 57 state of the illness and ‘relapse recovery’ cycles (8). It hypothesises that neuroinflammation of the
 58 hypothalamus’ stress centre within the paraventricular nucleus could be responsible for the unexplained
 59 prolonged and fluctuating symptom presentations (8).

60

61 The absence of a specific molecular diagnostic test, and also the fluctuating variations across patients in disease
 62 presentation and symptom severity, has made understanding ME/CFS and conclusions from molecular data
 63 difficult. When studies focus on large cohorts of patients, they often include very varied presentations of

ME/CFS, and different studies have often used different clinical case definitions for diagnosis. A recent study of the involvement of cytokines addressed the diversity within the patient cohorts on a molecular scale, and found that a large number of pro-inflammatory cytokines were found to be linearly associated with ME/CFS severity (9). This investigation highlighted the dilemma of many investigations targeting ME/CFS, since the cytokines associated with severity were not sufficient to distinguish the patients from the controls because of the abundance of mild ME/CFS cases in the study group. Other studies have approached this issue by classifying ME/CFS patients into different subgroups. A study in 2008 examining expression levels of transcripts classified 7 subtypes, through mean relative transcript quantities 88 transcripts that corresponded with clinical severity (10). In recent years DNA methylation has been applied to investigate the disease status of ME/CFS patients, and a number of studies have found important differences separating the patients from controls (11–16). A recent publication with this technology has identified four subtypes utilising DNA methylation and symptom severity (17), with key differentially methylated genes between subtypes having primarily immune and metabolic functions. This study indicated that molecular analyses could differentiate patients from the molecular changes reflecting physiology relevant to the observed symptoms.

The gradual trend to a more personalised approach taken by these investigations is an important step towards understanding the intricacies of ME/CFS. Many patients, once they have entered the chronic state of their disease following an initial acute period often of several years' duration, experience frequent extreme symptom fluctuations characteristic of a relapse event. No published molecular studies have yet followed patients through a 'relapse recovery' cycle. In order to understand in depth a disease as complex as ME/CFS this more personalized approach is more informative and appropriate, both for researchers in their studies and for patients in the management of their disease. Indeed precision medicine is becoming more readily accessible not only as a research tool to understand the impact of disease on an individual, but also how they will respond to a specific medical intervention (18,19). This seems especially relevant for the study of ME/CFS where patients have a wide-ranging level of functionality, for example in their ability to exercise, their cognitive deficits, and often different comorbidities.

How can precision medicine be applied to ME/CFS? DNA methylation is an important epigenetic modification that affects the expression of genes without altering the genomic code itself. This specific analysis is with increasing precision helping researchers to bridge the gap between understanding genetic risk and assessing

environmental contributions to disease. Changes are captured that are not permanently reflected in the genome but occur in individuals as a result of the disease. An excellent example of this is an in-depth investigation that followed 87 individuals that had transitioned from a pre-diabetic state to a diabetic state. It revealed methylation differences that occurred before the switch to the disease state (20). Since DNA methylation can capture transcriptional changes that reflect physiological variations, it is an ideal tool in ME/CFS to determine temporal changes in genomic regions that reflect the symptom fluctuations. These are not so easily detectable or may not yet be present in single time point proteome or transcriptome analyses. Understanding variation in an individual's dynamic epigenetic code with sampling over a precise time period time can provide an insight into the molecular activity and course of their disease.

103

104 **Results**

105

106 **Study Design and Participants:**

107

Blood was taken from two ME/CFS patients and a healthy age matched control at 5 spaced time points spanning an eleven-month period that captured a health relapse in the ME/CFS patients from their typical compromised health state (see Figure 1A). The participants gave a subjective numerical assessment on a scale of 1- 10 as an indicator of their relative health (21). As is shown in Figure 1A patient 1 (hereafter referred to as P1) showed a drop from a relatively good health state, self-ranked as '7' ; 'well' - at the first sample time (A) to an off scale '-3' and '-2', 'fragile' indicating a severe relapse condition at time points (B) and (C). She then showed relative recovery to a health status '7' again at sample times (D) and (E). Patient 2 (hereafter referred to as P2) in contrast had a more fragile steady state '4-6', mainly 'fragile' - across the five separate sample collection points with a drop into a relapse '2' at the time of sample (C). The control remained in excellent health '10' throughout the timeline. The term 'recovery' is used in this investigation to define the time points following the relapse where the individual returns to the state of health they experienced prior to the relapse. It is also used as the term, for convenience, to define the time point(s) prior to the relapse event.

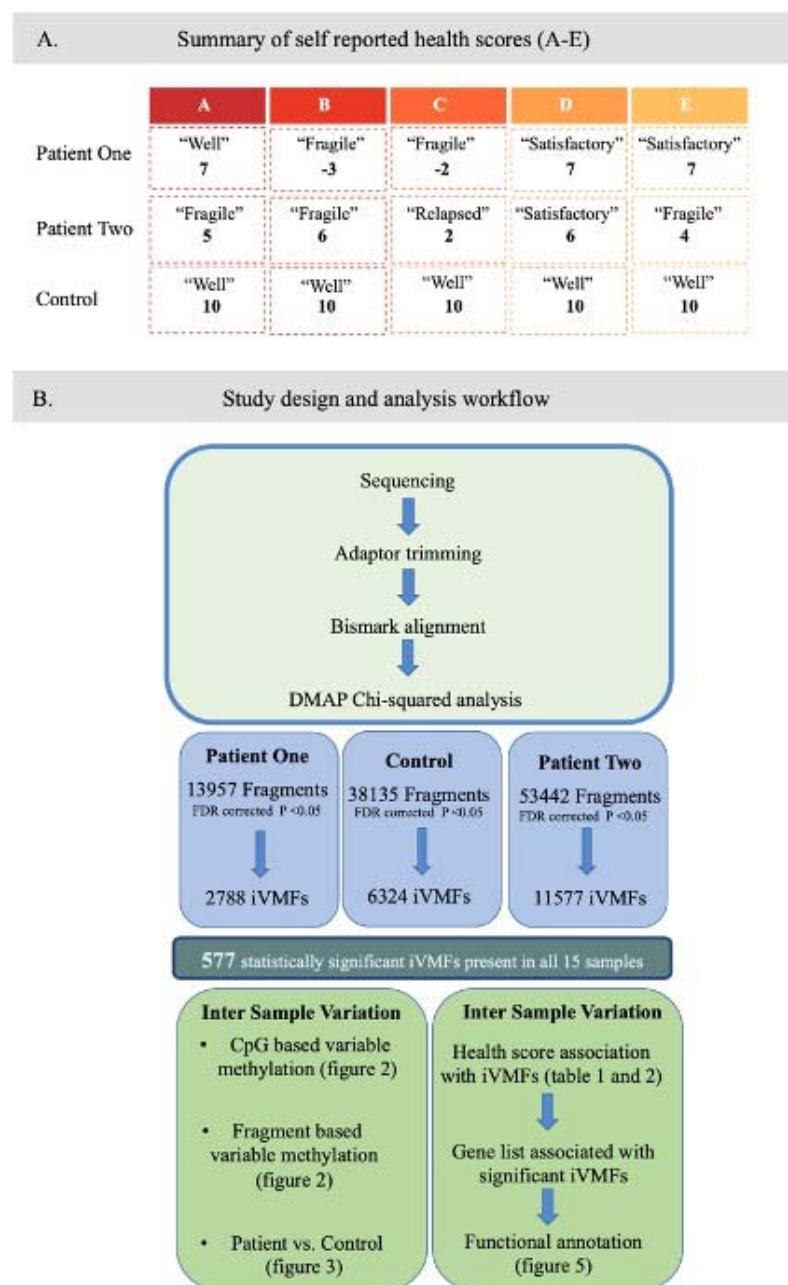


Figure 1. Study design. A. Summary of self-reported health status of study subjects through a relapse and relative recovery cycle. The timeline of the health status of two patients over an 11 month period that spanned a relapse in each patient is shown along with the matched control. Self-reported information from the patients that they supplied on the day of blood donation indicated their health status (between -3 and 10). **B. The study analysis for longitudinal analysis.** Following RRBS, adaptor trimming and alignment to human reference genome hg19 using Bismark, the samples were analysed. Initial estimates of variation utilised genome wide CpG methylation data before the samples were analysed utilising the DMAP platform where a Chi squared analysis was used to identify methylation variation. The fragment methylation was also used to estimate variation but at key genomic locations in addition to comparisons between the patients and control. Continued analysis utilised the 577 statistically significant variable fragments identified across all 15 samples (FDR corrected $P < 0.05$) where correlation was calculated with health scores and 'relapse' and 'recovery' events to identify intra-individual variably methylated fragments (iVMFs). Genes associated with the individuals iVMFs were determined and functional categories investigated.

The data from Reduced Representation Bisulphite Sequencing (RRBS) of each of the 5 samples from the three subjects were analysed using the DMAP platform with intra-individual variably methylated fragments (iVMFs) identified for each patient and the control (Figure 1B) -named ME-iVMFs. Fragments suitable for analysis had 10 or more reads of at least 2 CpG sites. A total of 13954, 53442 and 38135 met these criteria for P1, P2 and the C respectively. The statistical parameters applied to the analysis are described in methods. For the control individual, there were a total of 6324 fragments that met the significance threshold FDR <0.05. The qualifying fragments had a median size of 80bp and an average size of ~84bp. These fragments contained 52791 CpG sites, with an average of 8.4 CpGs in each fragment. For P1, there were a total of 2788 statistically significant variably methylated fragments with a total of 22550 CpG sites and an average of 8.1 per fragment (FDR corrected P <0.05). These fragments themselves had a mean length of ~78bp with a median length of ~75bp. For P2, 11577 fragments with a total of 87734 CpGs and an average of 7.6 CpGs per fragment met the same significance threshold, and had a median fragment size of 75bp, and an average fragment size ~80bp.

Dynamic analysis of DNA methylation variation

Initially the variation in the methylomes across the time points within each individual was investigated by performing a comparison of the variability at each CpG site. This was calculated from the number of statistically significantly differently methylated CpG sites (P<0.05, methylation difference >15% compared with the other time points) that were unique to each time point. The percentage of unique DNA methylation variations within each sample was derived from this number compared with the total number of CpG sites analysed for that individual. For example, in the healthy control at time point A 1276 statistically significantly differentially methylated CpGs were identified from a total of 119931 CpGs compared with the other four time points giving a unique variation of 0.12% (see Figure 2A). This was a consistent pattern, with each time point of the control having a similarly low level of unique differential methylation at CpG sites (ranging from 0.12-0.15%). By contrast the patients showed about a 20-fold greater variation -P1, at time point A compared to their other four time points had 2.06% of sites differentially methylated uniquely, and P2 at time point A had a similar level of variation of 2.67%). The unique differential methylation at the five time points ranged from 2.06-3.78% in P1 and 1.91-2.67% in P2.

Further predictions of variability based on the DMAP-produced fragment methylation data on further analysis showed that the patients were again more variable than the control, though not as distinctly obvious as from the individual CpG site methylation comparisons described above. The number of statistically significant variably fragments for each individual was divided by the total number of fragments assessed for that individual to produce a variability score. For C, P1, and P2, respectively the overall variability scores were 0.17, 0.20, and 0.22. In order to assess the relative variability of the two patients and the control across functionally important regions of the genome their individual data were extracted across regions of interest, such as within gene bodies, Transcriptional Start Sites (TSS) upstream regions of 10,000bp (see Figure 2B), and relative CpG island regions of <500bp with more than 55% GC content position) (Figure 2C).

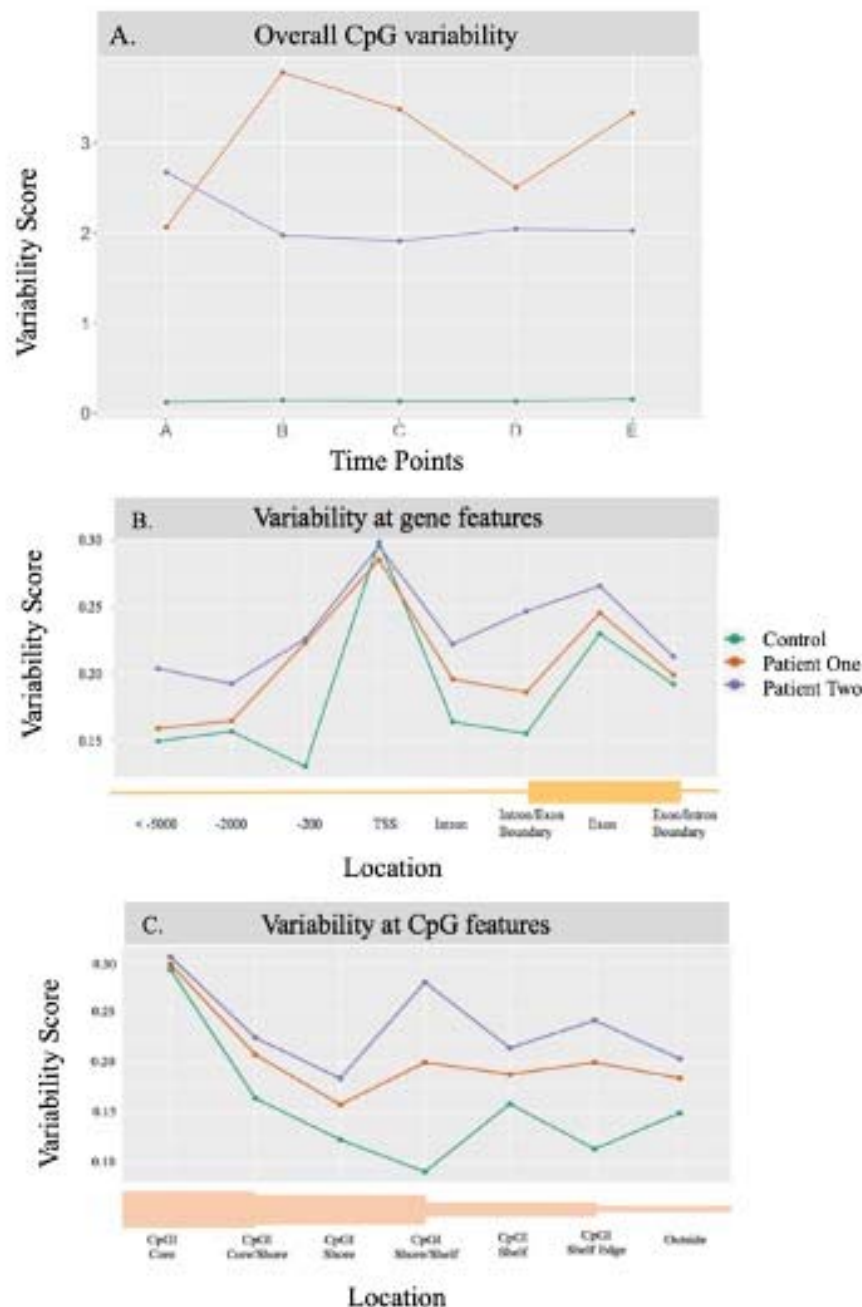


Figure 2 Line plots showing the variability calculated for each individual. **A** summary of the statistically significant unique differentially methylated CpG sites at each time point (A)-(E) for the two patients and the healthy control. The level of methylation variation was calculated utilising individual CpG methylation scores, with the significant differentially methylated CpGs found at that time point compared with the other time points, divided by the overall number of CpGs analysed for that individual, to give the percentage that were uniquely differentially methylated in each sample from each time points (A) to (E). **B**. Line plots showing the variability scores calculated for each individual across the gene features indicated on the x axis. Variability score was calculated by dividing the number of statistically significant ($q < 0.05$) variable fragments by the total number of fragments analysed at that feature **C**. Line plots showing the variability scores calculated for each individual across the features indicated on the x axis related to CpG Islands. CpG islands were defined as regions less than 500bp with more than 55% GC content, CpG shores are defined as regions 2Kb from the island with shelves 4Kb away, the boundaries between these features are included (CpGI core/shore, shore/shelf and shelf edge). Variability score was calculated by dividing the number of statistically significant ($q < 0.05$) variable fragments by the total number of fragments analysed at that feature.

197

198

199 As seen in Figure 2B the healthy control showed a lower level of variability compared to both the patients at
200 almost every gene related site analysed apart from the TSS where all three individuals in the analysis show
201 similar levels of variability. P2 with the more fragile health showed a higher level of variability compared to
202 both the control and P1.

203

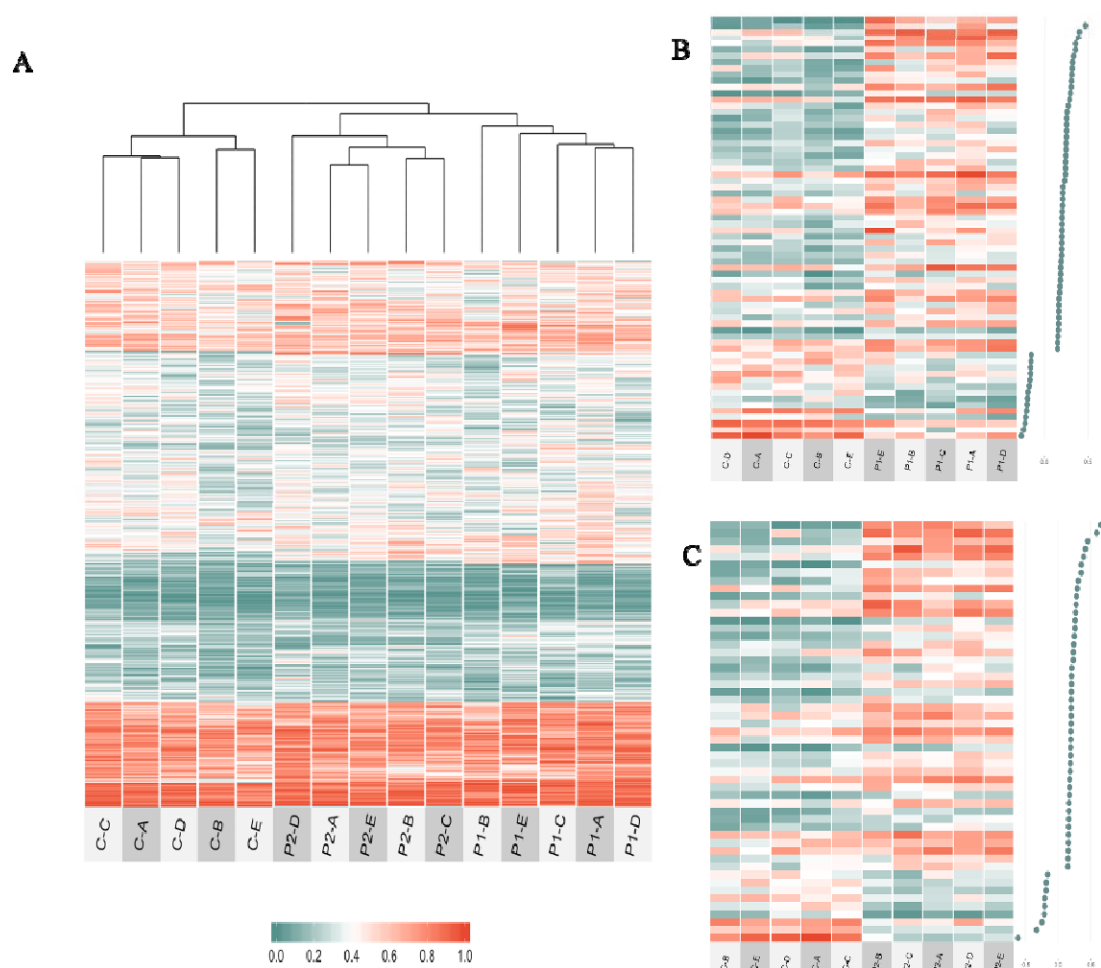
204 As seen in Figure 2C the variability of the three individuals across the CpG island related features reflected the
205 same pattern as in Figure 2B at the gene related regions. The control showed the lower variability score
206 compared to both patients. As before, P2, who had the more debilitating ongoing ME/CFS, showed consistently
207 higher levels of variability across the features analysed when compared to P1.

208

209

210

211 **Figure 3. Heatmap showing the individual methylation variation at ME-iVMFs of interest. A.** Shows the



212 methylation percent variation across all 577 statistically significant ME-iVMFs detected across all 15 samples.
 213 **B.** shows the methylation variation in 68 ME-iVMFs where the mean methylation difference between the P1
 214 and C groups is greater than 15% and **C.** shows the methylation variation in 53 ME-iVMFs where the mean
 215 methylation difference between the P2 and C groups is greater than 15%. The dot plots associated with **B** and **C**
 216 on the right of the figure show the degree of differential methylation. The scale below Figure 3A shows the
 217 corresponding colour associated with the methylation scores.

218 **Common ME-iVMFs methylation patterns in patients**

219
 220 Continued analysis of ME-iVMFs included only those fragments that were present in all 15 samples of the three
 221 individuals, which resulted in a total of 577 common fragments. Figure 3 is a heatmap that shows the
 222 methylation variation across these fragments. It identified the hierarchical clustering of the individual samples
 223 based on methylation percent values at each segment, with the associated dendrogram in Figure 3A clearly
 224 showing the relationships between the 15 individual samples. The variation within each patient and within the
 225 control sample (A to E samples) is lower than the variation among the samples (P1,P2, C) , since the heatmap
 226 and associated dendrogram clearly grouped the 15 samples into three separate groups that relate to each

individual (P1, P2,C) . However, it also shows visually the variation within each individual with the 5 different time points clearly showing differences in methylation at a number of fragments.

Initial investigations of these 577 common fragments involved comparing the two patients individually with the matched control in a differential methylation analysis. If the mean methylation difference was greater than 15% between the patient and control methylation scores at these fragments, they were then investigated further. Figure 3 B shows the differentially methylated fragments across 68 such selected fragments for P1 vs. C and Figure 3 C shows the 53 selected for P2 vs. C.

To further analyse the data for the differentially methylated fragments in the patients compared with the control that fell within gene bodies, STRING.org pathway enrichment analysis was performed. Of the 26 genes that contained differentially methylated fragments in P1 compared with C, two pathways were identified; Nicotine addiction and Morphine addiction due to the presence of genes GNAS, CACNA1A and GABRD. The 23 genes that contained differentially methylated fragments in P2 compared with C showed two pathways with the protein domains; Transforming growth factor-beta (TGF-beta) family and Immunoglobulin C-2 Type, that were identified from the genes; IGSF9B, OPCML, GDF7, CERS1 and LINGO3.

Identifying methylation pattern associated with the relapse condition

In order to find fragments with changes relevant to the relapse in the patients, the data from the overall 577 ME-iVMFs were correlated with the patients self-reported health scores. A Pearson's correlation coefficient was calculated using the association between the methylation percent at each fragment to the individuals self-reported health score (as seen in Figure 1A). A minimum Pearson's correlation coefficient of 0.9 was set. In order to further filter the fragments and select those that reflect the greatest changes in methylation between the patients self-assessed 'relapse' and better health 'recovery' conditions, a methylation difference was calculated based on the average mean methylation percentages of the relapse and recovery samples, for example, for P1 time points "B" and "C" were 'relapse' and "A", "D" and "E" were classified as 'recovery'. A minimum differential methylation of +/- 15% was set. This correlation analysis for P1 identified 17 fragments (Table 1). There were a total of 14 fragments identified using this method for P2 (Table 2).

Table 1 Fragments associated with the relapse condition for P1. Fragments with a Pearson's correlation coefficient of at least 0.9 and a mean methylation difference between the "relapse" and "recovery" time points

of at least 15% are shown. The table describes the location of each fragment, and the gene id if appropriate. It lists any overlapping regulatory elements (promoters/enhancers) recorded in Genehancer and the gene id associated with clusters of regulatory interaction determined using UCSC genome browser. Additionally, the table shows the methylation percent recorded for each time point (P1-A to P1-E) at each fragment with the relapse time points italicised.

Frag	Chr	Start	End	Position	GeneID	Genehancer	Regulatory Interactions	P1-A	P1-B	P1-C	P1-D	P1-E
1	19	33885306	33885381	On Intron	PEPD	GH09J033388	CEBPG:PEPD	78	45	44	71	64
2	X	150565438	150565527	On Intron	VMA21	GH0XJ151395	VMA21	50	33	31	45	49
3	19	55464080	55464189	On Intron	NLRP7	GH19J054952	NLRP2	78	55	48	81	75
4	7	5741705	5741780	On Intron	RNF216	GH07J005687	ACTB:CCZ1:RNF216:USP42	86	60	74	84	88
5	X	135579269	135579310	On Intron	HTATSF1	-	-	39	28	22	46	40
6	X	152908188	152908279	On Intron	DUSP9	-	DUSP9	46	18	30	37	45
7	17	45925149	45925204	On Exon	SP6	GH17J047846	SP2:CDK5RAP3:OSBPL7:SCR N2	46	28	29	46	44
8	X	23761294	23761378	On Exon	ACOT9	GH0XJ023741	ACOT9	37	18	18	34	37
9	8	145003618	145003684	On Exon	PLEC	GH08J143914	ZC3H3:EEF1D:PLEC	70	34	37	58	59
10	X	149106531	149106576	On Exon	CXorf40B	GH0XJ149937	LINC00B94:CXorf40B	48	28	36	54	58
11	22	42316243	42316306	Intergenic	-	GH22J041918	WBP2NL:CYP2D8P:CENPM: CYP5D6:TNFRSF13C	44	30	27	44	44
12	1	17199256	17199369	Intergenic	-	-	NECAP2:CROCC	55	35	44	59	57
13	2	232348597	232348713	Intergenic	-	-	NMUR1:NCL	57	38	34	69	62
14	13	114918456	114918525	Intergenic	-	-	CDC16:UPF31:RASA3	87	65	62	96	84
15	2	26521360	26521433	Intergenic	-	GH02J026298	HADHB:HADHA:ADGRF3	53	38	35	60	53
16	3	10334731	10334778	Intergenic	-	GH03J010291	GHRLOS:GHRL	33	13	22	40	41
17	15	22095431	22095475	Intergenic	-	-	-	51	40	34	51	56

P1-A to P1-E in Table 1 represents percent methylation data from the samples taken from P1 at each of the time points (A) to (E). As can be seen at time points (B) and (C) during which there was a self-reported severe relapse (Figure 1A) there was a much lower methylation rate than in the samples from the ‘recovery’ times (A), (D) and (E).

Table 2. Fragments associated with the relapse condition for P2. Fragments with a Pearson’s correlation coefficient of at least 0.9 and a mean methylation difference between the “relapse” and “recovery” time points of at least 15% are shown. The table describes the location of each fragment and the gene id if appropriate. It lists any overlapping regulatory elements (promoters/enhancers) recorded in Genehancer and the gene id associated with clusters of regulatory interaction determined using UCSC genome browser. Additionally, the

275 table shows the methylation percent recorded for each time point (P2-A – P2-E) at each fragment with the
 276 relapse time point italicised. Fragments hypermethylated in the relapse condition are shown in bold.
 277

Frag	Chr	Start	End	Location	GeneID	Genehancer	Regulatory Interactions	P2-A	P2-B	P2-C	P2-D	P2-E
1	19	4543716	4543762	On Intron	SEMA6B	GH19J004539	YJU2:PLIN5:SEMA6B:LRG1	80	81	<i>47</i>	91	67
2	X	15353393	15353501	On Intron	PIGA	GH0XJ015333	ZRSR2:PIGA	40	41	<i>18</i>	39	33
3	22	17640812	17640923	On Intron	CECR5	GH22J017157	HDHD5	69	82	<i>60</i>	80	71
4	14	105936238	105936292	On Exon	MTA1	GH14J105464	IGHGP:CDCA4:CRIP2:MTA1:ENSG00000257270	84	80	<i>63</i>	87	76
5	22	18027985	18028072	On Exon	CECR2	-	-	49	46	<i>68</i>	38	51
6	X	102565776	102565848	Intron – Exon Boundary	BEX2	GH0XJ103310	BEX2	42	45	<i>26</i>	43	41
7	1	155098923	155098964	Intergenic	-	GH01J155123	DAP3:CLK2:DPM3:GBAP1:THBS3:EFNA1	54	67	<i>35</i>	72	60
8	3	10334731	10334778	Intergenic	-	GH03J010291	GHRLOS:GHR	49	56	<i>32</i>	61	41
9	9	38687682	38687760	Intergenic	-	-	-	61	62	<i>39</i>	65	42
10	7	100882140	100882220	Intergenic	-	GH07J101231	FIS1:CLDN15	80	80	<i>65</i>	90	75
11	2	219233608	219233704	Intergenic	-	GH02J218366	AAMP:SCL11A1:TM6IM1:CATIP	49	55	<i>38</i>	62	52
12	6	170403979	170404085	Intergenic	-	-	WDR27	62	66	<i>81</i>	63	70
13	X	129299533	129299622	Intergenic	-	GH0XJ130164	ELF4:AIMF1:ZNF280C	42	36	<i>61</i>	31	49
14	X	135579192	135579268	Intergenic	-	-	-	28	27	<i>61</i>	30	44

278
 279 Whereas the ME-iVMFs that associated with the relapse condition of P1 were all hypomethylated (see Table 1),
 280 for P2, while the majority of the fragments 1-11 were also hypomethylated, three (12-14) by contrast were
 281 hypermethylated (shown in italics in Table 2) in the relapse condition (P2-C) compared to the recovery
 282 conditions.

283
 284 The fragments for both patients show a number of interactions with genomic elements including direct overlaps
 285 with gene bodies as well as regulatory elements as recorded in GeneHancer, and USCS genome browser
 286 recorded clusters of regulatory interactions between regulatory elements and gene bodies. As these fragments
 287 show clear changes in the methylation state of the individual across their relapse and recovery states it has
 288 important implications on the regulatory behaviour of a number of associated genes.

289

Relapse associated methylation signature exhibits striking variation compared to control

Figure 4 gives examples of the top 6 fragments across the five time points for both patients that had the greatest level of differential methylation between their relapse and recovery states. Figure 4 clearly shows that there are clear changes in methylation within the two patients at the relapse condition. In Figure 4A the relapse condition is shown for P1 at time points (B) & (C) compared to recovery time points (A), (D) & (E)), and in Figure 4B for P2 with relapse at time point (C) and (A),(B), (D) & (E) for recovery. Due to their genomic location, these fragments have important functional implications, for example from P1 shown in Figure 4A is a fragment that is located within the first intron of NLRP7 gene. It also overlaps with an enhancer (GH19J054952) and directly overlaps a region of regulatory interaction for NLPR2 in addition to being located within a region of Dnase hypersensitivity. As previously mentioned all these fragments were hypomethylated in the relapse state of P1, illustrated in the examples shown in Figure 4A, indicating that the corresponding regulatory features likely have a downstream up-regulation on associated genes.

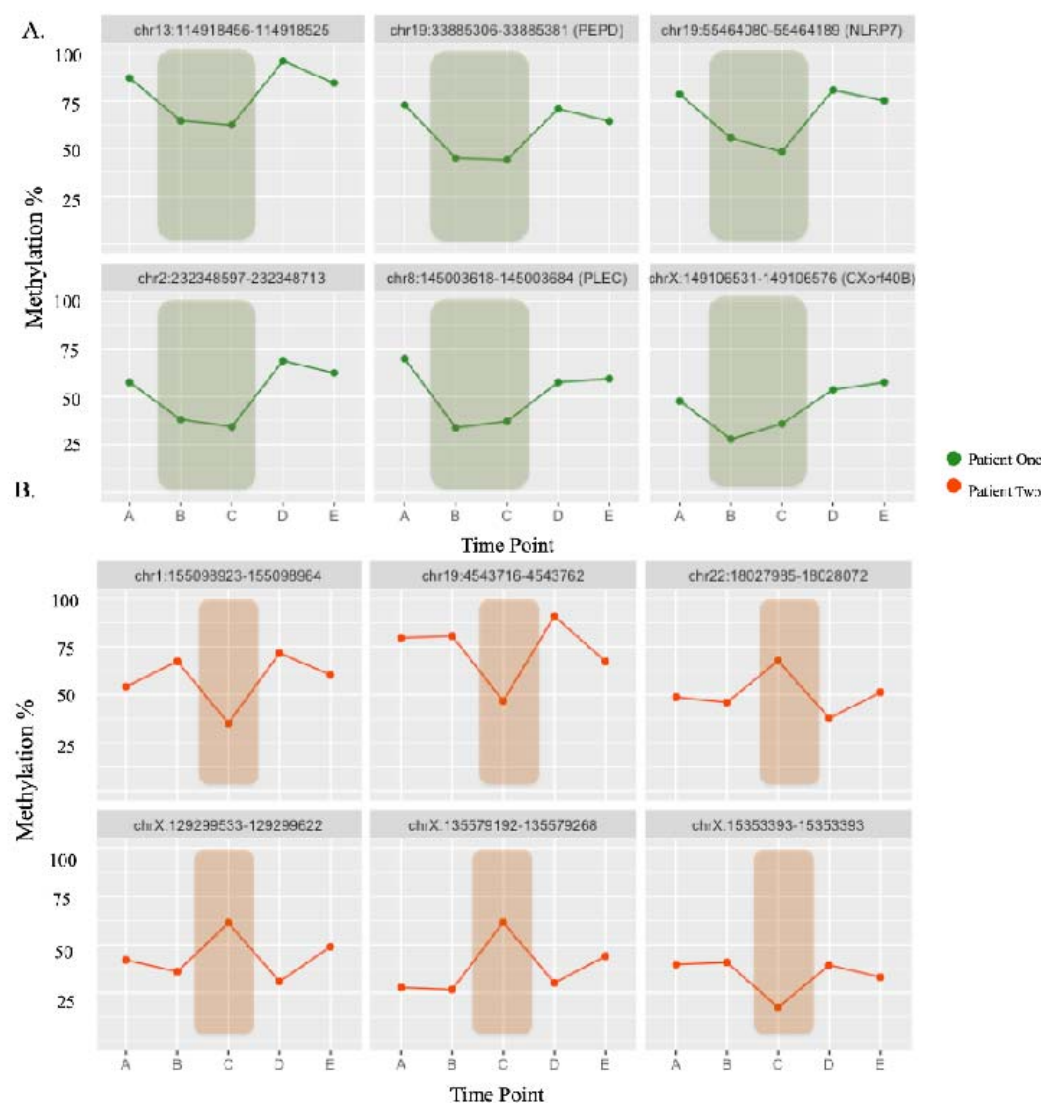


Figure 4. Dynamic DNA methylation variation correlated to self-reported patient health status.

Methylation percentages are shown across all five time points (A)-(E) for each fragment of interest -fragment co-ordinates are shown above each block and a gene id is shown in brackets if the fragment directly overlaps a gene feature (intron/exon). **A.** P1-green highlight indicates a period of relapse, **B.** P2-orange highlight indicates period of relapse captured in one blood sampling.

Both hypomethylation and hypermethylation is shown in Figure 4B with the examples from P2. P2 also has a number of fragments of regulatory importance such as a fragment located on chr1:155098923-155098964 that is located within an archived promoter region (GH01J155123) and has 37 target genes. Additionally, this fragment overlaps with a number of clustered interactions between Genehancer regulatory elements and genes for; DAP3, CLK2, DMP3, GBAP1, THBS3, EFNA1. Another fragment of interest shown in Figure 4B is located within the 17th (last) exon of SEMA6B, as it encodes a protein that may be involved in both peripheral

and central nervous system development. Additionally, this fragment overlaps with a Genehancer archived promoter region (GH19J004539). It overlaps with a DNase hypersensitivity cluster and four clustered interactions of Genehancer regulatory elements and genes (YJU2, PLIN5, SEMA6B and LRG1).

In order to investigate the potential impact of the methylation variation on the patient's molecular activity during relapse recovery cycles a total gene list was built that associated with the regulatory elements found to be overlapping with the fragments described in Tables 1 and 2. There were a total of 157 genes associated with the 17 ME-iVMFs identified on relapse in P1, and 127 genes associated with the 14 ME-iVMFs identified in P2 (see supplementary excel file 'Supplementary.xlsx' sheets 'Genes associated with P1' and 'Genes associated with P2').

Simulated relapses for the control subject identified fewer variable methylated genes than the patients

To determine whether the ME-iVMFs and associated genes were actually due to the ME/CFS relapse and recovery conditions and not simply a result of random chance and random methylation variation, the control sample was also analysed in two separate determinations as though the healthy control *also* had the relapse health scores of the two patients respectively at the appropriate time points. Thus each patient relapse health scores were assigned 'artificially' to the relevant control time points in separate analyses to calculate the correlation and determine differential methylation from the control data between these simulated 'relapse' and 'recovery' states. From these analyses, 11 fragments with only 39 genes associated with them (see supplementary excel file 'Supplementary.xlsx' sheet 'Control_Filt Correl with P1' and, see supplementary excel file 'Supplementary.xlsx' sheet 'Genes_assoc with C-P1 condition') met the filtering requirements with P1's health scores, and 14 fragments with 53 genes with P2's health scores (see supplementary excel file 'Supplementary.xlsx' sheet 'Control_Filt Correl with P2' and, see supplementary excel file 'Supplementary.xlsx' sheet 'Genes_assoc with C-P2 condition'). This compares with 157 genes in the data comparison and health states comparison for P1 and 127 genes in the analysis for P2. To determine whether the relationship between the number of associated genes identified per fragment was significantly higher in the patients relapse times, the number of genes identified from each fragment from these 'simulated' control analyses and the 'real' patient analyses were subjected to an unpaired t-test. The number of genes per fragment were significantly higher for the patient group compared to the two control simulations, with a p -value = 0.0053

(see Figure 5A). This implies that while a proportion of the methylation changes seen in the patients may be due to random methylation variation, most of the identified changes during relapse in the patients are due to their physiological relapsed state and are associated with important regulatory regions of the genome linked to ME/CFS disease presentation.

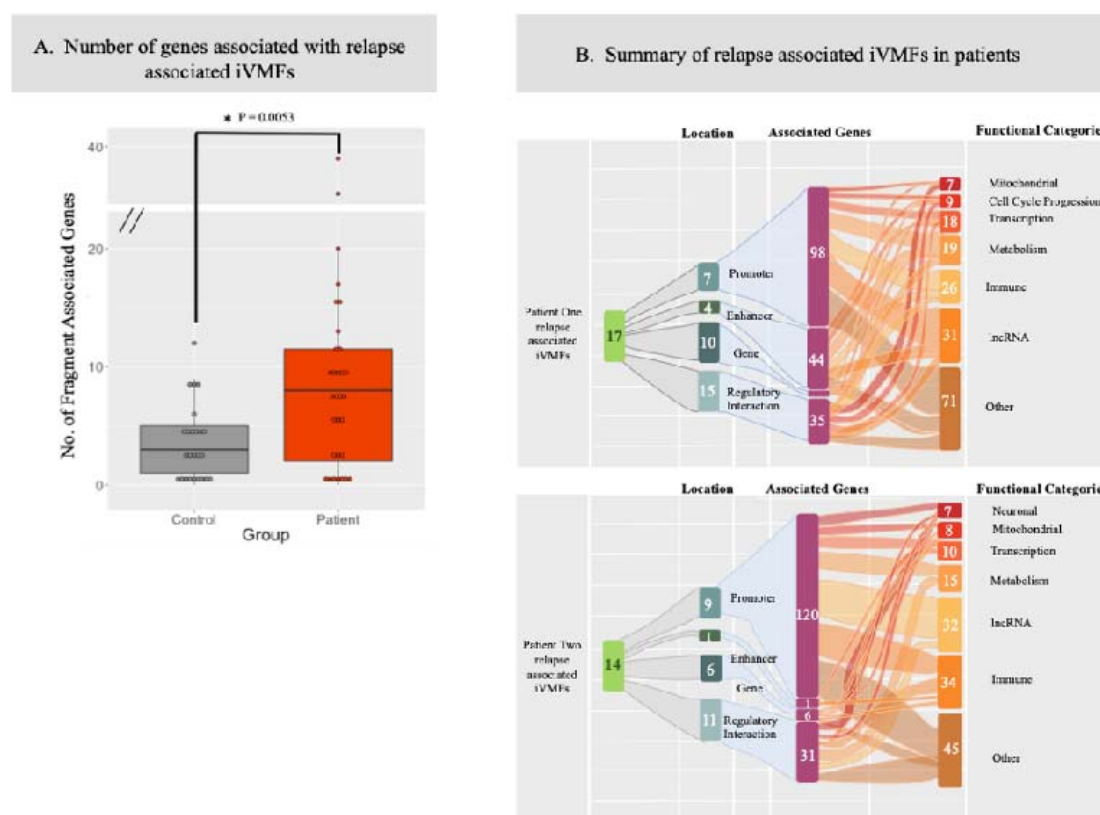


Figure 5. Investigation of the genes associated with relapse associated ME-iVMFs. **A** Box plots showing the number of genes linked with each statistically significant variably methylated fragment associated with a ‘simulated’ relapse event for the control in grey (in two analyses using each of the patient health relapse time points) and the ‘real relapse events’ of the two patients in orange. Each point represents a fragment with the number of associated genes shown on the y-axis. The mean number of genes associated with the identified fragments for the patient and control groups is shown with the line. An unpaired t-test resulted in a significance value of $p = 0.0053$. **B**. Sankey plot showing relationship between the variably methylated fragments identified in each patient associated with a relapse event and the biological functions they associate with through various regulatory genomic elements of relevant genes. From the statistically significant variably methylated fragments identified for each individual the location was determined and relevant regulatory interactions were recorded from UCSC genome browser. A gene list was compiled of genes associated with these regulatory interactions and the functional annotations were utilised to place them into categories. Some genes fell into multiple categories with others having no known function. Tables showing the full gene list, function and functional category is included in supplementary excel file ‘Supplementary.xlsx’ sheet ‘Genes_associated with P1’ and ‘Genes associated with P2’.

The potential functional associations of each gene that were linked to the variably methylated fragments associated with a 'relapse' event were determined. The gene functions identified by this analysis associated with patient symptom fluctuations indicate a change primarily in 'immune response' for both patients. Additional functions associated with the genes involved metabolism and transcription for both patients. P1 also had a number of genes involved in cell cycle progression while P2 had a larger number of neuronal related genes. Gene annotations for each gene are listed in the supplementary Excel file 'Supplementary.xlsx' sheets 'Genes associated with P1' and 'Genes associated with P2'. Of the immune related genes identified, a number were associated with activities implying increases in the inflammatory response in individuals, with specific functions linked to NF-kappa B activity, wound healing, cytokine release, and angiogenesis observed multiple times. This suggests that during a period of relapse the patient's immune systems are in an enhanced inflammatory state compared to their relative 'recovery' periods.

Discussion

Previous studies (11-15), including our own study that described the first RRBS methylome of ME patients (16) have established ME/CFS patients display an altered DNA methylome in comparison to matched controls. This current study is the first of its kind to analyse the DNA methylome of ME/CFS in individual patients across a longitudinal timeline to investigate a change in health status. Utilising the principles of precision medicine it has identified two key features: (i) the number of variably methylated sites and fragments of the genome are much greater in the two ME/CFS patients than in the control at each time point of the longitudinal study and, (ii) the severity of ME/CFS symptoms during a relapse is associated with methylation variation at key genomic features. The variable methylated DNA fragments enabled us to identify statistically important features specifically associated with a significant 'relapse' in the health of the two patients, compared with their prior health and their recovery after the relapse. The genomic features implicated regulatory changes affecting primarily immune functions with associated inflammation, but also metabolic, neurological and mitochondrial functions in patients as they experience symptom fluctuations along the course of their disease.

Benefits of DNA methylation for a precision investigation of ME/CFS

The individuals selected to participate in this study were within a similar weight, age range with the same gender, ethnicity and lifestyles. This was done to prevent any potential confounding factors, since DNA

methylation is a dynamic epigenetic mark known to vary due to environmental factors. The specific criteria for patient selection utilised aimed to ensure that the variation in methylation would be primarily due to fluctuations in ME/CFS symptom severities.

While previous epigenetic studies have utilised primarily array based methods, prior studies involving the same method of RRBS and analysis platforms described here have been performed successfully previously (21,22). Indeed, our recent study (16) with this method with ME/CFS patients gave methylation changes that significantly overlapped with the other similar studies of this disease that used the array technology (11-15). Utilising RRBS technology a large number of changes were identified that differentiated ME/CFS patients from controls. The use of RRBS here has followed extensive in house development and experience with the platform used (21-23). The advantage of using RRBS is that it identifies changes not captured by the array-based studies as it is not limited to the set number of sites in the array, allowing wider coverage of the whole genome.

DNA methylation is an excellent method to investigate physiological changes as it is reflective of transcriptional changes linked to the disease state, and so is very appropriate to study the relapse recovery cycle of ME/CFS). DNA methylation is a versatile method to investigate an individual's physiology (20,27). Small observed changes often reflect much larger changes occurring in a subpopulation of cells that are obscured by the broader range and number of cells from which the DNA is taken. Notably, previous research has indicated that even small measured methylation changes can have large impacts on the associated expression levels of a gene (25). For example, a recent investigation found that even a small change in methylation percent of (1%) was associated with a two-fold change in expression of insulin like growth factor -2 (IGF2) (26). A key study relevant to our 'relapse and recovery' in ME/CFS patients showed DNA methylation to be changed in at risk individuals *before* their transition to diabetes (20). This application of personalised medicine allows DNA methylation variation to be utilised not only to distinguish patients from healthy controls, but also to provide a more specific pathophysiological understanding of an individual patient's disease trajectory.

Inter-individual differences indicate increased epigenetic variation linked to disease severity

As we develop a deeper understanding of the onset of ME/CFS, it is becoming clear that there is an underlying genetic predisposition in combination with an environmental trigger to precipitate an altered homeostatic state

or compromised health ‘baseline’ in patients (27). Once the disease progresses past the initial acute stage ~75% of patients can transition to a chronic state but the partial recovery is interspersed with frequent periods of relapse followed by relative recovery to the initial compromised health state again. This ‘new normal’ chronic state of ME/CFS for patients may leave them more vulnerable to even minor changes in their environment that would not affect a healthy person. For ME/CFS patients in their altered homeostatic state a dramatic change in physiological state can easily be precipitated.

Initial analysis of the DNA methylation of the genome-wide CpGs of the patients and the healthy control in this study supported the idea that ME/CFS patients are more vulnerable to environmental changes. Month to month, the unique variability in methylation in the healthy control, who had stable excellent health throughout the longitudinal study, was low <1 in 500 sites, but both of the patients had a much higher level of unique variability at each sampling time point at between 1 in 20-50 sites. A similar estimate of variability was performed utilising the DMAP fragment methylation each containing multiple sites, and produced similar but less dramatic trends with the patients having 0.20 and 0.22 variability scores compared with 0.17 for the healthy control (based on the number of statistically significant variable fragments divided by the overall number of fragments in each individual). A key determinant however, is not the extent of variation but the variation at regions of functional importance across the genome, such as in proximity to CpG islands (often associated with regulatory regions), and upstream of and within gene bodies. This investigation found that the patients were consistently more variable than the control at all regions investigated (Figure 2) with the exception of the Transcriptional Start Site (TSS) where both patients and control had similar levels of variability. Of importance to note is that, while both patients were much more variable than the control, P2, in a more compromised state of health throughout the longitudinal timeline, was consistently the more variable of the two patients. The results from this study indicate that not only are patients more epigenetically variable than a healthy control, but also illness severity may be positively associated with methylation variation.

Intra-individual variation identifies regulatory regions

An individual is often their own best control for personalised medical applications, especially in studies like this where there is fluctuating health during a longitudinal disease course. Individuals have fluctuating baseline

DNA methylation, so important changes occurring within an individual could be obscured when compared to a control (28). In this analysis however, when the fifteen samples were clustered based on the methylation scores of the 577 significant fragments identified in all samples (Figure 3), while intra-individual sample variation was indeed revealed, the heatmap and associated dendrogram produced by hierarchical clustering showed that the inter-individual variation clearly differentiated the three individuals and was greater than this intra-individual variation.

For this reason, while a healthy control was included in the analysis, the ‘relapse recovery’ study focused primarily on variation within each individual patient and within the control as three separate individuals, utilising each as their own ‘control’ along a longitudinal time scale (for example Figure 4). Variably methylated fragments were identified in both the patients that strongly associated with the individuals self-reported health scores ($r > 0.9$) with a distinct methylation percentage difference between the ‘relapse’ and ‘recovery’ conditions ($\pm 15\%$). These thresholds enabled us to capture the more relevant changes occurring in the DNA methylation as a result of the relapse condition as discussed above since even small changes in methylation are often indicative of larger transcriptomic changes. The control was also analysed in the same manner ‘simulating’ a relapse by analysing samples B & C (as though they were a relapse as experienced by P1), and sample C (as in P2). This determined how many variably methylated fragments are likely to associate, by chance alone independent of disease, when the patient health scores are arbitrarily assigned to the control. A number of variably methylated fragments were identified and were further investigated to identify any functional associations. However, it was clear that the downstream gene associations were much lower when compared to the two patients during relapse (as shown in Figure 5).

From the variably methylated fragments identified in the patients a large number of downstream genes were associated through either direct physical overlap with the variable fragment, association with a promoter or enhancer, or within a region of regulatory interaction as recorded on UCSC genome browser. They were functionally relevant to physiological changes occurring in the patients as they experience fluctuations in health in a ‘relapse’ and ‘recovery’ cycle. The large majority of the intra-individual variable methylated fragments (ME-iVMFs) were hypomethylated in the relapse condition compared to the recovery condition (only three hypermethylated from P1) (Tables 1 & 2) indicating that there would be a corresponding increase in transcription in the downstream genes associated with the regulatory features. As there are such a large number

of genes associated with the ME-iVMFs identified in this investigation it suggests there are consequentially wide-ranging regulatory changes occurring in patients.

Immune and inflammatory changes implicated in relapse-recovery cycle

While there was a broad range of functional roles identified that were performed by the genes associated with the significant ME-iVMFs (see supplementary excel file ‘Supplementary.xlsx’ sheets ‘Genes associated with P1’ and ‘Genes associated with P2’) the largest category identified encompassed genes involved immune/inflammatory functions, then in metabolic pathways. As these biological systems have been implicated from previous ME/CFS research studies (6,29,30) it was not surprising that such functional categories would be highlighted as ME/CFS patients experienced fluctuations in their health.

The immune functions identified have important functional relevance to the presentation of ME/CFS. In their relapse compared to their recovery states, P2 had with 34 immune related genes affected. Of these genes, CXCR2 and CXCR1 indicated the potential activation of the interleukin-8-mediated signalling pathway. IL8 has already been observed as the gene most differentially expressed between ME/CFS and controls (7,30). Other previous studies also having observed a significantly higher level of IL-8 in severely affected ME/CFS patient group compared to both healthy controls and moderately affected ME/CFS patients (31).

P1 also showed a number of affected genes that like IL8 are known to be associated with inflammatory responses, for example, NLRP7, and genes associated with NF-kappa-B function (COMMD5, LRRC14, TONSL). P2 additionally also showed a similar relationship with a number of the immune related genes having inflammatory roles including (TICAM1 and IL17RA) which are involved in the positive regulation of cytokine production in inflammatory responses. Significantly a number of the immune related genes from P2 are associated specifically with inflammatory disorders including genes involved in the neutrophil degranulation pathway including; TMBIM1, SLC11A1, MOSPD2, CR2, CR1 and LRG1.

Among the additional genes of interest identified during relapse in the ME/CFS patients were seven mitochondrial genes in P1 that included ACOT9, which is a member of the acyl-CoA family involved in the hydrolysis of Coenzyme A. HADHA, HADHB are both involved in mitochondrial beta-oxidation of long chain fatty acids into either 3-etoacyl-CoA if NAD is present, or acetyl CoA if both NAD and coenzyme A are present

(32). In ME/CFS patients it has been hypothesised that a number of factors may be interfering with the production of coenzyme A as a result of inflammation, and with reactive oxygen species through the pyruvate dehydrogenase kinase pathway in the mitochondria (33). As the activity of mitochondrial beta-oxidation is key to cellular energy production, P1 may be showing the effects of reduced mitochondrial function in the relapse condition that reflects the severity of her relapse state.

While the majority of previous work investigating DNA methylation in ME/CFS patients has focused on the differences between the patients and healthy controls, a 2018 study classified patients into 4 subgroups based on DNA methylation patterns associated with symptom severity (17). DNA methylation from 1939 genomic sites was utilised as a signature to differentiate the four subgroups. Of these, the top differentially methylated sites had associations related to immune signalling. The subtypes of ME/CFS with the more severe symptom presentation in terms of post exertional malaise were the sites with the highest differential methylation indicating changes in metabolic and immune responses. When considered together with the outcomes of this current study, where the relapse events also highlighted regions potentially affecting the function of immune, inflammatory and metabolic activity, it reinforces the importance of fully understanding the dysfunction of these pathways, not only in patients compared to healthy controls but in individual patients along their disease course.

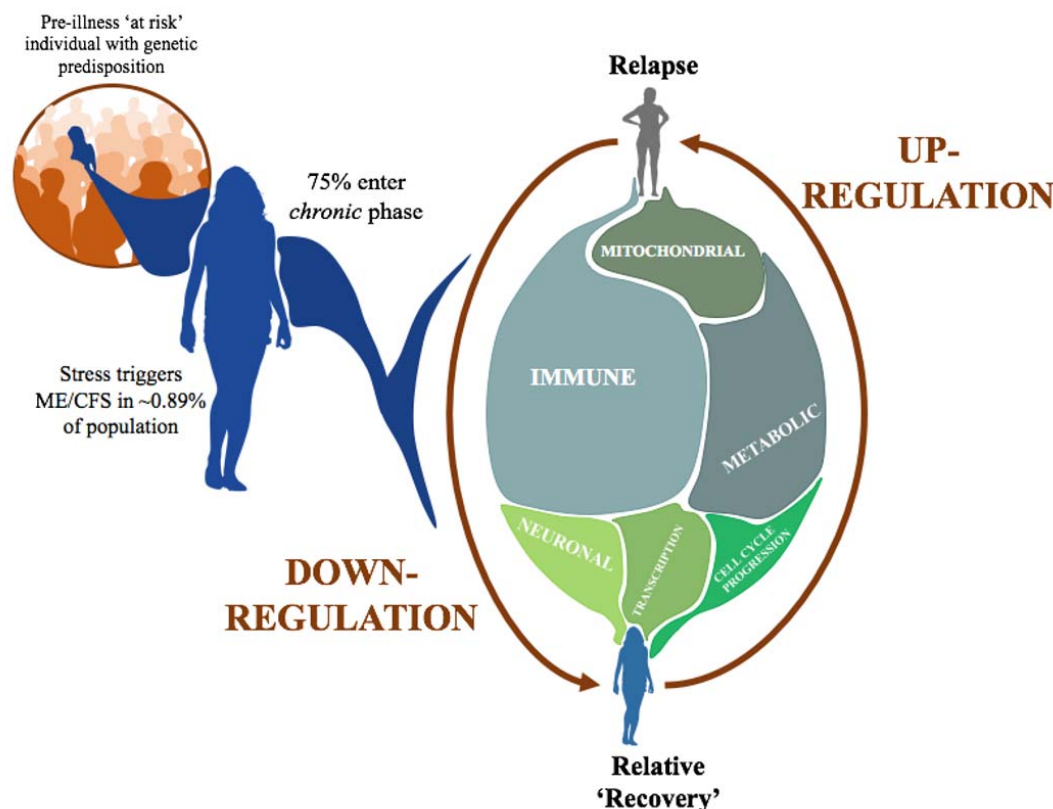


Figure 6. Summary of the longitudinal timeline of a ME/CFS patient. The initial external trigger for ME/CFS is a ‘stress event’ (for example, a viral infection- upper left in Figure 6) in a genetically susceptible person. Following progression into the chronic state patients experience frequent relapse events, which as this investigation suggests are primarily associated with the up regulation of a number of key biological systems. This investigation has shown regulatory disruptions occurring in the patients associated with their self-reported relapse events. It is worth noting that, while both patients followed here did display a similar overall pattern of disrupted functional pathways associated with their relapse events, there were notable differences. These differences would likely have been obscured if they had been part of a larger scale patient vs. control analysis. As personalised medicine is becoming more accessible, ME/CFS patients remain a patient group that will greatly benefit further from this style of investigation. Affected ME/CFS patients would be able to contribute to the overall understanding of the activity of their disease, and with individual molecular assessments be able to adopt therapeutic and behavioural management strategies that might better manage their illness and decrease the frequency of relapses.

Conclusion

This study shows the benefits of precision medicine for individual patients with a disease as physiologically complex as ME/CFS. Currently, ME/CFS patients can respond quite differently to specific medications, for example supplements like vitamin B12, and to anti-inflammatory drugs like naltrexone, and to physiological states like pregnancy, with some showing marked improvement, some marked deterioration, and some seemingly no change in their condition. By considering individual patients over the course of their ME/CFS disease we can better understand not only the similarities within the overall patient group, but also develop an in depth understanding of the fluctuations for each patient that relates to their specific pathophysiology. Variable methylation of regulatory regions associated with the relapse condition has in this study identified a number of genes with key functional roles in immune, inflammatory, metabolic and mitochondrial pathways. For a disease that has proven challenging to diagnose and characterise, with the delay in diagnosis detrimental for the affected person, this kind of analysis provides not only further evidence of serious biological dysfunction, but importantly also ongoing systematic molecular changes that inform future targets for individual treatment or symptom management as we continue to unravel and understand the complex nature of ME/CFS.

Methods

Cohort recruitment

ME/CFS patients were recruited from Dunedin, New Zealand. Diagnosis was initially made by expert clinician, Dr Rosamund Vallings, of the Howick Health and Medical Centre, Auckland, NZ using the International Consensus Criteria (34). The two patients and the healthy control. The two patients and the healthy control were NZ European females (23-28 years of age) of similar weight. Each was asked to self-report on their health status at each blood sampling indicating whether they were in a stable health period or in a more fragile or relapsed health state. Details of these assessments from each patient and control can be found in Table 1. The study conforms to the ethics approval 17/STH/188 for ME/CFS patient studies from the Southern Health and Disability Ethics Committee of New Zealand. General consultation with Ngai Tahu Research Committee of the University of Otago was carried out before the beginning of this research.

PBMC isolation

The study involved sampling of blood on 5 occasions from two patients and a healthy control over an 11-month period with the aim of catching a ‘relapse/relative recovery’ cycle of their illness. The patients filled out a brief survey detailing their current condition at the time of each blood collection. These health indicators are seen in Figure 1. Blood was collected early to midmorning and the fractions were then processed within the same day. PBMCs were isolated from the whole blood by layering on Ficoll-Paque before separating plasma from PBMCs and other cells by centrifuging at 400 x g. The PBMC layer was pelleted (100 x g) through PBS and the resulting pellet resuspended in PBS and RNA later and stored at -80° (deg C).

DNA extraction

DNA was extracted from 200µl of the PBMC fraction using the Illustra blood Genomic Prep Mini Spin Kit according to the manufacturer’s instructions. DNA was eluted into the provided EB buffer. Concentration was determined utilising the Qubit 2.0 flurometer, following the Qubit dsDNA HS Assay Kit protocol.

Generating methylation map using RRBS

RRBS libraries were prepared as previously described (16). Briefly, genomic DNA (500ng) was digested with 160U of MSP1 restriction enzyme. Following end repair and adenylation of 3’ ends, adaptors were ligated to the DNA fragments. Bisulfite conversion was performed using the specifications of the EZ DNA methylation kit. Semi-Quantitative PCR was performed on the bisulfite converted DNA in order to determine the optimal amplification cycle needed for the final large scale PCR of the final library. Following PCR amplification of the DNA it was size selected using a 6% (w/v) NuSieve Gel in order to extract the 40-220bp desired fragments for RRBS libraries and to minimize adaptor contamination. Following purification and analysis of quality using a BioAnalyzer and Qubit measures, samples were further purified using AMPure XP Bead purification.

High-throughput Sequencing

The samples were sequenced through the Otago Genomics and Bioinformatics Facility. Following sequencing the raw fastq files were checked for adaptor presence and trimmed. The data were aligned to the human genome version GRCh37/hg19 using Bismark bowtie alignment generating BAM files utilised in the differential methylation analysis.

DNA methylation and Statistical analysis

Analysis was performed with the DMAP analysis program (35,36) run on a MAC OS X computer in order to investigate regions of methylation variability within each individual across fragments 40-220bp in length. DMAP applied a Chi-squared test comparison for each individual. The fragment-based analysis approach has been well described previously (37–39). All samples collected from each individual were included in these analyses. A minimum of two CpGs in each fragment had a minimum of at least 10 sequencing hits in order for the fragment to qualify. A Chi-square distribution test was performed on the five samples taken from each individual in this longitudinal study. False discovery rate corrected P values were calculated for each fragment and only fragments that met the significance threshold of FDR <0.05 were used in the remaining analysis. The genomic features overlapping with the fragments were identified using the DMAP Geneloc function.

Differential methylation was performed on each patient compared with the control producing gene lists, i.e. differentially methylated fragments directly overlapping with exon/intron regions. These gene lists were then analysed with pathway enrichment analyses using String.org (40). A FDR P value cut-off of 0.05 was applied to select the enriched pathways.

Fragments associated with patient relapse events were identified using 577 common ME-iVMFs detected across the three individuals. A fragment was associated with the relapse condition if it was found to have at least 15% average methylation difference between the relapse and recovery states, and if the methylation scores had a Pearson's correlation coefficient of at least 0.9. In order to identify the functional associations of each variably methylated fragment associated with patient relapse events the regions covered by the ME-iVMFs of interest were investigated using the UCSC genome browser to compile a list of archived overlapping enhancers, promoters and regions of regulatory interactions. The associated genes were determined using the Genehancer database (41). The functional roles of these genes were determined using Genecards (42) which was then used for determining appropriate functional categories for downstream analyses.

List of abbreviations

CpG: 5': Cytosine-3'phosphate-5'Guanosine-3'

650

651 **iVMF:** Intra-individual variably methylated fragment

652

653 **ME/CFS:** Myalgic Encephalomyelitis/Chronic Fatigue Syndrome

654

655 **DMA:** Differential methylation analysis program

656

657 **PBMC:** Peripheral blood mononuclear cells

658

659 **RRBS:** Reduced representation bisulfite sequencing

660

661

662

663 **References**

- 664 1. Shepherd C. ME/CFS/PVFS: an exploration of the key clinical issues. ME Association; 2013. 52 p.
- 665 2. Lim EJ, Ahn YC, Jang ES, Lee SW, Lee SH, Son CG. Systematic review and meta-analysis of the
666 prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). J Transl Med. BioMed
667 Central Ltd.; 2020;18:100.
- 668 3. Valdez AR, Hancock EE, Adebayo S, Kiernicki DJ, Proskauer D, Attewell JR, et al. Estimating
669 Prevalence, Demographics, and Costs of ME/CFS Using Large Scale Medical Claims Data and Machine
670 Learning. Front Pediatr; 2018;6:412.
- 671 4. Chu L, Valencia IJ, Garvert DW, Montoya JG. Onset Patterns and Course of Myalgic
672 Encephalomyelitis/Chronic Fatigue Syndrome. Front Pediatr; 2019;7:12.
- 673 5. Naviaux RK, Naviaux JC, Li K, Bright AT, Alaynick WA, Wang L, et al. Metabolic features of chronic
674 fatigue syndrome. Proc Natl Acad Sci USA; 2016;113:E5472–80.
- 675 6. Missailidis D, Annesley SJ, Allan CY, Sanislav O, Lidbury BA, Lewis DP, et al. An isolated Complex
676 V inefficiency and dysregulated mitochondrial function in immortalized lymphocytes from ME/CFS
677 patients. Int J Mol Sci;2019:1–38.
- 678 7. Sweetman E, Ryan M, Edgar C, MacKay A, Vallings R, Tate W. Changes in the transcriptome of
679 circulating immune cells of a New Zealand cohort with myalgic encephalomyelitis/chronic fatigue

680 syndrome. *Int J Immunopathol Pharmacol*; 2019;33:205873841882040.

681 8. Mackay A, Tate WP. A compromised paraventricular nucleus within a dysfunctional hypothalamus: A
682 novel neuroinflammatory paradigm for ME/CFS. *Int J Immunopathol Pharmacol*; 2018 1;32.

683 9. Montoya JG, Holmes TH, Anderson JN, Maecker HT, Rosenberg-Hasson Y, Valencia IJ, et al.
684 Cytokine signature associated with disease severity in chronic fatigue syndrome patients.. *Proc Natl*
685 *Acad Sci USA* ; 2017;114:E7150–8.

686 10. Kerr JR, Burke B, Petty R, Gough J, Fear D, Matthey DL, et al. Seven genomic subtypes of chronic
687 fatigue syndrome/myalgic encephalomyelitis: A detailed analysis of gene networks and clinical
688 phenotypes. *J Clin Pathol*; 2008;61:730–9.

689 11. de Vega WC, Vernon SD, McGowan PO. DNA methylation modifications associated with chronic
690 fatigue syndrome. *PLoS One*; 2014;9:e104757.

691 12. de Vega WC, Herrera S, Vernon SD, McGowan PO. Epigenetic modifications and glucocorticoid
692 sensitivity in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *BMC Med Genomics*;
693 2017;10:11.

694 13. Herrera S, de Vega WC, Ashbrook D, Vernon SD, McGowan PO. Genome-epigenome interactions
695 associated with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Epigenetics*; 2018;13:1174–90.

696 14. Ekua W Brenu EWB. Methylation Profile of CD4+ T Cells in Chronic Fatigue Syndrome/Myalgic
697 Encephalomyelitis. *J Clin Cell Immunol*; 2014;05.

698 15. Helliwell AM, Sweetman EC, Stockwell PA, Edgar CD, Chatterjee A, Tate WP. Changes in DNA
699 methylation profiles of myalgic encephalomyelitis/chronic fatigue syndrome patients reflect systemic
700 dysfunctions. *Clin Epigenetics*; 2020;12:167.

701 16. Trivedi MS, Oltra E, Sarria L, Rose N, Beljanski V, Fletcher MA, et al. Identification of Myalgic
702 Encephalomyelitis/Chronic Fatigue Syndrome-associated DNA methylation patterns. *PLoS One*;
703 2018;13:e0201066.

704 17. de Vega WC, Erdman L, Vernon SD, Goldenberg A, McGowan PO. Integration of DNA methylation
705 & health scores identifies subtypes in myalgic encephalomyelitis/chronic fatigue syndrome.
706 *Epigenomics*. 2018;10:539–57.

707 18. Gameiro GR, Sinkunas V, Liguori GR, Auler-Júnior JOC. Precision Medicine: Changing the way we
708 think about healthcare. *Clinics*; 2018;73:e723

709 19. Miksad RA, Samant MK, Sarkar S, Abernethy AP. Small But Mighty: The Use of Real-World Evidence

710 to Inform Precision Medicine. Clin Pharmacol Ther.;2019;106:87–90.

711 20. Johnson RK, Vanderlinden LA, Dong F, Carry PM, Seifert J, Waugh K, et al. Longitudinal DNA
712 methylation differences precede type 1 diabetes. Sci Rep; 2020;10:1–13.

713 21. Eriksson I, Undén A-L, Elofsson S. Self-rated health. Comparisons between three different measures.
714 Results from a population study. Int J Epidemiol; 2001;30:326–33.

715 22. Chatterjee A, Stockwell PA, Rodger EJ, Morison IM. Genome-scale DNA methylome and
716 transcriptome profiling of human neutrophils. Sci Data; 2016;3:1–9.

717 23. Chatterjee A, MacAulay EC, Ahn A, Ludgate JL, Stockwell PA, Weeks RJ, et al. Comparative
718 assessment of DNA methylation patterns between reduced representation bisulfite sequencing and
719 Sequenom EpiTyper methylation analysis. Epigenomics; 2017;9:823–32.

720 24. Chatterjee A, Rodger EJ, Stockwell PA, Weeks RJ, Morison IM. Technical Considerations for Reduced
721 Representation Bisulfite Sequencing with Multiplexed Libraries. J Biomed Biotechnol;2012;2012.

722 25. Wang Z, Yin J, Zhou W, Bai J, Xie Y, Xu K, et al. Complex impact of DNA methylation on
723 transcriptional dysregulation across 22 human cancer types. Nucleic Acids Res; 2020;48:2287–302.

724 26. Murphy SK, Adigun A, Huang Z, Overcash F, Wang F, Jirtle RL, et al. Gender-specific methylation
725 differences in relation to prenatal exposure to cigarette smoke. Gene; 2012;494:36–43.

726 27. Pheby D, Saffron L. Risk factors for severe ME/CFS. Res Artic Biol Med;2009;1:50–74.

727 28. Tang J, Xiong Y, Zhou H-H, Chen X-P. DNA methylation and personalized medicine. J Clin Pharm
728 Ther;2014;39:621–7.

729 29. Naviaux RK, Naviaux JC, Li K, Bright AT, Alaynick WA, Wang L, et al. Metabolic features of chronic
730 fatigue syndrome. Proc Natl Acad Sci U S A.; 2016;113:E5472–80.

731 30. Sweetman E, Kleffmann T, Edgar C, de Lange M, Vallings R, Tate W. A SWATH-MS analysis of
732 Myalgic Encephalomyelitis/Chronic Fatigue Syndrome peripheral blood mononuclear cell proteomes
733 reveals mitochondrial dysfunction. J Transl Med; 2020;18:365.

734 31. Hardcastle S, Brenu E, Johnston S, Nguyen T, Huth T, Ramos S, et al. Hardcastle SL, Brenu EW,
735 Johnston S, et al. Longitudinal analysis of immune abnormalities in varying severities of Chronic
736 Fatigue Syndrome/Myalgic Encephalomyelitis patients. J Transl Med. 2015;13:299.

737 32. Carpenter K, Pollitt RJ, Middleton B. Human liver long-chain 3-hydroxyacyl-coenzyme A
738 dehydrogenase is a multifunctional membrane-bound beta-oxidation enzyme of mitochondria. Biochem
739 Biophys Res Commun. 1992;183:443–8.

33. Fluge Ø, Mella O, Bruland O, Risa K, Dyrstad SE, Alme K, et al. Metabolic profiling indicates impaired pyruvate dehydrogenase function in myalgic encephalopathy/chronic fatigue syndrome. JCI Insight. American Society for Clinical Investigation; 2016;1.
34. Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al. Myalgic encephalomyelitis: International Consensus Criteria. J Intern Med; 2011;270:327–38.
35. Stockwell PA, Chatterjee A, Rodger EJ, Morison IM. DMAP: Differential methylation analysis package for RRBS and WGBS data. Bioinformatics; 2014;30:1814–22.
36. Chatterjee A, Stockwell PA, Rodger EJ, Morison IM. Comparison of alignment software for genome-wide bisulphite sequence data. Nucleic Acids Res; 2012;40:e79.
37. Chatterjee A, Rodger EJ, Ahn A, Stockwell PA, Parry M, Motwani J, et al. Marked Global DNA Hypomethylation Is Associated with Constitutive PD-L1 Expression in Melanoma. iScience.; 2018;4:312–25.
38. Chatterjee A, Macaulay EC, Rodger EJ, Stockwell PA, Parry MF, Roberts HE, et al. Placental hypomethylation is more pronounced in genomic Loci Devoid of retroelements. G3 Genes, Genomes, Genet; 2016;6(7):1911–21.
39. Chatterjee A, Stockwell PA, Rodger EJ, Duncan EJ, Parry MF, Weeks RJ, et al. Genome-wide DNA methylation map of human neutrophils reveals widespread inter-individual epigenetic variation. Sci Rep; 2015;5(1):17328.
40. Snel B, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res. 2000;15(28):3442–4.
41. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford); 2017 Jan 1;2017.
42. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: From gene data mining to disease genome sequence analyses. Curr Protoc Bioinforma. John Wiley and Sons Inc.; 2016;2016:1.30.1–1.30.33.

770

771

772

773 **Declarations**

774

775 **Competing interests**

776 The authors declare they have no competing interests

777

778 **Author's contributions**

779 AH, AC and WT conceived and planned the analysis. PS performed the sequencing alignment. TE performed
780 the blood processing. AH performed the experiments and analysis. AH, AC and WT wrote the manuscript. All
781 authors read and approved the final manuscript

782

783 **Funding**

784 The project was supported by a grant from the Healthcare Otago Charitable Trust, from the Associated New
785 Zealand Myalgic Encephalomyelitis Society (ANZMES), and from generous private donations from ME/CFS
786 families.

787

788 **Acknowledgments**

789 The authors gratefully acknowledge the two ME/CFS patients who generously participated in this study

790

791 **Ethics approval and consent to participate**

792 This study was carried out under an ethics approval from the Southern Health and Disability Ethics Committee
793 of New Zealand (ethics approval number 17/STH/188)

794

795 All participants in this study provided informed written consents for participation in molecular studies from
796 their donated blood samples, after viewing documentation about the studies.

797

798 **Consent for publication**

799 All participants in this study provided informed written consent for the results for the molecular studies they
800 were involved in to be published.

801

802 **Availability of data and materials**

803 All datasets generated and analysed during this current study are available in the GEO database NCBI
804 ([GSE166592](#))

805

806

807

808

809

810

811

812

813

814

815