Differential T cell reactivity to hypocretin and flu antigen epitopes provides clues to the pathophysiology of Type 1 narcolepsy

Guo Luo[1], Aditya Ambati[1†], Ling Lin[1†], Mélodie Bonvalet[1], Xuhuai Ji[2], Holden Terry Maecker[2], Emmanuel Mignot[1*]

[1]Stanford University Center for Sleep Sciences and Medicine Palo Alto CA 94304, USA.

[2]Immune Monitoring Core, Stanford University Institute of Transplantation and Immunology, Palo Alto, CA 94304, USA.

†: Equally contributing second author

*Correspondence should be addressed to: Emmanuel Mignot, Director, Center For Sleep Sciences and Medicine, 3165 Porter Drive, #2178, Palo Alto CA 94304, USA.

(1)650-725-6517 (tel), (1)650-725-7341 (fax), mignot@stanford.edu.

Abstract

Type 1 narcolepsy (T1N) is caused by a loss of hypocretin (HCRT) neurons. Association with HLA DQB1*06:02/DQA1*01:02 (98% vs 25%) heterodimer (DQ0602), T cell receptor (TCR) and other immune loci suggest autoimmunity but autoantigens are unknown. Onset is seasonal in children and associated with influenza-A. During the 2009-10 season, cases were triggered by pH1N12009 Pandemrix® vaccination in Europe. We screened 966 peptides derived from 1) Pandemrix® vaccine 2) H1N1, and 3) HCRT and RFX4, autoantigens primarily expressed in HCRT neurons for DQ0602 binding, identifying 109 binders. Screening cognate tetramer-specific $CD4^+$ T cells in 6 DQ0602 T1N (5 post-Pandemrix®, one recent onset) and 4 post-Pandemrix® controls showed differential reactivity to 3 immunodominant influenza epitopes, one from pH1N12009 ($pHA_{273-287}$) and two from vaccine backbone strain PR8 (NP_{17-31} and $NP_{261-275}$). Among autoantigens, we discovered extensive reactivity to C-amidated but not native version of $HCRT_{54-66}$ and $HCRT_{86-97}$, two highly homologous peptides, suggesting reduced $CD4^+$ tolerance for post-translational modifications of HCRT. Replication in 35 cases and 22 controls showed higher frequency of tetramer positive $CD4^+$ cells for $pHA_{273-287}$, NP_{17-31} and $HCRT_{54-66-NH2}$ but not $NP_{261-275}$ or $HCRT_{86-97-NH2}$. Single cell TCR sequencing of TCRα/β revealed usage biases consistent with *in-vitro* clonal expansion. TCRα/β CDR3 motifs found in $pHA_{273-287}$, NP_{17-31} tetramer positive $CD4^+$ cells were also found in INFγ secreting $CD4^+$ cells stimulated with Pandemrix®, confirming dominance in DQ0602-mediated influenza responses. Usage of NP_{17-31} specific $CD4^+$ cells was biased toward Vβ4-2, a segment increased by narcolepsy-associated Single Nucleotide Polymorphism (SNP) rs1008599. TCRα/β CDR3 motifs of $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$ tetramers were often shared and more diverse. Particularly notable was extensive sharing of a single CDR3α (associated with various CDR3β) using TRAJ24, a chain modulated by SNPs rs1154155 and rs1483979, two SNPs strongly associated with T1N. This particular CDR3 was not enriched in TCRs isolated with $pHA_{273-287}$ or other flu epitopes so that mimicry with these sequences is still uncertain. Higher $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$ positive $CD4^+$ T cell numbers in T1N together with J24 usage in the corresponding TCRs, suggest that DQ0602-mediated $CD4^+$ responses to HCRT are

causal to T1N. Our results provide the first evidence for autoimmunity and flu involvement in T1N pathophysiology.

Main text

Whereas genetic[1-6] epidemiological[7-11] and pathophysiological[12-14] studies implicate autoimmunity in response to flu infection in type 1 narcolepsy (T1N), a disease caused by hypocretin (HCRT) neuronal loss[15-17] the identity of the autoantigen is unknown. Potential autoantigens explored have included HCRT sequences[18-21], HCRT receptor 2 sequences[22] and the protein TRIB2[23] but these results have been either withdrawn[24] or found not to be conclusive or replicated[25-27]. Similarly, staining of hypothalamic slices with T1N sera has been found negative[28].

Extending on prior work[1-6] a large recent GWAS analysis in over 4,500 patients across multiple ethnic groups confirmed primordial importance of the Human Leukocyte Antigen (HLA) DQ0602 heterodimer in disease predisposition, with important secondary associations in T cell receptor loci α and β and other loci suggesting importance of DQ0602 presentation of antigens to $CD4^+$ by Antigen Presenting Cells (APC) and likely subsequent $CD8^+$ destruction of hypocretin cells[29]. Hypocretin cell killing through $CD4^+$ and $CD8^+$ mediation has also been shown to cause narcolepsy in an animal model[30].

Much has recently been learned regarding what may trigger T1N autoimmunity. Onset of T1N in children, which often can be timed to the week, is seasonal and peaks in spring and summer[10]. Onset has also been associated with Streptococcus Pyogenes infections[9,31] and Influenza-A[10], suggesting trigger may be a winter infection, followed by a variable delay peaking at 5 months where complete hypocretin cell loss has occurred and symptoms appear. Most strikingly prevalence of T1N increased several folds in China and Taiwan with a similar delay following the 2009-2010 “swine” flu H1N1 influenza pandemic (pH1N1)[10,32,33], although this is less variable in other countries[11,34]. Similarly, cases of T1N have been triggered by Pandemrix® in Europe with relative risk increasing 5 to 14-fold in children and adolescents and 2 to 7-fold in adults after

vaccination[33]. As Pandemrix® is an adjuvanted vaccine containing the artificially produced reassortant strain X-179A, a mix of PR8, an old H1N1 strain derived from pre-2009 seasonal H1N1 and containing key H1N1 2009 surface proteins (Hemagglutinin-HA and Neuraminidase-NA)[35], it is likely flu protein are critically involved. The fact the HLA and TCR genetic associations are universal[6,36-40] is also consistent with a flu trigger, as these infections are worldwide[41]. The situation is nonetheless likely complex as even with Pandemrix® vaccination, only 1 for 16,000 of vaccinated children ever developed T1N, suggesting the need for additional factors. Similarly, as mentioned above, effects of the H1N12009 pandemic and vaccination on T1N have been variable across countries (and vaccines), suggesting the need for additional factors.

Results

DQ0602 restricted epitopes of H1N1 flu responses in T1N versus controls suggest a role for $pHA_{273-287}$ and NP_{17-31}

Complementing our previously published data[18,24], we screened overlapping 15-mers for DQ0602 binding for key proteins contained in X-179A and 2009H1N1 wild type, HA, NA, Influenza RNA-dependent RNA polymerase subunit 1 (PB1) and nucleoprotein (NP) (Supplementary Table 1). HA and NA were selected as these are key surface proteins enriched by design in flu vaccines such as Pandemrix® (to facilitate B cell responses)[35,42]. PB1 was selected because it is of H1N12009 origin in Pandemrix® and thus is also shared by the wild type H1N12009 virus and Pandemrix®, two known triggers of T1N[35,42]. We finally added NP as it is the most abundant protein in flu and is also present in large amount in Pandemrix® as enrichment process for surface proteins is not absolute in this vaccine (and many others, although some subunit specific vaccine were available, such as Focetria® for pH1N1) [35]. NP of Pandemrix® is of backbone PR8 origin[35,42]. Screening was primarily done using A/reassortant/NYMC X-179A (California/07/2009 x NYMC X-157, reference ADE2909) sequences complemented by wild type pH1N1 sequence (A/California/07/2009, reference AFM728) when polymorphism was evident, or sequences are of distinct origin (i.e. NP). Screening for

these proteins identified large number of binders (Supplementary Table 1) using the biotin-conjugated $EBV_{490–503}$ epitope (Bio-EBV, Bio-GGGRALLARSHVERTTDE), a known binder[18,43]. Peptides that outcompeted >75% of the reference peptide signal were considered binders, and these were selected for tetramer screening[18 24].

To assess for the presence of $CD4^+$ T cell reactive to these epitopes, likely 9 mer-MHC class II binding repertoires were determined for all binders using three Immune Epitope Database prediction algorithms (SMM_align, NN_align, and consensus) and similarity scores. Peptides (15-mers) surrounding the cores and displacing most of the bio-EBV peptide in competition assays were used to construct tetramers by peptide exchange using biotinylated DQ0602 (Bio-DQ0602) cleaved with thrombin as described in the methods section. These tetramers were then used to FACS stain 10-day expanded peripheral blood mononuclear cells (PBMCs) cultures of 6 cases and 4 DQ0602 controls stimulated with Pandemrix® (100ng/ml) or the corresponding single peptide (6.26 nM) (see methods and staining results in Supplementary Figure 1A). The control and narcoleptic subjects were all known past Pandemrix® vaccinees except for one case who developed T1N independent of vaccination 6 months prior to blood sampling (recent onset case).

Figure 1A shows the number of controls and T1N patients (from 10 subjects) positive for each tetramer staining as well as example staining (Figure 1B) with selected tetramers. Tetramer positivity ranged from no positive to 10 positives for some tetramers (Figure 1A), with mean number of positives per flu protein in post Pandemrix® controls (4) and T1N (5) cases shown in Figure 1C. Post Pandemrix® cases and controls were selected for Figure C comparison to ensure similar flu antigen exposure and as expected no difference was found across cases and controls for these antigens overall (all epitopes). In contrast, lower reactivity with PB1 versus HA, NP and NA was found, likely the consequence of the lower concentration of PB1 in the Pandemrix® vaccine (and all vaccines) in comparison to these other proteins[35].

X179A derived tetramers with the most staining included $pHA_{273-287}$, $pNA_{217-231}$, NP_{17-31}, $NP_{261-275}$ (Figure 1D) suggesting these epitopes are immunodominant (Figure 1A, NP epitopes from pH1N1 instead of PR8 indicated by stars). Other immunodominant peptides include $pHA_{209-223}$, $pHA_{321-335}$, $pHA_{437-451}$, $pNA1_{45-59}$, $pNA1_{217-231}$, $NP_{261-275s}$ (shared epitope between pH1N1 and PR8), $NP_{437-451-pH1N1}$*, and $NP_{365-379}$. $HA1_{273-287}$ is notable as it was previously reported to have homology with prepro-hypocretin sequences $HCRT_{54-66}$ and $HCRT_{86-97}$ and suggested be involved in a possible pathogenic mimicry using Enzyme-Linked ImmunoSpot (ELISpot) results, a publication that was retracted based on flawed ELISpot results (DQ0602 binding results were all verified)[18,24] (note that none of the authors of this publication were not involved in the fraudulent results, as determined by a review of Stanford's internal investigation report by NIH Office of Research Integrity). Interestingly, among immunodominant epitopes, responses to three epitopes $pHA_{273-287}$ (of H1N12009 origin), NP_{17-31}, and $NP_{261-275}$ (of PR8 origin) were stronger in Pandemrix® vaccinated patients versus controls and positive in the early onset cases (Figure 1A) and were thus selected for additional investigation.

Additional analysis of 29 cases and 23 DQ0602 positive controls confirmed significant overabundance of T cells positive for $pHA_{273-287}$ and NP_{17-31} but not $NP_{261-275}$) (Figure 2A and 2B, Supplementary Figure 1B and Supplementary Figure 2) suggesting that these two epitopes could be involved in the recruitment of differential $CD4^+$ T cell populations that could be involved in triggering T1N. Sub-analyses confirmed that these epitopes also gave significantly more response in a total of 11 post-Pandemrix® cases (mean age 27±6yo) versus 8 post-Pandemrix® controls (17±2 yo), an important control as these subjects had a more similar exposure to pH1N1 around the first wave of the pH1N1 pandemic infection (Supplementary Figure 2D). Findings with $pNA_{217-231}$, $NP_{261-275}$ and $NP_{17-31-pH1N1}$, the homologous sequence of PR8 NP_{17-31} in pH1N1 (2 amino acid difference in the core, see Supplementary Table 2) showed no differences overall and in subgroups (Supplementary Figure 2A-C). This further demonstrate that increased reactivity in T1N versus controls is not broadly shared across all H1N1 epitopes, confirming screening results presented in Figure 1C.

Broad $CD4^+$ T cell autoreactivity to amidated $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$ fragments when presented by DQ0602

We next tested whether T1N was associated with autoreactivity involving two proteins known to be primarily colocalized with hypocretin, hypocretin itself and RFX4, a transcription factor primarily expressed in testis and also exceptionally abundant in hypocretin cells. Hypocretin and RFX4 binders were tested in the 10 subjects mentioned above, with results suggesting almost no or little cross reactivity, consistent with T cell tolerance toward autoantigen peptide sequences (Figure 1B). Autoreactivity was significantly lower than for all viral proteins (Figure 1C).

Because the pre-prohypocretin peptide is cleaved and post-translationally modified in the region where homologous sequences $HCRT_{54-66}$ and $HCRT_{86-97}$ are located (resulting in secreted hypocretin-1 and -2), most notably C-amidated, a transformation necessary to its biological activity, we also tested $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$, modified peptides we have previously shown also bind DQ0602 with similar or higher affinity[18]. To our surprise, T cell reactivity to these fragments was high in both DQ0602 controls and T1N subjects (Figure 2B). Of notable interest was the fact $HCRT_{54-66-NH2}$ staining was generally restricted to distinct populations (not unlike what was observed with flu antigens) while $HCRT_{86-97-NH2}$ staining was more tail-like, suggesting large amounts of low affinity tetramer cross-reactivity (Figure 2D). This data was later expanded to 30 cases and 23 DQ0602 positive controls, and a significantly higher percentage of $HCRT_{54-66-NH2}$ but not $HCRT_{86-97-NH2}$ tetramer positive cells was found in T1N (Figure 1C). Sub-analysis of 13 post-Pandemrix® cases (mean age 17±1yo) versus 10 post-Pandemrix® controls (mean age 27±6 yo) confirmed significantly higher reactivity for amidated HCRT peptides (Supplementary Figure 2E). Even more strikingly, in 7 recent onset cases (22±6 yo) versus 21 controls (23±3yo) higher reactivity was found to both $HCRT_{54-66-NH2}$ and $HCRT_{86-97}$ (non-amidated) suggesting an important role in the cause of T1N (Supplementary Figure 2F).

T Cell receptor sequencing of flu tetramer positive cells indicate bias in TCRα/β usage consistent with *in-vitro* clonal expansion

TCR positive cells of 13-14 subjects were next FACS single sorted and sequenced in control and T1N patients (including the 9 post Pandemrix® subjects and one early onset case mentioned above) for $pHA_{273-287}$, NP_{17-31}, $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$. Sporadic additional sequencing (a few subjects, data not shown) was performed for other epitopes. Subjects sequenced are denoted in red in Figure 1C. For comparison, approximately one million of the same subject's peripheral memory $CD4^+$ T cells were sequenced for TCRA and TCRB V and J segment baseline usage (data from: Ollila et al. [29]).

As expected, we found strong significant bias in TCRα/β usage (mostly for TRAV and TRBV segments) specific of each epitope, except for $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$ where large amount of sequence sharing was found across the two highly homologous HCRT epitope and sequence usage diversity was higher in comparison to flu epitopes (Supplementary Figure 3, Supplementary Table 3). For example, TCRA usage of $pHA_{273-287}$ was notable for a most significant enrichment in TRAJ23 (Mantel Haenszel [MH] OR=12.6, $p<10^{-300}$),TRAV13-1 (OR=8.3 , $p<10^{-247}$) and TRAV35 (OR=8.3, $p<10^{-247}$) while TCRB usage for the epitope was enriched in TRBV19 (OR=29.4, $p<10^{-300}$) and TRBV4-3 (OR=30.3, $p<10^{-300}$) in comparison to the same subject's CD4+ cell peripheral usage, without differences in control versus T1N. Similarly, NP_{17-31} positive TCR cells were enriched in TRAV12-2 (OR=15.3, $p<10^{-300}$), TRAV8-6 (OR=10.3, $p<10^{-284}$), TRBV4-2 (OR=60.0, p< $p<10^{-300}$), TRBV7-9 (OR=13.1, $p<10^{-289}$), TRBV20-1(OR=4.1, $p<10^{-92}$) (Supplementary Figure 3, Supplementary Table 3) without differences in control versus T1N. Of particular interest was enrichment in TRBV4-2, as T1N is strongly associated with rs1008599G (OR=0.81, $p=6.63\ 10^{-12}$), a SNP that is the main eQTL for increased expression of this segment [29].

To explore differential CDR3 in control versus narcolepsy with these epitopes we also used the Grouping of Lymphocytes Interactions by Paratopes Hotspots (GLIPH) [44] in

CDR3βs from $pHA_{273-287}$ and NP_{17-31}. In $pHA_{273-287}$ sorted cells, significant ($p<0.001$) enriched motifs were TGSP and QVSN in controls and QGR, SQGR, QGRT, LYG, SID, DQGL and ATNE in patients, suggesting more diversity. In NP_{17-31} sorted cells, specific T cells, DSN and TDS were enriched in both controls and narcolepsy cases, but TNYG, GTDS, KL, TNY and SHS motifs were only significantly enriched in narcolepsy cases.

T Cell receptor sequencing of C-amidated HCRT tetramer positive cells indicate broad TCR reactivity and has consequently more limited bias in TCRα/β usage

Sequencing of amidated-HCRT positive tetramers gave a different picture (Supplementary Figure 3, Supplementary Table 3) in comparison to flu-positive tetramer. First, although increased usage of specific TCR segments was found, usage diversity was much higher (Figure 2), notably when $HCRT_{86-97-NH2}$ tetramers were used, giving a typically less distinct population with likely lower TCR affinity (Figure 1B, supplementary Figure 1). In addition, in cases where $HCRT_{54-66-NH2}$ and more rarely $HCRT_{86-97-NH2}$ tetramers gave a clearly separated population on FACS plots (Supplementary Figure 1), sequencing resulted in almost pure oligoclonality of TCR sequences (even less diverse than with flu antigens) with minimal usage sequence sharing (patients 22, 948 and controls 51 and 685 for $HCRT_{54-66-NH2}$ and controls 567 and 391 for $HCRT_{86-97-NH2}$ in Supplementary Figure 3, with corresponding FACS plots in Supplementary Figure 1). In agreement with this, GLIPH analysis did not find significant motif enrichement.

Single cell sorting and TCR sequencing of INFγ positive, Pandemrix® reactive $CD4^+$ T cells confirms the importance of $pHA_{273-287}$ and NP_{17-31} as immunodominant peptides of the HLA class II response to Pandemrix® and suggest mimicry with HCRT $CD4^+$ responses

expanded CFSE tagged PBMCs from 8 patients and 8 controls (overlapping with the sample above) in the presence of Pandemrix® in a 7 day period following which CFSElow $CD4^+$ cells were sorted and expanded further for a 4day period and then

stimulated with autologous monocytes prepulsed with Pandemrix®, following which single cell sequencing of INFγ-positive (Pandemrix®-reactive), CFSE low (having divided in the presence of Pandemrix® during the culture) $CD4^+$ cells was performed. This experimental design amplifies all HLA Class II mediated responses, with DQ0602 mediated responses likely in minority, notably because DRB1 responses are typically dominant among ClassII responses. To limit this problem, one patient was also cultured with Pandemrix® in the presence of anti DR and anti DP antibodies. In this case, presentation and reactivity should be limited or enriched for DQ response (DQ0602 and the other trans located allele). Upon completion of these experiments, CDR3 TCRα and TCRβ sequences were matched with tetramer-derived sequences to search for commonality (Supplementary Table 4). Strikingly, we found that $pHA_{273-287}$ and NP_{17-31} tetramer sequences, in many cases sharing both TCRα and β chains were present in these cultures, most notably in the cultures than had been blocked by anti DR and DP antibodies. Interestingly, we also found a significant number of sequences shared with $HCRT_{54-66-NH2}$, $HCRT_{86-97-NH2}$ tetramer sequences suggesting mimicry with one or several unknown peptide variants in Pandemrix®. As a control, no sharing was found with single sorted $CD8^+$ cell TCR sequences reactive to Pandemrix® were isolated using a similar procedure (data not shown).

Shared CD3 analysis across subjects per tetramer indicate extensive sharing of J24 CDR3α CAVETDSWGKLQF

We next analyzed sharing of CDR3α, CDR3β and CDR3α/β pairs across patients and diagnosis (Figure 3). Not surprisingly, sharing of CDR3α was more frequent than sharing of CDR3β and exact CDR3α/β pairs were very rarely shared across subjects. Sharing for viral epitope $pHA_{273-287}$ and NP_{17-31} tetramers was as expected from strong usage preference reported in Supplementary Figure 3 (for example TRAJ23 for $pHA_{273-287}$CDR3α CATDAYNTDKLIF). The most striking finding was extensive sharing of J24 CDR3α CAVETDSWGKLQF in 3 and 2 patients with $HCRT_{86-97-NH2}$ and $HCRT_{54-66-NH2}$ respectively as well as in 4 controls with $HCRT_{86-97-NH2}$ and $HCRT_{54-66-NH2}$ (Figure 3G

and J). As J24 is a low usage segment (0.8%), this is highly significant. Of notable importance was the fact these J24 CDR3 α did not share CDR3β (see supplementary Table 5) except for 2 patients (26 occurrences) sharing CDR3 α CAVETDSWGKLQF: CDR3 β CATSSGGGGKAYGYTF (Figure 3). This suggests that TCRβ is less important for interaction of these TCR α/β with $HCRT_{86\text{-}97\text{-}NH2}$ and $HCRT_{54\text{-}66\text{-}NH2}$ when presented by DQ0602. In addition, we found that a specific CDR3β CASSQERGSYNEQFF using Vβ4-2 was the second top CDR3β recognizing $NP_{17\text{-}31}$ across subjects, with positivity in 4 patients (48 occurrences) but no controls. Sharing of Vβ4-2 containing CRD3β CASSQETQGRNYGYTF (labelled in red) was also seen in $pHA_{273\text{-}287}$ and $HCRT_{86\text{-}97\text{-}NH2}$ associated tetramers. These did not share similar CRD3α however (see supplementary Table 6). For pairing and details of all TRAJ24 and TRVB4-2 positive clones obtained with various tetramers see below.

Shared CDR3 analysis across tetramers does not reveal evident mimicry between tested flu epitopes and $HCRT_{86\text{-}97\text{-}NH2}$ and $HCRT_{54\text{-}66\text{-}NH2}$

As mentioned above, $pHA_{273\text{-}287}$ and $HCRT_{54\text{-}66\text{-}NH2}$/ $HCRT_{86\text{-}97\text{-}NH2}$ have been suggested to share epitope homology[18,24], a result confirmed using http://tools.iedb.org/cluster/reference/[45], with three EVB sequences as anchor (Figure 3E). Based on this finding, mimicry was proposed but not demonstrated[18,24]. To search for mimicry across $pHA_{273\text{-}287}$, $NP_{17\text{-}31}$, $HCRT_{54\text{-}66\text{-}NH2}$ and $HCRT_{86\text{-}97\text{-}NH2}$, we first analyzed sharing of CDR3α, CDR3β and CDR3α/β pairs across epitopes. Supplementary Table 7 shows cross tabulations of shared sequences across epitopes.

A large number of sequences were shared between $HCRT_{54\text{-}66\text{-}NH2}$ and $HCRT_{86\text{-}97\text{-}NH2}$ tetramers across CDR3α (6%), CDR3β, (10%) and both CDR3 α/β (3.4%), a result that is not surprising considering that $HCRT_{54\text{-}66\text{-}NH2}$ and $HCRT_{86\text{-}97\text{-}NH2}$ only differ in their N-amidated C-terminal end, L-$_{NH2}$ (hypocretin-1) and M-$_{NH2}$ (hypocretin-2) in P9 of the binding repertoire (Supplementary Table 7). In contrast, $pHA_{273\text{-}287}$, a viral epitope which has been suggested to be a mimic of HCRT and has some homology with $HCRT_{54\text{-}66\text{-}NH2}$

and $HCRT_{86\text{-}97\text{-}NH2}$ (see Figure 3E) shared very few of the same clones with $HCRT_{54\text{-}66\text{-}NH2}$ and $HCRT_{86\text{-}97\text{-}NH2}$, for example only 6% for CDR3α, 1.6 % for CDR3β (sharing at that level could be just above random) and none and 2 for both CDR3α/β for $HCRT_{54\text{-}66\text{-}NH2}$ and $HCRT_{86\text{-}97\text{-}NH2}$ respectively. Potential sharing was even lower with $NP_{17\text{-}31}$, which shared a few clones with $pHA_{273\text{-}287}$, but at even lower level with $HCRT_{86\text{-}97\text{-}NH2}$ and $HCRT_{54\text{-}66\text{-}NH2}$. Overall, the result suggests that the more distant the homology (as illustrated in Figure 3E), the less sharing occurred, and this can only be tested formally with full CDR3α/β heterodimers as there is background sharing by chance of CDR3α and, more occasionally, CDR3α. As some of the cross-reactivity could involve only CDR3α (our main $HCRT_{NH2}$ clone is CDR3α specific and has variable CDR3β) or only, CDR3β these experiments did not allow us to conclude clearly whether or not $pHA_{273\text{-}287}$ or $NP_{17\text{-}31}$ could act as mimics.

TRAJ24, TRAJ28 and TRBV 4-2 specific analysis

As mentioned above, TRAJ24, TRAJ28 and TRBV 4-2 are of special interest as QTL locations modulating usage of these segments. For these reasons, we explored whether single cell sequencing directed against these particular segments showed deviation in T1N versus controls. While TRAJ28 did not show any particular over representation in our data, we observed extensive usage of TRAJ24 and TRBV4.2. Supplementary Table 5 reports on all TRAJ24 positive clones found in all cultures. We focused on clones containing J24 CDR3α CAVETDSWGK<u>L</u>QF the major CDR3α shared across patients and controls following $HCRT_{NH2}$ stimulation (Figure 3). All these clones also use TRAV2. As can be seen, this CDR3α generally occurs in the context of diverse CDR3βs containing TRBV15, TRBV2, TRBV29-1, TRBV9 in controls and TRBV15, TRBV3-1, TRBV7-2 and TRBV20-1 in patients with only TRAV15-CATSSGGGGKAYGYTF CDR3β shared by 2 patients and one control following $HCRT_{NH2}$ culture and tetramer isolation. Of note, CAVETDSWGK<u>F</u>QF was not observed in any subject, control or T1N. Supplementary Table 6 reports on all TRBV4-2 positive clones found in all cultures. One potentially interesting finding is the frequent

CDR3β CASSQERGAYNEQFF in association with TRAV12-2 in 4 patients (with TRAJ22 in 3) and no controls after NP_{17-31} stimulation.

Discussion

In this work, we screened two potential autoantigens for DQ0602 binding and T cell reactivity using tetramer staining of PBMCs of T1N and control cases. We found limited TCR reactivity directed toward any segment of the native HCRT sequence, but strong reactivity to the N-Amidated, post-translationally modified (PTM) C-terminal end of hypocretin-1 and 2 (denoted $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$ or, as a group as $HCRT_{NH2}$). To the best of our knowledge, this is the first study to find genuine reactivity in narcolepsy cases and controls to the amidated $HCRT_{NH2}$. These two peptides, also called orexin-A and orexin -B are homologous peptides (differing only in their C-terminal amidated residue) secreted by hypocretin neurons, that undergo amidation of their C-terminal end, a PTM necessary for biological activity. This extends other observations made in other autoimmune diseases such as Rheumatoid Arthritis, celiac disease and type 1 diabetes[46] indicating that PTMs are frequently present in autoantigens. Although this is not proven, it may be that PTMs do not occur in the thymus during central tolerance establishment.

Relevant to T1N, we also found that patients have higher numbers of $CD4^+$ T cells binding $HCRT_{54-66-NH2}$ when presented by DQ0602, the major HLA allele associated with T1N (97% versus 25% in controls). Recent onset cases also showed significantly higher numbers of T cells recognizing $HCRT_{86-97-NH2}$, a peptide differing from only one amino-acid with $HCRT_{86-97-NH2}$. These results indicate that autoreactivity to amidated $HCRT_{NH2}$ fragments is common in both patients and controls but occurs at significantly higher levels in patients. These findings are consistent with other publications reporting cross-reactivity to HCRT in both cases and controls[19,20].

The findings above could reflect causality, cross-reactivity secondary to hypocretin cell loss via another mechanism, or an epiphenomenon. Strikingly however, the top

CDR3α CAVETDSWGKLQF involved in the recognition of $HCRT_{NH2}$ in controls and patients uses J24, a segment marked by a SNP strongly associated with T1N risk [rs1154155G (OR=1.64, $p=1.4810^{-73}$)]. This genetic association is the strongest association after HLA across all ethnicity[1-6] and as J24 is normally only used rarely (0.8% of repertoire) [29], it is a very striking finding. TCRA rs1154155G is linked (r^2=0.99) with rs1483979 (J24F/L), a substitution that changes an aminoacid in the TRAJ24 CDR3 region and could thus potentially affect epitope binding. The polymorphism also globally decreases TRAJ24 so that, for example, TRAJ24F/L heterozygous subjects contain 60% TRAJ24L and 40% TRAJ24F when their repertoires are sequenced[29]. Considering these associations, our data suggest that T1N is caused by autoimmunity to $HCRT_{NH2}$ either through dysregulation of regulatory T cells or through increased pathogenicity of TRAJ24F. Because we found that many controls have the same CDR3α, however, it is likely not sufficient in most cases to escalate to full hypocretin cell loss and establishment of the T1N phenotype. Additional phenotyping of autoreactive cells carrying CDR3α CAVETDSWGKLQF will be needed to answer these questions, together with $CD8^+$ T cell studies. Genetic [29] and animal studies[30] indeed suggest a $CD4^+$ helper $CD8^+$ cytotoxic model as the most likely to explain hypocretin cell loss in T1N.

We also demonstrate that T1N is associated with differential reactivity to at least two flu peptides derived from X-179A, the reassortant strain used in Pandemrix®. This vaccine has been significantly associated with T1N onset in Northern Europe[34], although risk varies across countries[34], a result that suggest interactions with other factors. Of interest was the fact one peptide, $pHA_{273-287}$, is specific of 2009H1N1, while another peptide NP_{17-31} is specific of the vaccine backbone strain PR8, an old seasonal H1N1 strain originally derived from 1918H1N1. The $pHA_{273-287}$ sequence has been shown to have significant homology with $HCRT_{54-66}$ and $HCRT_{86-97}$ sequences and binding repertoire, and could therefore be involved in mimicry with $HCRT_{NH2}$, although this could not be clearly demonstrated in this study. The NP_{17-31} sequence is representative of seasonal H1N1 prior to the 2009 Pandemic, a strain currently circulating at a lower frequency. TCR sequencing of NP_{17-31} tetramer positive cells suggests recruitment of

Vβ4-2 in the recognition of this peptide when presented by DQ0602. As Vβ4-2 is positively regulated by a T1N GWA-associated Quantitative Trait Loci (QTL), causality is also possible with this epitope, providing downstream mimicry or another mechanism is involved with this segment. Unlike with $pHA_{273-287}$ however, sequence comparison with $HCRT_{54-66}$ and $HCRT_{86-97}$ does not reveal obvious sharing of aminocids presented to TCR is the context of this epitope binding to HLA, making it difficult to explain mimicry with the simple hypothesis of hotspot binding. A possibility maybe that HCRT amidation induces broad cross reactivity via CDR3α contacts without the need for strong CDR3β contacts; CAVETDSWGKLQF is paired with various CDR3β in our dataset. In this scenario, flu responses more driven by CDR3β (possibly involving multiple epitopes) could occasionally amplify a TCR heterodimer containing CAVETDSWGKLQF, inducing autoimmunity via cross-reactivity to $HCRT_{NH2}$. Additional studies involving sequencing TCRs recognizing more flu epitopes in the context of DQ0602, cross cultures and eventually structural studies will be needed to pursue to address this hypothesis.

Our data also suggest that flu epitopes derived from multiple strains may be needed to favor immune response toward T1N immunity. If this is the case, a particular history of coincident flu infections could trigger T1N, explaining why the H1N12009 pandemic may have increased cases more clearly China and Taiwan[10,33] versus other countries. Similarly, the diversity of flu epitopes involved may explain differential risk after Pandemrix® vaccination, as it has been suggested vaccination coincidentally with the H1N12009 pandemic was a possible factor explaining why this vaccine in Northern Europe, and not administration of a very similar adjuvanted vaccine to Pandemrix®, Arepanrix®, in Canada was not strongly associated with T1N[47]. Our result is a first step in attempting to unravel this complexity. More extensive tetramer sequencing studies using bar coded tetramers will be needed to fully understand the spectrum of cross reactivity of flu and hypocretin autoreactive T cells involved in T1N.

Methods

Ethics Statement

This study was reviewed and approved by the Stanford University Institutional Review Board (Protocol # 14325, Registration # 5136). Informed consent was obtained from each participant.

Participants

All patients with narcolepsy (n = 32; 44% male; mean age ± SEM, 18.9 ± 1.7 years; range, 6 to 53 years) had cataplexy with an onset age of 15.6 ± 1.6 years (ranging from 5 to 52 years) and met criteria for International Classification of Sleep Disorders 3 (ICSD3) for type 1 narcolepsy. Cerebrospinal fluid (CSF) hypocretin-1 levels were tested from available 7 patients and all had low level with mean (SEM): 28.04pg/mL ± 13.51, Table 8. Thirteen out of 32 patients had onset post pandemrix vaccination, less than a year following vaccination (all in 2009-2010 except for one case vaccinated with Pandemrix® the following season). Controls (n = 25; 60% male; age, 22.8 ± 2.5 years; range, 10 to 56 years) were unrelated subjects, healthy twin or vaccinated siblings. Fourteen of 25 controls are post pandemrix controls and were vaccinated. All patients and controls were DQB1*06:02-positive. Patient's peripheral Mononuclear Cells (PBMCs) were collected by apheresis at the Stanford Blood Center or through phlebotomy, followed by ficolling and stored in liquid nitrogen until used.

Vaccines

The Pandemrix® vaccine (A/California/7/2009 (H1N1) NYMC X-179A monovalent bulk (inactivated, sterile)) (Batch# AFLSFDA280) used throughout the study was manufactured with HA content at 139 µg/ml (determined with single radial diffusion (SRD)) by GlaxoSmithKline (GSK) Dresden in January 2010. This batch has been used during the 2009-2010 vaccination campaign in Europe. Summary of Pandemrix® characteristics can be found in European medicines agency document (http://www.ema.europa.eu/docs/en_GB/document_library/Other/2010/05/WC500091295.pdf).

Peptide library

Overlapping 15-mer peptide (11-amino acid overlap) library covering full length of HA, NA, PB1, HCRT, 23 variants of NP derived from A/California/07/2009 and A/Puerto Rico/8/1934, and 6 isoforms of RFX4 (Supplementary Table 1) were synthesized with >90% purity at GenScript NJ and dissolved in dimethyl sulfoxide (DMSO) at a stock concentration of 5 mM or 10 mM as applicable.

Constructs of DQ0602 and HLA-DM

Truncated coding sequence (CDS) (deletion of signal peptide and transmembrane domain) of HLA-DQB1*06:02 (residues 33-230 of NP_002114.3) together with upstream CLIP peptide and thrombin cleavage site and HLA-DQA1*01:02 (residues 24-217 of NP_002113.2) were amplified with specific primers (DQB1-F: 5'-TATGGATCCCGGGGCCGGCCCAGTGTCCAAGAT-3', DQB1-R: 5'-TATATGAATTCGCCACCTCTAGACTTGGACTGAGCGCTCTCGC-3'; DQA1-F: 5'-TATGGATCCCGAGGATATCGTCGCCGACCATGTCGC-3', DQA1-R: 5'-TATATGAATTCGCCACCTCTAGACTCTGTGAGTTCGCTCATAG-3') using plasmid C (provided by the Emory University NIH tetramer core facility) as a template. The resulting DQB1*06:02 and DQA1*01:02 alleles were then next cloned downstream of a GP67A signal sequence and upstream of a 6*His tag in baculovirus expression plasmids A and B (gifts from Garcia lab at Stanford), followed by acidic and basic leucine zipper sequences, respectively. As above, HLA-DMA (residues 1-231 of NM_006120.3) and HLA-DMB (residues 1-218 of NM_002118.4) were cloned into plasmid B and A, respectively. Human cDNA library was used as a template with specific primers (DMA-F1: 5'-TCCCCCCGGGATGGGTCATGAACAGAACCAAG-3', DMA-R1: 5'-CCGGAATTCCACATTCTCCAGCAGATCTGAGG-3'; DMB-F1: 5'-TCCCCCCGGGATGATCACATTCCTGCCGC-3', DMB-R1: 5'-CCGGAATTCCTTCAGGGTCTGCATGGGG-3'.

Transfection

DQA1*01:02, DQB1*06:02, HLA-DMA and HLA-DMB were transfected to Sf9 cells (gift from the Garcia lab at Stanford) using a BestBac® 2.0 kit (Cat# 91-002, Expression Systems) according to manufacturer's instructions. Briefly, Sf9 cells were cultured in complete SFM (Sf-900® III SFM (Cat# 12658-019, Gibco) supplemented with 10% FBS, 1% L-glutamine (Cat# 25030-081, Gibco) and 0.04% gentamicin (Cat# 15710-064, Gibco)) at 120rpm, 27°C. 0.9×10^6 Sf9 cells/well were plated in a 6-well plate and incubated for 30 minutes. For each construct, 5 µl of BestBac® DNA and 2 µg of plasmid DNA were diluted in 100 µl of transfection medium, while 6 µl of Expres^{2}TR transfection reagent was diluted in another 100 µl of transfection medium. After incubation for 5 minutes at room temperature, these were combined gently and incubated for 20 more minutes at RT. The mixture was then transferred to a well and incubated with Sf9 cells for 5 hours at 27°C in a water bath while the plate was sealed in a plastic bag. Medium was replaced with fresh complete SFM and plate was incubated for 5 days. Contents were collected and spun down at 1,300rpm for 5 minutes. The supernatant was P0 stock virus. To generate P1 stock virus, 300 ml of Sf9 cells (~1.0×10^6 cells/ml) were seeded and cultured overnight. 200 µl of P0 stock was added when they expanded up to 2.0×10^6 cells/ml. Sf9 Cells were resuspended back to 2.0×10^6 cells/ml every day in the following 3 days by adding complete SFM and then sat down for 7 more days. Cultures were then harvested and spun down at 2,300rpm for 10 minutes and supernatant was stored as P1 stock. To test if the transfection was successful, high five (Hi5) cells (gift from Garcia lab at Stanford) were cultured in protein-free insect cell medium (Cat# 12-730Q, Lonza) supplemented with 0.02% gentamicin at 120rpm, 27°C. 2 ml of Hi5 cells (2×10^6 cells/ml) were plated in a 6-well plate and 50 µl of P0 or P1 stock was added and incubated for 3 days. Contents of each well were harvested and precipitated by adding 25 µl of 2 M Tris-HCl (pH=8.0), 5 µl of 1 M $CaCl_2$ and 1 µl of 1 M $NiCl_2$ per 1 ml, followed by centrifugation for 15 minutes at 12,000rpm. Supernatant was then incubated with 50 µl of Ni-NTA agarose slurry (Cat# 745400.500, Macherey-Nagel) overnight at 4°C by end-to-end rotation. Agarose beads were washed with 1.5 ml of 50 mM Tris-HCl (pH=8.0), 100 mM NaCl and 10 mM imidazole. Proteins were eluted with

elution buffer (50 mM Tris-HCl (pH=8.0), 100 mM NaCl and 300 mM imidazole) and analyzed by coomassie blue (Cat# 161-0803, Bio-Rad) (Fig S4A).

Purification of monomer DQ0602

5 L of Hi5 cells (2.0×10^6 cells/ml) were transfected with 12 ml each of P1 stock of DQA1*01:02 and DQB1*06:02 and cultured for 3 days at 120rpm, 27°C. Cultures were rested down overnight at RT and supernatant was precipitated by adding 125 ml of 2 M Tris-HCl (pH=8.0), 25 ml of 1 M $CaCl_2$ and 5 ml of 1 M $NiCl_2$ while stirring for 1 hour, and then rested down overnight at RT. Supernatant was filtered through a 0.45 μm membrane (Cat# 295-4545, Nalgene) and incubated with 50 ml of Ni-NTA agarose slurry overnight at RT while stirring. Agarose beads were collected using a filter apparatus (Cat# 300-4100, Nalgene) with a 10 μm membrane (REF# AN1H04700, EMD Millipore) and washed once with ~25 ml of 50 mM Tris-HCl (pH=8.0), 100 mM NaCl and 10 mM imidazole. Proteins were eluted with ~20 ml of elution buffer 5 times by end-to-end rotation and elutions were transferred to a centrifuge tube top filter unit (Cat# SCGP00525, EMD Millipore) to remove beads and tested by coomassie blue staining (Fig S4B). Elutions were combined and concentrated in PBS (pH=7.4) using a 10 kDa centrifugal filter (Cat# UFC901024, EMD Millipore). Sample was then injected into a Superdex® 200 increase 10/300 GL column (Cat# 28-9909-44, GE Healthcare) for fast protein liquid chromatography (FPLC) by AKTA® pure 25 (GE Healthcare). Fractions with spectrum peaks were collected and tested by coomassie blue using NuPAGE® Bis-Tris gel (Cat# NP0321BOX, invitrogen) and LDS sample buffer (Cat# NXB31010, Expedeon) without boiling. Similarly, purification and detection of HLA-DM (DMA, 35.8 kDa; DMB, 35.7 kDa) were exactly same as DQA1*01:02/DQB1*06:02 without biotinylation and FPLC.

Biotinylation in vitro

Biotinylation of DQ0602 was conducted in vitro using BirA standard reaction kit (Cat# BirA500, Avidity) according to the manufacturer's instructions. Firstly, buffer was exchanged to 50 mM bicine (pH=8.1). Then, 312.16 μl of purified DQ0602 (4.485 mg/ml; molecular weight: 70.4 kDa) was incubated with 50 μl of BiomixA, 50 μl of

BiomixB, 5 µl of BirA (1 mg/ml), and 82.84 µl of 50 mM bicine (pH=8.1) overnight at 30°C. Free biotin was decreased and buffer was exchanged back to PBS using a 10K centrifugal filter. One aliquot was incubated with streptavidin (Cat# 1244-360, PerkinElmer) and biotinylation quality was determined by coomassie blue. Sample was then run on FPLC and fractions were tested with a NuPAGE® Bis-Tris gel. Fractions 23-26 were combined and adjusted to 2 mg/ml with PBS for peptide exchange and tetramer staining.

Peptide exchange

Cleavage was conducted by incubation of 250 µl of 2 mg/ml biotinylated DQ0602 (Bio-DQ0602) (fractions 23-26), 50 µl of 10X cleavage buffer and 5 µl of thrombin (Cat# 69671, EMD Millipore) in 195 µl of ddH_2O overnight at RT. 5 µl of protease inhibitor (Cat# P8849-1ML, Sigma) was added to the mixture. Peptide exchange was performed with four 2,4-dinitrophenyl (DNP)-conjugated peptides: Epstein-Barr virus (EBV) nuclear antigen-1: $EBV_{486-500}$, [(*GGG*)RALLARSHVERTTDE], $EBV_{486-500}$*, [(*GLY*)RALLARSHVERTTDE(Y)], cytomegalovirus (CMV) phosphoprotein pp65 (CMV_{41-54}), [(*GGGY*)LLQTGIHVRVSQPS], and herpes simplex virus type 2 (HSV-2) VP16 ($HSV\text{-}2_{33-50}$), [(*GGG*)PPLYATGRLSQAQLMPSP] (non-virus amino acid residues were shown in italic). 5 µl of cleaved Bio-DQ0602 (1 mg/ml) was incubated with 0.5 µl of 10% octyl-β-D-glucoside (Cat# O8001-1G, Sigma), 4 µl of 500 mM citrate and 500 mM NaH_2PO_4 (pH=5.2), 0.4 µl of 20 mg/ml DNP-peptide, and 1.85 µl of 11 mg/ml DM (molar ratio of DM:DQ=4) for 3 days at 37°C (volume was brought up to 20 µl with ddH_2O). The reaction was neutralized by adding 1/5 volume of 1 M Tris-HCl (pH=8.0). 12.5 µl of streptavidin microspheres (Cat# 24160, Polysciences) (mixing well) was washed 3 times with 1% BSA in PBS and incubated with neutralized reaction for 1 hour at RT. Microspheres were spun down at 10,000g for 3 minutes, washed twice, and split up into 2 equal aliquots. One aliquot was stained with 10 µl of allophycocyanin (APC) mouse IgG1κ isotype control (Cat# 17-4714-41, eBioscience) and 10 µl of fluorescein isothiocyanate (FITC) anti-mouse IgE (Cat# ab99572, Abcam) for 1 hour on ice in dark. The other aliquot was stained with 10 µl of APC anti-HLA-DQ (Cat# 17-9881-42, eBioscience) and 10 µl of anti-DNP (Cat# ab24319, Abcam; isotype, mouse IgE),

followed by 20 µl of FITC anti-mouse IgE for 1 hour on ice in dark. Microspheres were resuspended with 1% BSA in PBS and run on BD LSRII. Data were analyzed with FlowJo.

Competition binding assay

Peptide binding assay was conducted by incubation of 25 nM cleaved DQ0602, 100 nM HLA-DM, 1 µM biotin-conjugated EBV epitope (Bio-EBV, Bio-(*GGG*)RALLARSHVERTTDE), and 40 µM competitor peptide in reaction buffer (100 mM acetate, pH=4.6, 150 mM NaCl, 1% BSA, 0.5% Nonidet P-40 (IGEPAL CA-630, Sigma), 0.1% NaN_3) with duplicate for 3 days at 37°C. The reaction was quenched by adding two volumes of neutralization buffer (100 mM Tris-HCl, pH=8.6, 150 mM NaCl, 1% BSA, 0.5 Nonidet P-40 (IGEPAL CA-630, Sigma), 0.1% NaN_3). Monoclonal anti-DQ (SPV-L3) antibodies (Cat# BNUM0200-50, Biotium) (1:400 in 100 mM carbonate-bicarbonate buffer, pH=9.5) was coated onto a high binding 96-well plate (REF# 9018, Corning), and incubated with neutralized reaction for 1-2 hours at RT. 100 µl/well Europium (Eu)-labelled streptavidin (Cat# 1244-360, PerkinElmer) (1:1000, 1% BSA in PBS) was added to each well and incubated for 1 hour at RT. Plate was washed 5 times with 300 µl/well wash buffer (0.05% Tween-20 in PBS). 100 µl/well enhancement solution (Cat# 1244-105, PerkinElmer) was then added. DELFIA® time-resolved fluorescence (TRF) intensity was detected using a Tecan Infinite® M1000.

Tetramer generation

The neutralized reaction (above) was incubated with PE-streptavidin (Cat# 405204, BioLegend) at 6:1 molar ratio of Bio-DQ0602:PE-streptavidin (1 µg of Bio-DQ0602:0.62 µl of PE-streptavidin) for 1 hour at RT in dark. 10 µl of streptavidin agarose slurry (Cat# N-1000-005, Solulink) (for every 5 µg of Bio-DQ0602) was spun down at 8,000rpm for 3 minutes, washed once with 1 ml of PBS, and incubated with the mixture for 30 minutes at RT (mixing well for every 10 minutes). Beads were spun down at 12,000rpm for 5 minutes. Supernatant was transferred to a centrifugal filter unit (REF# UFC30GV00, EMD Millipore) and centrifuged at 12,000rpm for 2 minutes. Tetramer DQ0602 was

concentrated in PBS (pH=7.4) by centrifugation at 13,000rpm for 5 minutes using an 100 kDa centrifugal unit (REF#, UFC510096, EMD Millipore) for 3 times.

Pandemrix Cell culture and tetramer sorting

Tetramer staining was previously described[48,49]. Briefly, PBMCs were stimulated with100 ng/ml (final concentration of HA) Pandemrix® in a 50 ml culture flask (2.5×10^6 cells/ml) or 6.25 μM in a 96-well plate ($1\text{-}2.5\times10^6$ cells/ml) for 10 or 14 days at 37°C, 5% CO_2. 20 IU/ml IL2 (Cat# 14-8029-81, eBiosciences) was supplemented from the 8th day in complete RPMI medium (RPMI (Cat# 61870-036, Gibco) supplemented with 10% FBS and 1% Pen/Strep). Medium was changed every 2-4 days. Immediately before staining, tetramer pDQ0602 was spun down at 12,000rpm for 5 minutes if kept at 4°C (less than 2 weeks) and aggregates (if any) removed. Fresh or supernatant tetramer DQ0602-peptide (>10 μg/ml) was incubated with cells in 100 μl of complete RPMI medium for 90 minutes at 37°C, 5% CO_2. Cells were then washed once with cold staining buffer (Cat# 420201, Biolegend) and stained with Brilliant Violet 421® (BV421) anti-human CD3 antibody (Cat# 317344, BioLegend), Alexa Fluor® 488 (AF488) anti-human CD4 antibody (Cat# 317420, BioLegend) and Alexa Fluor® 700 (AF700) anti-human CD8 antibody (Cat# 301028, BioLegend) for 20 minutes on ice. Propidium iodide (PI) was added before running on BD LSRII or sorting on BD Aria at the Stanford shared flow FACS facility. Data was analyzed using FlowJo. For most experiments, multiple repeats were done, and if this was the case the mean of repeat observations for each subject was used in statistical analysis.

To complement tetramer data, $CD4^+$ T cells that had been activated by Pandemrix® were enriched utilizing an IFN-γ secretion assay. Briefly, 10-20 million PBMCs were tagged with 0.5μm CFSE (Cell Trace CFSE Life Technologies ref C34554) in the presence of 100ng (HA) of Pandemrix and expanded (Advanced RPMI supplemented with 10% FCS and IL-2 100IU/mL GIBCO ref 12633-012) for a 7-day period, whilst supplementing the culture as needed. On day 7, cultured PBMCs were harvested and an aliquot FACS phenotyped for quantification of T helper subsets as described in Becattini et al[50] on a FACS BD LSR 2. To sort CFSE-low cells, the remaining portion

was then stained with a combination of CD3 AF647 (Biolegend), CD4 PE Cy7 (Biolegend), CD8 AF 700 (BD Biosciences) and the viability dye 7AAD (Biolegend). The CFSE-low $CD4^+$ T cells, which contain the bulk of proliferating cells in response to Pandemrix® stimulation, were FACS sorted (minimum of 200,000 cells using a FACS Aria 3) into sterile tubes containing Advanced RPMI 10 % FCS with 25mM HEPES. As a control, an equal number of CFSE-high cells from the CD4 gate, were sorted as described above. The sorted $CD4^+$ T cells were further cultured in high dose IL-2 (500IU/mL) for 4-days, following which autologous monocytes (CD14 pre-pulsed with 1ug/mL Pandemrix® for 4 hrs) were added to stimulate sorted CFSE-low versus CFSE-high $CD4^+$ cells in a 1:2 ratio for a 16 hour period. Following this, viable IFNγ producing $CD4^+$ T cells were enriched using a IFN-γ secretion assay (Miltenyi Biotech, Germany) using the two different cell types (CFSE-low & CFSE-high). These were then single cell FACS sorted for TCR sequencing using a FACS Aria 3[51]. In subset of individuals (n=2) the pulsing of monocytes with Pandemrix was performed in the presence of 10ug/mL each of Anti HLA-DP (Leinco Technologies, clone B7/21, ref H260) and Anti HLA-DQ (Abcam clone SPV-L3, ref ab23632).

Single cell TCR sequencing

Single cell paired TCR sequencing was performed as described in Han et al.[51]. Briefly, single cells of $CD3^+CD4^+CD8^-Tetramer^+$ or IFN-γ^+CD4^+ were sorted directly into OneStep® Reverse transcription polymerase chain reaction (RT-PCR) buffer (Cat# 210215, Qiagen) in a 96-well plate, following which the plates were frozen at -80ºC until assayed. Paired TCR CDR3 sequences were obtained by a series of 3 nested PCR reactions. For the first reaction, reverse transcription and pre-amplification were performed with multiplex PCR in a 20 µl of reaction using multiple Vα/Vβ and Cα/Cβ region primers and using the following cycling condition: 50°C 30 minutes; 95°C 15 minutes; 94°C 30 seconds, 62°C 1 minute, 72°C 1 minute × 25 cycles; 72 °C 5 minutes; 4 °C 10 minutes. The second PCR was performed with HotStarTaq® DNA polymerase (Cat# 203205, Qiagen) in a 20 µl of reaction using 1 µl of the first reaction as template and the following cycling condition: 95°C 15 minutes; 94°C 30 seconds, 64°C 1 minute, 72°C 1 minute × 25 cycles (TCR); 72 °C 5 minutes; 4 °C 10

minutes. The third PCR incorporated barcodes to enable sequencing on the Illumina® MiSeq platform and was performed in a 20 µl of reaction using 1 µl of the second PCR product as template. Combined PCR products were run on a 1.2% agarose gel. DNA of around 350 to 380 bp was purified using a QIAquick® gel extraction kit (Cat# 28706, Qiagen) and then sequenced on MiSeq platform for a pair-ended 2x250 run. Fastq files were analyzed with a customized TCRseq pipeline at The Human Immune Monitoring Center (HIMC) at Stanford.

TCR Bulk sequencing

PBMCs were used to separate the $CD4^+$ memory cells (Memory CD4+T cell isolation kit, cat. 130-091-893. Milteny Biotec). TCR alpha and beta repertoires were amplified from total RNA by ARM-PCR procedure[52] and sequenced by 2x101 HiSeq 2000 and 2x151 HiSeq 4000 at the Stanford Genome Sequencing Service Center. Data analysis was as follows. Briefly, the quality of raw FASTQ files was assessed using FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Post quality control, FLASH was used to merge the paired end reads[53]. BLAST databases were constructed from IMGT TCR α/β sequences using local blast binaries[54]. The fastq file was converted into fasta and simultaneously processed into equal chunks of fasta records which were then processed using custom python-based analysis pipeline implemented on the Stanford University Genomics Cluster. Briefly each fasta chunk was used to query the BLAST database and Identify the TCR C, J & V Regions. The CDR3s were then extracted and their frequency along with usage of V, J families computed.

Statistical analyses

Statistical comparisons were performed using two-tailed Mann-Whitney U tests, paired Wilcoxon non parametric tests or Fisher's exact tests whenever most appropriate. Data were plotted using GraphPad PRISM 5. P-value <0.05 were considered statistically significant. All data wrangling and handling was implemented with custom python scripts. R function mantelhaen.test was used to perform Mantel-Haenszel chi-squared test to assess enrichment of TCR families comparing bulk sequencing to single cell sequencing.

Legend to Figures

Figure 1: Results of primary tetramer screening in 10 subjects: A. Tetramer positivity in post-Pandemrix® patients (red, from 5 total), post-Pandemrix® controls (grey, from 4 total) and one early onset patient (onset=0.4 year prior, green, single subject). Tetramers were built one by one using peptide exchange using a DQ602-CLIP tetramer (see Methods). *Viral protein tetramers:* HA=Hemmaglutinin, NA= Neuraminidase, NP= Nucleoproteins, PB1= Polymerase Basic 1. Binder sequences from influenza A/reassortant/NYMC X-179A (reference ADE2909) were primarily used. X179A is a reassortant of A/California/07/2009 (pandemic pH1N1 2009) with A/Puerto Rico/8/1934. A/Puerto Rico/8/1934, is an old pre-2009 H1N1 strain, commonly called PR8, representative of seasonal H1N1 strains prior to the 2009-2010 pandemic and still circulating at low levels. IN X-179A the primary sequence of HA, NA and PB1 is from pH1N1 while the NP sequence is from PR8. Any polymorphic peptide sequence differing in NP (A/California/07/2009) between PR8 (/Puerto Rico/8/and pH1N1) that still bound to DQ0602 was also tested and is labeled with a star (others are of PR8 origin). s=epitope of identical sequence in PR8 and pH1N1. Subjects were tested with tetramers after culture with Pandemrix 100 ng/ml for 10 days. *Autoantigen tetramers:* Prepro-HCRT= preprohypocretin, RFX4=all known isoforms of Regulatory Factor X4, Last six peptides= post-translationally modified sequences of prepro-HCRT modified as per secreted hypocretin-1 and 2, when binding DQ0602, including N-Pyroglutamate $HCRT_{34-48}$ plus $HCRT_{38-52}$, $HCRT_{33-47}$ and $HCRT_{37-51}$ with two double bonds (DB), C-terminal amidated $HCRT_{54-66}$ and $HCRT_{86-97}$ denoted $HCRT_{54-66-NH2}$ and $HCRT_{86-97-NH2}$, as in secreted, mature hypocretin-1 and 2. Subjects were tested after a 10-day culture with 6.25nM of each peptide. The main epitopes selected for further analysis are labelled above each bar (based on immunodominance and likely difference between narcolepsy and controls). Peptides $pHA_{209-223}$, $pHA_{273-287}$, $pHA_{321-335}$, $pHA_{437-451}$, pNA_{45-59}, $pNA_{217-231}$, $pNA_{217-231}$, NP_{17-31}, $NP_{261-275s}$ (shared epitope between pH1N1 and PR8), $NP_{437-451-pH1N1}$*, and $NP_{365-379}$ are responding in 60% of subjects and are considered immunodominants in the DQ0602 response to H1N1. $NP_{17-31-pH1N1}$ the sequence equivalent to NP_{17-31} but derived from pH1N1 instead of backbone PR8 sequence.

HA69 ($pHA_{273-287}$) does not have a strong binder in the same area for PR8. B. Distribution of positive subjects per antigen. Note that response to HA, NA and NP, the main antigens contained in Pandemrix® and known to be responsive to flu infections have the large number of positive subjects (≥ 2) whereas binders from autoantigens RFX4 and HCRT only show sporadic response in a maximum of 2 subjects. C. Mean percent of subjects positive for DQ0602 binders in each proteins in post Pandemrix® subjects. Note that viral peptides derived from HA, NA and NP peptides have a large number of responders (36±3%, n=55) in comparison to lower abundance protein PB1 (26±5%, n=7). The response of non-modified autoantigen sequence is lowest at 5±1%. (n=67). In contrast, the response to post translationally modified (PTM)-modified hypocretin fragment is high (40± 16%, n=6), approaching that of foreign, viral antigens (see t-test statistics in figure). D. Example of FACS staining obtained for various selected tetramers. The entire dataset of FACS staining plots is provided as Supplementary Figure 1. Note tail-like staining in many (but not all) cases with $HCRT_{86-97-NH2}$ but not with other tetramers.

Figure 2: Results of secondary tetramer screening using $pHA_{273-287}$, NP_{17-31} , $HCRT_{54-66}$, $HCRT_{54-66-NH2}$, $HCRT_{86-97}$, and $HCRT_{86-97-NH2}$ in 29 patients and 23 controls. Tetramers were built one by one using peptide exchange using a DQ602-CLIP tetramer (see method). A. Percent of subjects with $pHA_{273-287}$ and NP_{17-31} tetramer positivity in controls versus patients. PBMCs of subjects were grown in Pandemrix® or the cognate peptide for 10 and 14 days in two independent experiments, and as results were repeatable and did not differ statistically, the mean is presented in each dot. T: indicate rare occurrences where the experiment was repeated three times, in this case the mean of 3 is included. Subjects in red are those with single cell sorting and sequencing data available. Number of subjects tested is under parenthesis. B. Percent of subjects with $HCRT_{54-66}$/$HCRT_{54-66-NH2}$ and $HCRT_{86-97}$ /$HCRT_{86-97-NH2}$ tetramer positivity in controls versus patients. PBMCs of subjects were grown with the cognate peptide for 10 and 14 days in two independent experiments, and as results were repeatable and did not differ statistically, the mean is presented in each dot. T: rare occurrences where the experiment was repeated three times, in this case the mean of 3 experiment is used.

Subjects in red are those with single cell sorting and sequencing data available. C: Example of $pHA_{273-287}$ and NP_{17-31} tetramer staining in a few subjects (for all results, see Supplementary Figure 1). D. Example of $HCRT_{54-66}$, $HCRT_{54-66-NH2}$, $HCRT_{86-97}$, and $HCRT_{86-97-NH2}$ tetramer staining in a few subjects (for all results, see Supplementary Figure 1). Note the "tail-like staining" aspect of $HCRT_{86-97-NH2}$ often found with this peptide, although occasional distinct monoclonal population can be observed.

Figure 3: Shared CDR3α, CDR3β and CDR3α/β sequences for $pHA_{273-287}$ (A), NP_{17-31}, (B), $HCRT_{54-66-NH2}$ (C), and $HCRT_{86-97-NH2}$ (D), across subjects. Red: number of patients with corresponding CDR3, Grey, number of controls with same CDR3 (per horizontal axis). Box numbers indicate total of times when the CDR3 occurred across patients and controls. E: $pHA_{273-287}$, NP_{17-31}, $HCRT_{54-66-NH2}$, and $HCRT_{86-97-NH2}$ sequence homology assessed using http://tools.iedb.org/cluster/reference/[45] with three EVB sequences used as anchor. Note high similarity of $HCRT_{54-66-NH2}$, and $HCRT_{86-97-NH2}$, also share homology with $pHA_{273-287}$[18,24] as previously reported, but not NP_{17-31}. Note extensive sharing of TRAJ24 (labeled in red) CRD3 CAVETDSWGKLQF in patients and controls with $HCRT_{86-97-NH2}$ and $HCRT_{54-66-NH2}$. In most cases, these CRD3α CAVETDSWGKLQF do not share the same CDR3β (see supplementary Table 7). We also found sharing of TRBV4.2 containing CRD3β CASSQERGSYNEQFF (labeled in red) in patients with NP_{17-31} only and sharing of TRBV4.2 containing CRD3β CASSQETQGRNYGYTF (labeled in red) in $pHA_{273-287}$ and $HCRT_{86-97-NH2}$ associated tetramers although these do not share a similar CRD3α (see supplementary Table 8).

Acknowledgments

The study was primarily funded by gifts from Wake Up Narcolepsy, Jazz Pharmaceutical and individual patients. Samples have been gathered by defunct NIH grant NS-23724 to Emmanuel Mignot. This work used the Genome Sequencing Service Center by Stanford Center for Genomics and Personalized Medicine Sequencing Center, supported by the grant award NIH S10OD020141. We thank the Emory Tetramer facility for providing DQ0602 tetramers.

References

1 Han, F. *et al.* Genome wide analysis of narcolepsy in China implicates novel immune loci and reveals changes in association prior to versus after the 2009 H1N1 influenza pandemic. *PLoS genetics* **9**, e1003880, doi:10.1371/journal.pgen.1003880 (2013).

2 Faraco, J. *et al.* ImmunoChip study implicates antigen presentation to T cells in narcolepsy. *PLoS genetics* **9**, e1003270, doi:10.1371/journal.pgen.1003270 (2013).

3 Kornum, B. R. *et al.* Common variants in P2RY11 are associated with narcolepsy. *Nature genetics* **43**, 66-71, doi:10.1038/ng.734 (2011).

4 Hor, H. *et al.* Genome-wide association study identifies new HLA class II haplotypes strongly protective against narcolepsy. *Nature genetics* **42**, 786-789, doi:10.1038/ng.647 (2010).

5 Hallmayer, J. *et al.* Narcolepsy is strongly associated with the T-cell receptor alpha locus. *Nature genetics* **41**, 708-711, doi:10.1038/ng.372 (2009).

6 Ollila, H. M. *et al.* HLA-DPB1 and HLA class I confer risk of and protection from narcolepsy. *American journal of human genetics* **96**, 136-146, doi:10.1016/j.ajhg.2014.12.010 (2015).

7 Martinez-Orozco, F. J., Vicario, J. L., De Andres, C., Fernandez-Arquero, M. & Peraita-Adrados, R. Comorbidity of Narcolepsy Type 1 With Autoimmune Diseases and Other Immunopathological Disorders: A Case-Control Study. *J Clin Med Res* **8**, 495-505, doi:10.14740/jocmr2569w (2016).

8 Singh, A. K., Mahlios, J. & Mignot, E. Genetic association, seasonal infections and autoimmune basis of narcolepsy. *Journal of autoimmunity* **43**, 26-31, doi:10.1016/j.jaut.2013.02.003 (2013).

9 Aran, A. *et al.* Elevated anti-streptococcal antibodies in patients with recent narcolepsy onset. *Sleep* **32**, 979-983 (2009).

10 Han, F. *et al.* Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. *Annals of neurology* **70**, 410-417, doi:10.1002/ana.22587 (2011).

11 Dye, T. J., Gurbani, N. & Simakajornboon, N. Epidemiology and Pathophysiology of Childhood Narcolepsy. *Paediatr Respir Rev* **25**, 14-18, doi:10.1016/j.prrv.2016.12.005 (2018).

12 Feltelius, N. *et al.* A coordinated cross-disciplinary research initiative to address an increased incidence of narcolepsy following the 2009-2010 Pandemrix vaccination programme in Sweden. *J Intern Med* **278**, 335-353, doi:10.1111/joim.12391 (2015).

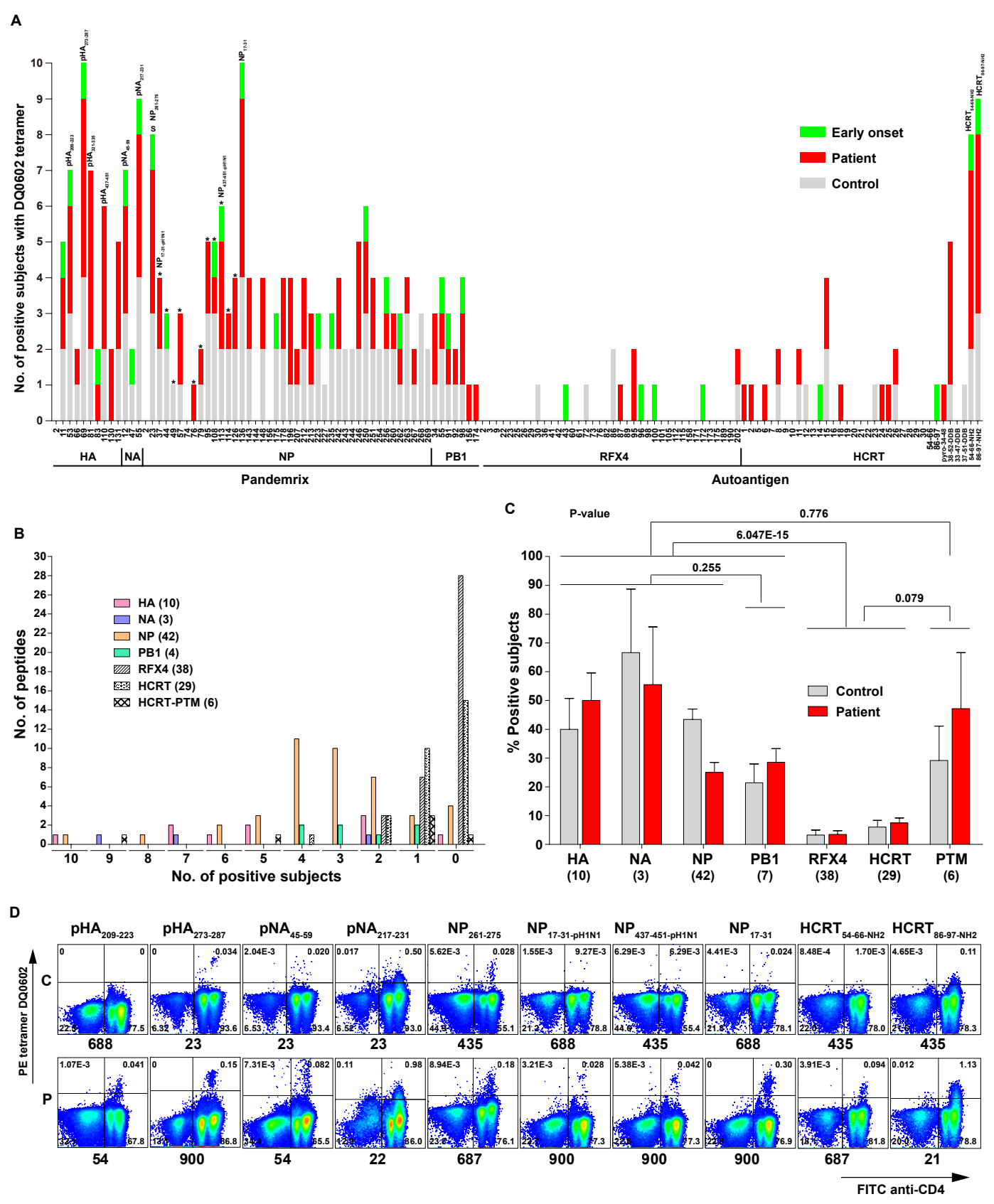
13 Ambati, A. *et al.* Increased beta-haemolytic group A streptococcal M6 serotype and streptodornase B-specific cellular immune responses in Swedish narcolepsy cases. *J Intern Med* **278**, 264-276, doi:10.1111/joim.12355 (2015).

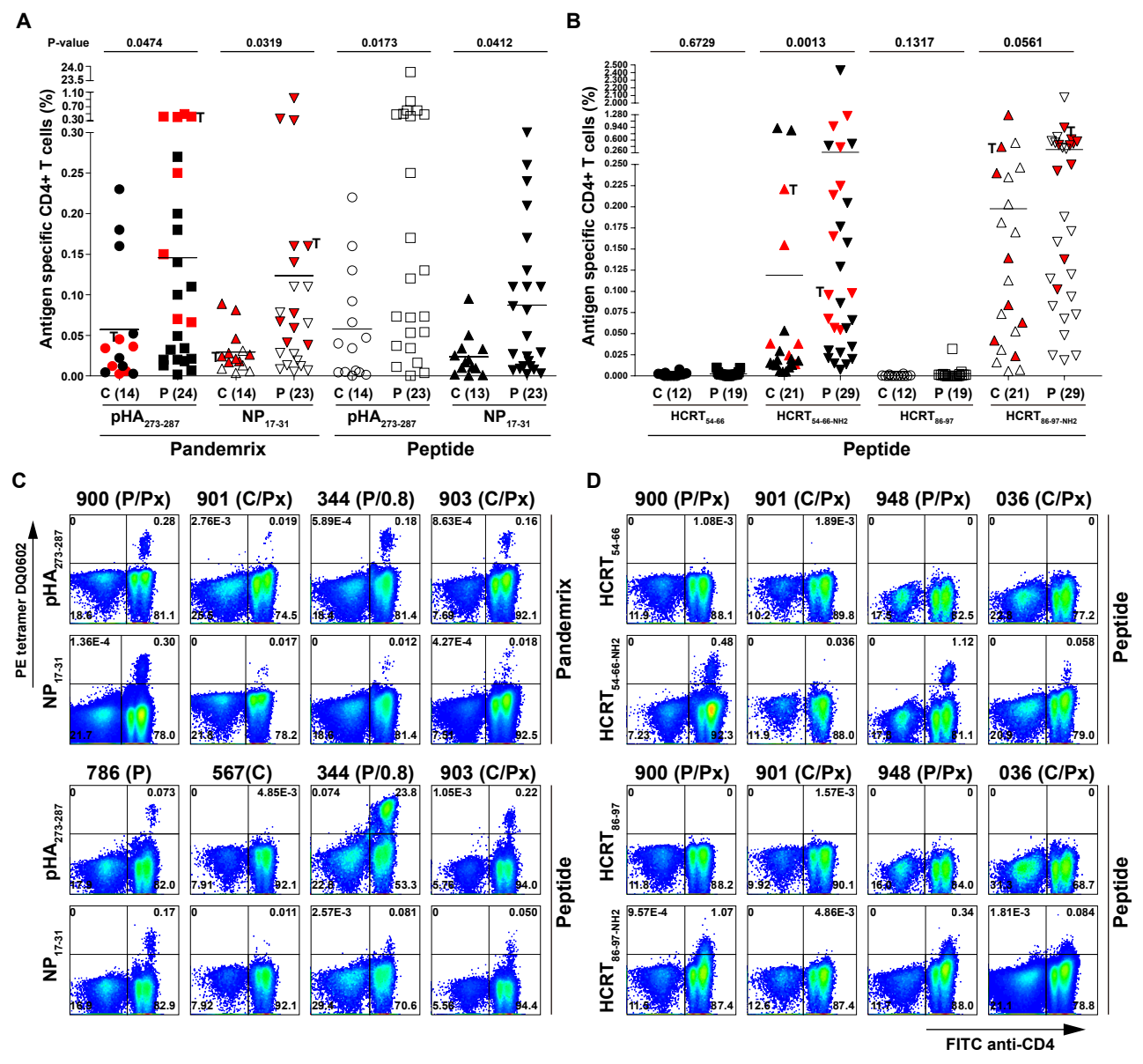
14 Vaarala, O. *et al.* Antigenic differences between AS03 adjuvanted influenza A (H1N1) pandemic vaccines: implications for pandemrix-associated narcolepsy risk. *PloS one* **9**, e114361, doi:10.1371/journal.pone.0114361 (2014).

15 Nishino, S., Ripley, B., Overeem, S., Lammers, G. J. & Mignot, E. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet (London, England)* **355**, 39-40, doi:10.1016/s0140-6736(99)05582-8 (2000).
16 Peyron, C. *et al.* A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nature medicine* **6**, 991-997, doi:10.1038/79690 (2000).
17 Thannickal, T. C. *et al.* Reduced number of hypocretin neurons in human narcolepsy. *Neuron* **27**, 469-474 (2000).
18 De la Herran-Arita, A. K. *et al.* CD4+ T cell autoimmunity to hypocretin/orexin and cross-reactivity to a 2009 H1N1 influenza A epitope in narcolepsy. *Science translational medicine* **5**, 216ra176, doi:10.1126/scitranslmed.3007762 (2013).
19 Kornum, B. R. *et al.* Absence of autoreactive CD4(+) T-cells targeting HLA-DQA1*01:02/DQB1*06:02 restricted hypocretin/orexin epitopes in narcolepsy type 1 when detected by EliSpot. *Journal of neuroimmunology* **309**, 7-11, doi:10.1016/j.jneuroim.2017.05.001 (2017).
20 Ramberger, M. *et al.* CD4+ T-Cell Reactivity to Orexin/Hypocretin in Patients With Narcolepsy Type 1. *Sleep* **40**, doi:10.1093/sleep/zsw070 (2017).
21 Tanaka, S., Honda, Y., Inoue, Y. & Honda, M. Detection of autoantibodies against hypocretin, hcrtrl, and hcrtr2 in narcolepsy: anti-Hcrt system antibody in narcolepsy. *Sleep* **29**, 633-638 (2006).
22 Ahmed, S. S. *et al.* Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. *Science translational medicine* **7**, 294ra105, doi:10.1126/scitranslmed.aab2354 (2015).
23 Cvetkovic-Lopes, V. *et al.* Elevated Tribbles homolog 2-specific antibody levels in narcolepsy patients. *The Journal of clinical investigation* **120**, 713-719, doi:10.1172/jci41366 (2010).
24 De la Herran-Arita, A. K. *et al.* Retraction of the research article: "CD4(+) T cell autoimmunity to hypocretin/orexin and cross-reactivity to a 2009 H1N1 influenza A epitope in narcolepsy". *Science translational medicine* **6**, 247rt241, doi:10.1126/scitranslmed.3009995 (2014).
25 Giannoccaro, M. P. *et al.* Antibodies Against Hypocretin Receptor 2 Are Rare in Narcolepsy. *Sleep* **40**, doi:10.1093/sleep/zsw056 (2017).
26 Luo, G. *et al.* Absence of anti-hypocretin receptor 2 autoantibodies in post pandemrix narcolepsy cases. *PloS one* **12**, e0187305, doi:10.1371/journal.pone.0187305 (2017).
27 Lind, A. *et al.* A/H1N1 antibodies and TRIB2 autoantibodies in narcolepsy patients diagnosed in conjunction with the Pandemrix vaccination campaign in Sweden 2009-2010. *Journal of autoimmunity* **50**, 99-106, doi:10.1016/j.jaut.2014.01.031 (2014).
28 van der Heide, A., Hegeman-Kleinn, I. M., Peeters, E., Lammers, G. J. & Fronczek, R. Immunohistochemical screening for antibodies in recent onset type 1 narcolepsy and after H1N1 vaccination. *Journal of neuroimmunology* **283**, 58-62, doi:10.1016/j.jneuroim.2015.04.008 (2015).
29 Ollila, H. *et al.* Narcolepsy risk loci are enriched in immune cells and suggest autoimmune modulation of the T cell receptor repertoire. *bioRxiv* (2018).

30 Bernard-Valnet, R. *et al.* CD8 T cell-mediated killing of orexinergic neurons induces a narcolepsy-like phenotype in mice. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 10956-10961, doi:10.1073/pnas.1603325113 (2016).
31 Longstreth, W. T., Jr., Ton, T. G. & Koepsell, T. D. Narcolepsy and streptococcal infections. *Sleep* **32**, 1548 (2009).
32 Han, F., Lin, L., Li, J., Dong, X. S. & Mignot, E. Decreased incidence of childhood narcolepsy 2 years after the 2009 H1N1 winter flu pandemic. *Annals of neurology* **73**, 560, doi:10.1002/ana.23799 (2013).
33 Wu, H. *et al.* Symptoms and occurrences of narcolepsy: a retrospective study of 162 patients during a 10-year period in eastern China. *Sleep medicine* **15**, 607-613, doi:10.1016/j.sleep.2013.12.012 (2014).
34 Gomez-Bautista, M., Rojo-Vazquez, F. A. & Alunda, J. M. The effect of the host's age on the pathology of Eimeria stiedai infection in rabbits. *Veterinary parasitology* **24**, 47-57 (1987).
35 Jacob, L. *et al.* Comparison of Pandemrix and Arepanrix, two pH1N1 AS03-adjuvanted vaccines differentially associated with narcolepsy development. *Brain, behavior, and immunity* **47**, 44-57, doi:10.1016/j.bbi.2014.11.004 (2015).
36 Ollila, H. M., Fernandez-Vina, M. & Mignot, E. HLA-DQ allele competition in narcolepsy: a comment on Tafti et al. DQB1 locus alone explains most of the risk and protection in narcolepsy with cataplexy in Europe. *Sleep* **38**, 147-151, doi:10.5665/sleep.4342 (2015).
37 Han, F. *et al.* HLA-DQ association and allele competition in Chinese narcolepsy. *Tissue antigens* **80**, 328-335, doi:10.1111/j.1399-0039.2012.01948.x (2012).
38 Miyagawa, T. *et al.* New susceptibility variants to narcolepsy identified in HLA class II region. *Human molecular genetics* **24**, 891-898, doi:10.1093/hmg/ddu480 (2015).
39 Mignot, E. *et al.* Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. *American journal of human genetics* **68**, 686-699 (2001).
40 Tafti, M. *et al.* DQB1 locus alone explains most of the risk and protection in narcolepsy with cataplexy in Europe. *Sleep* **37**, 19-25, doi:10.5665/sleep.3300 (2014).
41 Bedford, T., Cobey, S., Beerli, P. & Pascual, M. Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). *PLoS pathogens* **6**, e1000918, doi:10.1371/journal.ppat.1000918 (2010).
42 Nicolson, C. *et al.* An additional oligosaccharide moiety in the HA of a pandemic influenza H1N1 candidate vaccine virus confers increased antigen yield in eggs. *Vaccine* **30**, 745-751, doi:10.1016/j.vaccine.2011.11.081 (2012).
43 Demachi-Okamura, A. *et al.* Epstein-Barr virus nuclear antigen 1-specific CD4+ T cells directly kill Epstein-Barr virus-carrying natural killer and T cells. *Cancer science* **99**, 1633-1642, doi:10.1111/j.1349-7006.2008.00852.x (2008).
44 Glanville, J. *et al.* Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**, 94-98, doi:10.1038/nature22976 (2017).
45 Kim, Y. *et al.* Immune epitope database analysis resource. *Nucleic acids research* **40**, W525-530, doi:10.1093/nar/gks438 (2012).

46 Dendrou, C. A., Petersen, J., Rossjohn, J. & Fugger, L. HLA variation and disease. *Nat Rev Immunol* **18**, 325-339, doi:10.1038/nri.2017.143 (2018).
47 Montplaisir, J. *et al.* Risk of narcolepsy associated with inactivated adjuvanted (AS03) A/H1N1 (2009) pandemic influenza vaccine in Quebec. *PloS one* **9**, e108489, doi:10.1371/journal.pone.0108489 (2014).
48 Day, C. L. *et al.* Ex vivo analysis of human memory CD4 T cells specific for hepatitis C virus using MHC class II tetramers. *The Journal of clinical investigation* **112**, 831-842, doi:10.1172/JCI18509 (2003).
49 Newell, E. W., Klein, L. O., Yu, W. & Davis, M. M. Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. *Nature methods* **6**, 497-499, doi:10.1038/nmeth.1344 (2009).
50 Becattini, S. *et al.* T cell immunity. Functional heterogeneity of human memory CD4(+) T cell clones primed by pathogens or vaccines. *Science (New York, N.Y.)* **347**, 400-406, doi:10.1126/science.1260668 (2015).
51 Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol* **32**, 684-692, doi:10.1038/nbt.2938 (2014).
52 Wang, C. *et al.* High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1518-1523, doi:10.1073/pnas.0913939107 (2010).
53 Magoc, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics (Oxford, England)* **27**, 2957-2963, doi:10.1093/bioinformatics/btr507 (2011).
54 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410, doi:10.1016/s0022-2836(05)80360-2 (1990).


A
No. of positive subjects with DQ0602 tetramer
Early onset
Patient
Control
HA
NA
NP
PB1
RFX4
HCRT
Pandemrix
Autoantigen
B
No. of peptides
HA (10)
NA (3)
NP (42)
PB1 (4)
RFX4 (38)
HCRT (29)
HCRT-PTM (6)
No. of positive subjects
C
P-value
0.776
6.047E-15
0.255
0.079
% Positive subjects
Control
Patient
HA (10)
NA (3)
NP (42)
PB1 (7)
RFX4 (38)
HCRT (29)
PTM (6)
D
pHA209-223
pHA273-287
pNA45-59
pNA217-231
NP261-275
NP17-31-pH1N1
NP437-451-pH1N1
NP17-31
HCRT54-66-NH2
HCRT86-97-NH2
PE tetramer DQ0602
C
P
FITC anti-CD4


A
P-value
0.0474
0.0319
0.0173
0.0412
Antigen specific CD4+ T cells (%)
C (14)
P (24)
C (14)
P (23)
C (14)
P (23)
C (13)
P (23)
pHA273-287
NP17-31
pHA273-287
NP17-31
Pandemrix
Peptide
B
0.6729
0.0013
0.1317
0.0561
Antigen specific CD4+ T cells (%)
C (12)
P (19)
C (21)
P (29)
C (12)
P (19)
C (21)
P (29)
HCRT54-66
HCRT54-66-NH2
HCRT86-97
HCRT86-97-NH2
Peptide
C
900 (P/Px)
901 (C/Px)
344 (P/0.8)
903 (C/Px)
PE tetramer DQ0602
pHA273-287
NP17-31
Pandemrix
786 (P)
567(C)
344 (P/0.8)
903 (C/Px)
Peptide
D
900 (P/Px)
901 (C/Px)
948 (P/Px)
036 (C/Px)
HCRT54-66
HCRT54-66-NH2
Peptide
900 (P/Px)
901 (C/Px)
948 (P/Px)
036 (C/Px)
HCRT86-97
HCRT86-97-NH2
Peptide
FITC anti-CD4

A
CDR3α
CDR3β
CDR3α/β
pHA273–287
CAASRGTSYDKVIF
CATDAYNTDKLIF
CAVNARLMF
CIVRAGGTSYGKLTF
CAAPGANNLFF
CASAYTGTASKLTF
CAGATGTASKLTF
CAATGGYNKLIF
CAVSPGNNQGGKLIF
CATDSYNTDKLIF
CATLSYNTDKLIF
CIVRVLTNQGGKLIF
CAVGASAGNMLTF
CAGRPNSGGSNYKLTF
CATDGGGIGAGSYQLTF
CIVRVATNQAGTALIF
CIVRVATNQGGKLIF
CAATMDNYGQNFVF
CIVRVGTNQAGTALIF
CIVRVGTNMDSNYQLIW
CAGHTSNTGNQFYF
CAAPYNQGGKLIF
CAVGVADSWGKLQF
CSVRGLYNEQFF
CASSQETQGRNYGYTF
CAASRGTSYDKVIF:CSVRGLYNEQFF
B
NP17–31
CAVSAYNTDKLIF
CAVSPYNTDKLIF
CAVSQYNTDKLIF
CAVSDQPGGSYIPTF
CVVRGGSGGSYIPTF
CAMREGAGGSYIPTF
CAVRVYNTDKLIF
CVVSEGGYALNF
CAVTYYGNNRLAF
CAVSLYNTDKLIF
CAYRSAILGDDKIIF
CAVTYLGGKLIF
CAASNRDYKLSF
CAGPGGATNKLIF
CAVSPGFKTIF
CAVNKEGGKLIF
CAVAGGFKTIF
CAATYYGQNFVF
CAFMKLYFGNEKLTF
CAVSYGGSQGNLIF
CAVTYYGQNFVF
CAVNTGGFKTIF
CAVPLSGNTGKLIF
CAVIIYNTDKLIF
CASSHSTDTQYF
CASSQERGSYNEQFF
CASSHGGEKLFF
CASSHTTDTQYF
CASSHSADTQYF
CASSYGTDSTDTQYF
CASSEWGSVNEKLFF
CASSQGGTDTQYF
CASSFGTQHQETQYF
CASSHSVDEQFF
CASSFDSSSGNTIYF
CASSFDSSNQPQHF
CASSHSREKLFF
CASSQGGADTQYF
CAVSAYNTDKLIF:CASSHSTDTQYF
CAMREGAGGSYIPTF:CASSFGTQHQETQYF
CAVTYYGNNRLAF:CASSQGGADTQYF
CAVSPYNTDKLIF:CASSHSVDEQFF
CAVSQYNTDKLIF:CASSHSTDTQYF
CAVSQYNTDKLIF:CASSHSADTQYF
CAVRVYNTDKLIF:CASSHGGEKLFF
CAVSPYNTDKLIF:CASSHSTDTQYF
C
HCRT54–66–NH2
CAVETDSWGKLQF
CIVRSQGGSYIPTF
CAASDNDMRF
CAMNNNAGNMLTF
CAMTLYNFNFNKFYF
CILRNPLNFGNEKLTF
CAVNAGRSGGYQKVTF
CATLTLSGSARQLTF
CAYRSDWVSGTYKYIF
CAIPNSGGSNYKLTF
CAENPNNAGNMLTF
CAGRASYNNNDMRF
CAVTTVLDYGGSQGNLIF
CALSEGRGNQGGKLIF
CVVNTGTGGFKTIF
CAMREVYLGGKSTF
CAMSFPGANNLFF
CIALSGGYNKLIF
CAVREVYSGNTPLVF
CAENYGAGGTSYGKLTF
CAVRNTDKLIF
CAVRGYSGGGADGLTF
CATSSGGGGKAYGYTF
CASSVAAAHYGYTF
CASSVVGSYGYTF
CASRDSRAGELFF
CATSDPGRDEQFF
CASSTPTGVSTDTQYF
CASSASSNIQYF
CASSPPGRSGNTIYF
CASSPTGRPQETQYF
CASSTRGHEQFF
CASSFQPFNEQFF
CASSQETQGRNYGYTF
CASSIIPPLPSTSGGADTQYF
CAMSFPGANNLFF:CATSDPGRDEQFF
CAGRASYNNNDMRF:CASSPPGRSGNTIYF
CATLTLSGSARQLTF:CASRDSRAGELFF
CAENYGAGGTSYGKLTF:CASSASSNIQYF
CAVRGYSGGGADGLTF:CASSIIPPLPSTSGGADTQYF
CAVETDSWGKLQF:CATSSGGGGKAYGYTF
CAVREVYSGNTPLVF:CASSFQPFNEQFF
D
HCRT86–97–NH2
CAVETDSWGKLQF
CIVRSQGGSYIPTF
CAASDNDMRF
CAMNNNAGNMLTF
CAALSGGSYIPTF
CVVSPYTGKLIF
CAVDSGNTPLVF
CAENPNNAGNMLTF
CAVEDGGGFKTIF
CAVQATGGFKTIF
CAMREVYLGGKSTF
CVVTSDSSYKLIF
CAMREGSSNTGKLIF
CALFLMNRDDKIIF
CIVRVATNQAGTALIF
CAVRADNYGQNFVF
CALMNNNAGNMLTF
CVARGTNDYKLSF
CSGYALNF
CAENYGAGGTSYGKLTF
SUBJECT NUMBER
CATSSGGGGKAYGYTF
CASSEGTSDLYEQYF
CATSRDTMTSIGTDTQYF
CASSRGEGLSYGYTF
CASSQETQGRNYGYTF
E
pHA273–287
HCRT54–66–NH2
HCRT86–97–NH2
NP17–31
EBV486–500