TITLE

Regulation of mitophagy by the NSL complex underlies genetic risk for Parkinson's disease at Chr16q11.2 and on the MAPT H1 allele

AUTHOR INFORMATION

Marc P.M. Soutar[1,*], Daniela Melandri[1,*], Emily Annuario[1,*], Amy E. Monaghan[2,3], Natalie J. Welsh[4], Karishma D'Sa[1,5], Sebastian Guelfi[1], David Zhang[1], Alan Pittman[6], Daniah Trabzuni[1], Kylie S. Pan[1], Demis A. Kia[1], Magda Bictash[2,3], Sonia Gandhi[1,7], Henry Houlden[1], Mark R. Cookson[8], Nicholas W Wood[1], Andrew B. Singleton[8], John Hardy[1,3], Paul J. Whiting[2,3], Cornelis Blauwendraat[8], Alexander J. Whitworth[4], Claudia Manzoni[1,9], Mina Ryten[1,†], Patrick A. Lewis[1,9,†] & Hélène Plun-Favreau[1,†,‡]

[1] UCL Queen Square Institute of Neurology, London, UK
[2] UCL Alzheimer's Research UK, Drug Discovery Institute, London, UK
[3] UCL Dementia Research Institute, London, UK
[4] MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, UK
[5] KCL, London, SE1 1UL, UK
[6] St Georges University, London, UK
[7] Francis Crick Institute, London, UK
[8] Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
[9] School of Pharmacy, University of Reading, Reading, UK

* and † these authors contributed equally to the work
‡ corresponding author

RUNNING TITLE

New mitophagy Parkinson's risk genes

ABSTRACT

Parkinson's disease (PD) is a common incurable neurodegenerative disease. The identification of genetic variants via genome-wide association studies (GWAS) has considerably advanced our understanding of the PD genetic risk. Understanding the functional significance of the risk loci is now a critical step towards translating these genetic advances into an enhanced biological understanding of the disease. Impaired mitophagy is a key causative pathway in familial PD, but its relevance to idiopathic PD is unclear. We used a mitophagy screening assay to evaluate the functional significance of risk genes identified through GWAS. We identified two new regulators of PINK1-mitophagy, KAT8 and KANSL1, previously shown to modulate lysine acetylation. These findings establish PINK1-mitophagy as a contributing factor to idiopathic PD. *KANSL1* is located on chromosome 17q21 where the risk associated gene has long been considered to be *MAPT*. Our data provide evidence that this assignment is likely to be incorrect and that variability at *KANSL1* underpins this association. Finally, these results enrich our understanding of physiological events regulating mitophagy and establish a novel pathway for drug targeting in neurodegeneration.

KEY WORDS

GWAS / KANSL1 / KAT8 / mitophagy / Parkinson's disease

INTRODUCTION

Parkinson's disease (PD) is the most common movement disorder of old age and afflicts more than 125,000 in the UK [1]. Temporary symptomatic relief remains the cornerstone of current treatments, with no disease-modifying therapies yet available [2]. Until recently, the genetic basis for PD was limited to family-based linkage studies, favouring the identification of rare Mendelian genes of high penetrance and effect. However, genome-wide association studies (GWAS) have identified large numbers of common genetic variants linked to increased risk of developing the disease [3,4]. While these genetic discoveries have led to a rapid improvement in our understanding of the genetic architecture of PD [5], they have resulted in two major challenges for the research community. First, conclusively identifying the causal gene(s) for a given risk locus, and secondly dissecting their contribution to disease pathogenesis. Addressing these challenges is critical for moving beyond genetic insights to developing new disease-modifying strategies for PD.

Previous functional analyses of *PINK1* and *PRKN*, two genes associated with autosomal recessive PD, have highlighted the selective degradation of damaged mitochondria (mitophagy) as a key contributor to disease pathogenesis. In mammalian cells, the mitochondrial kinase PINK1 selectively accumulates at the surface of damaged mitochondria, where it phosphorylates ubiquitin, leading to the recruitment of the E3 ubiquitin ligase Parkin. The recruitment of autophagy receptors leads to the engulfment of damaged mitochondria in autophagosomes, and ultimately fusion with lysosomes [6–11]. It has subsequently become clear that other PD-associated Mendelian genes, such as *FBXO7*, *DJ-1* and *VPS35* [12], are implicated in the regulation of PINK1-mediated mitochondrial quality control. Based upon these data, we hypothesised that PD-GWAS candidate genes may also be involved in this process, providing a mechanistic link between these genes and the aetiology of idiopathic PD. In order to test that hypothesis, we used functional genomics to prioritise candidate genes on the PD GWAS loci, and we developed a biological screening assay as a tool to identify genes that regulate PINK1-mitophagy, and as such, are very likely to be genes that increase the risk of developing PD.

In this study, we show that *KAT8* and *KANSL1*, two genes that were previously shown to be part of the same lysine acetylase complex partially located at the mitochondria [13], are two new important regulators of PINK1-mediated mitochondrial quality control. These findings establish mitophagy as a contributing factor to idiopathic PD and provide a proof of principle for the value of screening approaches to identify causative genes in GWAS loci. Finally, these results suggest lysine acetylation as a potential new avenue for mitophagy modulation and therapeutic intervention.

RESULTS

Genomic analyses of PD have identified over 80 loci associated with an increased lifetime risk for disease. In contrast to Mendelian PD genes, however, the assignment of a causative gene to a risk locus is often challenging. In order to identify new risk genes for PD, we undertook a triage of PD GWAS candidate genes using a combination of methods: i) Colocalization (Coloc) and transcriptome-wide association analysis (TWAS) [14] using expression quantitative trait loci (eQTLs) information derived from Braineac [15], GTEx and CommonMind resources [16,17] to link PD risk variants with specific genes, ii) weighted protein-protein interaction (PPI) network analysis (WPPINA) [18] based on Mendelian genes associated with PD, and iii) the prioritised gene set as described in PD-GWAS [3,19]. 31 open reading frames (ORFs) were nominated as putatively causal for associations at PD risk loci. 55% of these genes were prioritised through multiple techniques (Fig. 1a), with three out of 31 genes (*KAT8, CTSB* and *NCKIPSD*) identified through all three prioritization methods (Fig. 1b, c). The 31 genes, together with 7 PD Mendelian genes and lysosomal storage disorder genes, previously shown to be enriched for rare, likely damaging variants in PD [20], were then taken forward for functional analysis (Supplementary Table 1).

Based upon extensive data implicating impaired mitophagy in the aetiology of familial PD [7], we hypothesized that additional PD-GWAS candidate genes, involved in the most common, idiopathic form of the disease, may play a role in this process. In order to test whether the 38 prioritised genes have a role in PINK1-mitophagy, we developed and optimized a high content screening (HCS) assay for phosphorylation of ubiquitin at serine 65 (pUb(Ser65)), a PINK1-dependent mitophagy marker [21], following mitochondrial depolarization (Extended Data Fig. 1a and 2). The 38 prioritised genes were individually knocked down (KD) using siRNA in Parkin over-expressing (POE)-SHSY5Y human neuroblastoma cells. Increased mitochondrial clearance following mitochondrial depolarization induced by treatment with 10 µM of oligomycin/antimycin A (O/A) was validated as an endpoint for mitophagy (Extended Data Fig. 1b). Over 97% of the pUb(Ser65) signal colocalised with the TOM20 mitochondrial marker in O/A treated cells (Extended Data Fig. 1c, d). siRNA KD efficiency was validated using both a pool of *PINK1* siRNA, which decreased O/A induced pUb(Ser65) and subsequent TOM20 degradation (Extended Data Fig. 1e-g) without decreasing cell viability (Extended Data Fig. 3a, b), and a pool of Polo-like kinase 1 (PLK-1) siRNA that decreased cell viability by apoptosis (Extended Data Fig. 3a, b). The siRNA pools for the 38 candidate genes, together with controls, were screened in duplicate on each plate, across three replicate plates per run. Hits were identified as those wells where O/A-induced pUb(Ser65) was decreased or increased at greater than two standard deviations from the mean of the scramble (SCR) negative control siRNA.

KAT8 was selected based on reproducible downregulation of O/A-induced PINK1-dependent pUb(Ser65) across all three replicates (Fig. 2a and Extended Data Fig. 1h), without affecting cell

viability (Extended Data Fig. 3c). Notably, *KAT8* was selected as a candidate gene on the basis of all three prioritization criteria – namely, proximity of the lead SNP to an ORF, colocalization of a brain-derived eQTL signal with a PD GWAS association signal (Extended Data Fig. 4) and evidence of PPI with a known PD gene (Fig. 1). Furthermore, we find that colocalization and Transcriptome-wide Association Study (TWAS) [22] analyses at this locus are consistent with the KD models in the HCS assay (Supplementary Tables 2 and 3) [17]. Both methods predict that the risk allele operates by reducing *KAT8* expression in PD cases versus controls. The effect of KAT8 KD on pUb(Ser65) was further validated in POE SHSY5Y cells treated with either 1 or 10 µM O/A, using both immunofluorescence (IF) and immunoblotting (IB) (Fig. 3a-d and Extended Data Fig. 5a-e). In order to assess whether other lysine acetyltransferases (KATs) could regulate PINK1-dependent mitophagy, the pUb(Ser65) screen was repeated in POE SHSY5Y cells silenced for 22 other KATs [23,24]. Only *KAT8* KD led to a decreased pUb(Ser65) signal, emphasising the specificity of the KAT8 KD effect on pUb(Ser65) (Fig. 3e and Supplementary Table 4).

These functional data complement and support the Omic prioritization of *KAT8* as a causative gene candidate for the chromosome 16q11.2 PD associated locus (Fig. 2b). To gain further insight into a possible role for KAT8 in the aetiology of PD, we explored the known functional interactions of this protein. KAT8 has previously been shown to partially localise to mitochondria as part of the NSL complex together with KANSL1, KANSL2, KANSL3, and MCRS1 [13] (Fig. 4a). To test whether other components of the NSL complex also modulate mitophagy, the pUb(Ser65) screen was repeated in POE SHSY5Y cells silenced for each of the nine NSL components (HCFC1, KANSL1, KANSL2, KANSL3, KAT8, MCRS1, OGT, PHF20, WDR5) (Fig. 4a). Notably, reduction of KANSL1, KANSL2, KANSL3, MCRS1 and KAT8 expression led to decreased pUb(Ser65) after 1.5 or 3 h O/A treatment (Fig. 4b,c). Strikingly, *KANSL1* is another PD GWAS candidate gene [3]. The effect of KANSL1 KD on pUb(Ser65) was further validated in POE SHSY5Y cells treated with 1 µM O/A, using both IF and IB (Fig. 5a-d). A time course experiment was designed to compare the effect of KAT8 and KANSL1 KD on PINK1-dependent pUb(Ser65) (Fig. 6a,b and Supplementary table 5) and subsequent mitophagy (Fig. 7a,b). While individual KD of either KANSL1 or KAT8 affect phosphorylation of ubiquitin, KANSL1 KD decreased pUb(Ser65) and mitophagy, as assessed by measuring the puncta of the mitochondrial marker PMPCB [25], more efficiently than KAT8 KD (Fig. 6a,b and Fig. 7a,b). To assess the role of KAT8/KANSL1 in neuronal function and survival *in vivo*, we used *Drosophila* as a simple model system. Notably, the NSL complex was originally discovered in *Drosophila* through the homologs of *KAT8* and *KANSL1* (*mof* and *nsl1*, respectively), but null mutants for these genes are associated with developmental lethal owing to profound transcriptional remodelling during development [26]. Therefore, we utilised inducible transgenic RNAi strains to target the KD of *mof* and

nsl1 specifically in neuronal tissues. Using behavioural assays as a sensitive readout of neuronal function we found that pan-neuronal KD of *mof* or *nsl1* caused progressive loss of motor (climbing) ability (Extended data Fig. 6a, b), and also significantly shortened lifespan (Extended data Fig. 6c, d). Interestingly, loss of *nsl1* had a notably stronger effect than loss of *mof*. Consistent with this, KD of *nsl1* but not *mof*, in either all neurons or only in dopaminergic (DA) neurons, caused the loss of DA neurons (Extended data Fig. 6e, f).

KANSL1 is located within the extensively studied inversion polymorphism on chromosome 17q21 (Extended data Fig. 7a, b), which also contains *MAPT* - a gene frequently postulated to drive PD risk at this locus [27]. While the majority of individuals inherit this region in the direct orientation, up to 25% of individuals of European descent have a ~1mb sequence in the opposite orientation [28,29], inducing a larger ~1.3–1.6 Mb region of linkage disequilibrium (LD). Since this inversion polymorphism precludes recombination over a region of ~1.3–1.6 Mb, haplotype-specific polymorphisms have arisen resulting in the generation of two major haplotype clades, termed H1 (common haplotype) and H2 (inversion carriers), previously strongly linked to neurodegenerative disease [30,31]. Due to high LD, the genetics of this region have been hard to dissect, and robust eQTL analyses have been challenging due to the issue of polymorphisms within probe sequences in microarray-based analyses or mapping biases in RNA-seq-based analyses. Several variants (rs34579536, rs35833914 and rs34043286) are in high LD with the H1/H2 haplotype and are located within *KANSL1* (Fig. 8a,b), which could directly impact on KANSL1 protein function. In particular, one of the missense variants is a serine to proline change in KANSL1 protein sequence (S718P), and would therefore be predicted to alter the gross secondary structure of the KANSL1 protein (Fig. 8b). Furthermore, we used allele-specific expression (ASE) analysis to explore the possibility that PD risk might be mediated at this locus through an effect on *KANSL1* expression. Using RNA sequencing data generated from 84 brain samples (substantia nigra n=35; putamen n=49), for which we had access to whole exome sequencing and SNP genotyping data thus enabling mapping to personalised genomes [32], we quantified the variation in expression between the H1 and H2 haplotypes (Supplementary Table 6). While we identified ASE sites within *MAPT* (Extended Data Fig. 8 and Supplementary Table 7), we also identified 4 sites of allele-specific expression in *KANSL1* (Fig. 8a), suggesting that the high PD risk H1 allele is associated with lower *KANSL1* expression, consistent with our functional assessment. Interestingly, sequence analysis of the human *KANSL1* haplotype revealed that the high risk H1 haplotype is the more recent "mutant" specific to *Homo sapiens*, and that other primates and mammals share the rarer non-risk ancestral H2 haplotype (Figure 8b).

To assess the specificity of the KANSL1 KD effect on PINK1-mitophagy, 32 open reading frames in linkage disequilibrium on the H1 haplotype at the 17q21 locus (Extended data Fig. 7a, b

and Supplementary Table 8) were knocked down individually and their effect on pUb(Ser65) was assessed. While the effect of *KANSL1* KD on pUb(Ser65) was confirmed, neither the KD of *MAPT*, nor the KD of each of the other 30 genes on this locus, led to a decreased in the pUb(Ser65) signal (Fig. 9). These data confirm the selectivity of our mitophagy screening assay and suggest that KANSL1 is likely to be a key PD risk gene at the 17q21 locus.

DISCUSSION

Since the first PD GWAS study was performed in 2006 [33] GWAS have identified about 90 independent loci for PD [4]. However, translating GWAS findings into a new molecular understanding of PD-associated pathways and new therapeutic targets has remained a major challenge for the scientific community. In order to screen for PD GWAS candidate genes that play a role in PINK1-mitophagy, and thus are likely to be genuine risk genes for PD, we have set up and optimised a HCS for pUb(Ser65), a marker of PINK1-dependent mitophagy, a key pathway in PD pathogenesis. This approach allowed the successful identification of two new genes associated with increased PD risk, that play a role in mitophagy. Interestingly, these two genes were previously shown to be part of the same complex, the NSL complex.

This study demonstrates the substantial potential of functional screens to exploit genetic data by providing orthogonal information that can confidently identify new risk genes. This is particularly important in genomic regions with uniformly high linkage disequilibrium, such as the 17q21 inversion region which includes 32 ORFs of which many are highly expressed in brain and where existing fine-mapping and functional genomic analyses have been inconclusive. Interestingly, while *MAPT* has long been considered the risk associated gene at this locus, this has recently been questioned by Dong and colleagues, who also raised the significance of KANSL1 in driving PD risk at the locus [34]. Furthermore, functional screening can simultaneously provide mechanistic insights as exemplified in this case by the novel insights we provide into the molecular events regulating mitochondrial quality control and which support a role for mitophagy as a contributing factor to sporadic PD. While our data demonstrate an unequivocal role for KAT8 and KANSL1 in PINK1-dependent ubiquitin phosphorylation and subsequent mitophagy, the precise mechanism of regulation remains to be determined. Our results identify KAT8 and KANSL1 as modifiers of PINK1-dependent pUb(Ser65), suggesting that they may be regulating PINK1 kinase activity. It will be important to determine whether phosphorylation of other PINK1 substrates, such as Parkin and Rab8A [9,35], are modulated by KAT8 and KANSL1. The KAT8 and KANSL1-containing NSL complex functions to promote histone acetylation and as such, is a master regulator of transcription [36]. Therefore, another possibility is that KAT8/KANSL1 may regulate PINK1 transcription, and subsequent translation. Interestingly, depletion of KAT8/KANSL1 was shown to cause significant downregulation of mitochondrial DNA transcription and translation, and ultimately impaired mitochondrial respiration [13]. Thus, it is also

possible that KAT8/KANSL1-dependent modulation of mitochondrial DNA indirectly regulates PINK1 mitochondrial accumulation and subsequent mitophagy. Finally, an intriguing possibility is that the KAT8/KANSL1 complex directly acetylates ubiquitin, which has previously been shown to be acetylated on six out of its seven lysines (K6, K11, K27, K33, K48, K63) [37]. It has been proposed that the KAT8/KANSL1 complex has targets in the mitochondria other than the mitochondrial DNA [13]. Further experiments are required to determine whether KAT8, which has been shown to partially localise at the mitochondrial outer membrane, where ubiquitin is phosphorylated [13], may contribute to ubiquitin acetylation.

Important genetic discoveries in PD, in particular, the identification of the *PINK1* [38] and *PRKN* genes [39], opened the field of selective mitophagy [7]. However, there is still a clear need for a better molecular understanding of mitochondrial quality control. Here we provide new insights into the mechanism by identifying two new molecular players, KAT8 and KANSL1. These new regulators of mitophagy provide the first direct evidence for a role of the PINK1-mitophagy pathway in idiopathic PD and the convergence between familial and idiopathic pathways in disease. Taken together, these findings open a novel avenue for the therapeutic modulation of mitophagy in PD, with potential implications across drug discovery in frontotemporal dementia and Alzheimer's disease, where mitophagy also plays an important role in disease pathogenesis [40].

METHODS

Reagents

Oligomycin (mitochondrial complex V inhibitor) was purchased from Cayman Chemicals (11341) and from Sigma-Aldrich (O4876), and antimycin A (mitochondrial complex III inhibitor) was purchased from Sigma-Aldrich (A8674). All siRNAs were purchased as pre-designed siGENOME SMARTpools from Dharmacon: non-targeting (D-001206-13), PINK1 (M-004030-02), PLK1 (L-003290-00), KIF-11 (L-003317-00), KAT8 (M-014800-00), KANSL1 (M-031748-00), KANSL2 (M-020816-01), KANSL3 (M-016928-01), HCFC1 (M-019953-01), MCRS1 (M-018557-00), OGT (M-019111-00), PHF20 (M-015234-02), WDR5 (M-013383-01). The following antibodies were used for immunocytochemistry: mouse anti TOM20 (Santa Cruz, sc-17764, 1:1000), rabbit anti phospho-ubiquitin (Ser65) (Cell Signaling, 37642, 1:1000), rabbit anti PMPCB (Proteintech 16064-1-AP, 1:1000), AlexaFluor 488 goat anti rabbit (Invitrogen, A11008, 1:2000), AlexaFluor 568 goat anti mouse (Invitrogen, A11004, 1:2000). The following antibodies were used for immunoblotting: mouse anti TIM23 (BD Biosciences, 611223, 1:1000), rabbit anti phospho-ubiquitin (Ser65) (Merck Millipore, ABS1513-I, 1:1000), mouse anti GAPDH (Abcam, ab110305, 1:1000), rabbit anti KAT8 (Abcam, ab200600, 1:1000), IRDye 680LT donkey anti mouse (LI-COR Biosciences, 925-68022, 1:20000), IRDye 800CW donkey anti rabbit (LI-COR Biosciences, 925-32213, 1:20000).

Selection of genes for High Content Screening

Candidates for High Content Screening were selected based on i) WPPINA; ii) complex prioritization; and, iii) coloc analysis. WPPINA analysis is reported in [18] where the 2014 PD GWAS [19] was analysed; candidate genes where selected among those prioritised and with an LD r2 ≥ 0.8. The same pipeline has then been additionally applied to the 2017 PD GWAS [3] to update the list of candidate genes. Briefly, a protein-protein interaction network has been created based on the Mendelian genes for PD (seeds) using data from databases within the IMEx consortium. The network has been topologically analysed to extract the core network (i.e. the most interconnected part of the network). The core network contains the proteins/genes that can connect >60% of the initial seeds and are therefore considered relevant for sustaining communal processes and pathways, shared by the seeds. These processes have been evaluated by Gene Ontology Biological Processes enrichment analysis. The top SNPs of the 2017 PD GWAS have been used to extract open reading frames (ORFs) in cis-haplotypes defined by LD r2 ≥ 0.8 (analysis performed in October 2017) [3]. These ORFs have been matched to the core network to identify overlapping proteins/genes in relevant/shared pathways. Results of complex prioritization (neurocentric prioritization strategy) were gathered from

[3] where this strategy was applied to the 2017 PD GWAS [3]. The coloc analysis was performed as reported in [17], posterior probabilities for the hypothesis that both traits, the regulation of expression of a given gene and the risk for PD share a causal variant (PPH4), were calculated for each gene, and genes with PPH4 ≥ 0.75 were considered to have strong evidence for colocalization. Summary statistics were obtained from the most recent PD GWAS [4] and were used for regional association plotting using LocusZoom [41].

Cell Culture and siRNA transfection

POE SH-SY5Y cells are a kind gift from H. Ardley [42]. Cells were cultured in Dulbecco's Modified Eagle (DMEM, Gibco, 11995-065) and supplemented with 10% heat-inactivated foetal bovine serum (FBS, Gibco) in a humidified chamber at 37 °C with 5% CO_2. For siRNA transfection, cells were transfected using DharmaFECT1 transfection reagent (Dharmacon, T-2001-03) according to the manufacturer's instructions (for concentrations of siRNA, see sections below).

ASEs

Sites of ASE were identified as described by Guelfi and colleagues [32] by mapping RNA-seq data to personalised genomes, an approach specifically chosen because it aims to minimise the impact of mapping biases. RNA-seq data generated from 49 putamen and 35 substantia nigra tissue samples from the UK Brain Expression Consortium was used for this analysis. All samples were obtained from neuropathologically normal individuals of European descent and sites with greater than 15 reads in a sample were tested for ASE. Only sites present in non-overlapping genes were considered and data from both the tissues were considered together to increase power. Sites with minimum FDR < 5% across samples were marked as ASE sites. Plots were generated using Gviz3, with gene and transcript details obtained from Ensembl v92.

High Content siRNA Screen

Cell plating and siRNA transfection

siRNA was dispensed into Geltrex-coated 96-well CellCarrier Ultra plates (Perkin Elmer) at a final concentration of 30 nM using the Echo 555 acoustic liquid handler (Labcyte). For each well, 25 µl of DMEM containing 4.8 µl/ml of DharmaFECT1 transfection reagent was added and incubated for 30 min before POE SH-SY5Y cells were seeded using the CyBio SELMA (Analytik Jena) at 15,000 cells per well, 100 µl per well in DMEM + 10% FBS. Cells were incubated for 72 h before treatment with 10 µM oligomycin/10 µM antimycin for 3 h to induce mitophagy.

IF and Image Capture and Analysis

Cells were fixed with 4% PFA (Sigma-Aldrich, F8775), then blocked and permeabilised with 10% FBS, 0.25% Triton X-100 in PBS for 1 h, before immunostaining with pUb(Ser65) and TOM20 primary antibodies (in 10% FBS/PBS) for 2 h at room temperature. After 3x PBS washes, AlexaFluor 568 anti-mouse and 488 anti-rabbit secondary antibodies and Hoechst 33342 (Thermo Scientific, 62249) were added (in 10% FBS/PBS, 1:2000 dilution for all) and incubated for 1 h at room temperature. Following a final 3x PBS washes, plates were imaged using the Opera Phenix (Perkin Elmer). 5x fields of view and 4x 1 µm Z-planes were acquired per well, using the 40X water objective, NA1.1. Images were analysed in an automated way using the Columbus 2.8 analysis system (Perkin Elmer) to measure the integrated intensity of pUb(Ser65) within the whole cell (see Extended Data Fig. 1 and 2 for image analysis workflow and parameters).

Screen quality control, data processing and candidate selection

Screen plates were quality controlled based on the efficacy of the PINK1 siRNA control and O/A treatment window (minimum 3-fold). Data were checked for edge effects using Dotmatics Vortex visualization software. Raw data was quality controlled using robust Z prime > 0.5. Data were processed using Python for Z score calculation before visualization in Dotmatics Vortex. Candidates were considered a hit where Z score was ≥ 2 or ≤ -2, and where replication of efficacy was seen both within and across plates.

siRNA libraries

The siRNA libraries were purchased from Dharmacon as an ON-TARGETplus SMARTpool Cherry-pick siRNA library, 0.25 nmol in a 384-well plate. siRNAs were resuspended in RNase-free water for a final concentration of 20 µM. SCR, PINK1 and PLK1 or KIF11 controls were added to the 384-well plate at a concentration of 20 µM before dispensing into barcoded assay-ready plates.

Mitochondrial enrichment and Western blotting

POE SH-SY5Y cells were transfected with 100 nM siRNA and incubated for 72 h. Cell lysates were fractionated into cytoplasmic and mitochondria-enriched samples, and run on SDS-PAGE before IB with the Odyssey® CLx Imager (LI-COR Biosciences). Mitochondrial enrichment and Western blotting protocols were described previously [43].

Immunofluorescence

POE SH-SY5Y cells were reverse transfected with 50 nM siRNA in 96-well CellCarrier Ultra plates according to the manufacturer's instructions and incubated for 72 h. Cells were then treated, fixed and stained as per the screening protocol detailed above (for treatment concentrations and times, see figures). Image intensities were not modified or pre-processed. Images are presented as

maximum projections of the pUb(Ser65) or PMPCB-488 channel for one field of view. Insets show the Hoechst 33342 channel for the same field.

Drosophila stocks and husbandry

Flies were raised under standard conditions in a humidified, temperature-controlled incubator with a 12h:12h light:dark cycle at 25°C, on food consisting of agar, cornmeal, molasses, propionic acid and yeast. The following strains were obtained from the Bloomington *Drosophila* Stock Center (RRID:SCR_006457): *mof* RNAi lines, P{TRiP.JF01701} (RRID:BDSC_31401); and P{TRiP.HMS00537} (RRID:BDSC_58281); *nsl1* RNAi lines, P{TRiP.HMJ22458} (RRID:BDSC_58328); the pan-neuronal *nSyb-GAL4* driver (RRID:BDSC_51941); and dopaminergic neuron driver (TH-GAL4; RRID:BDSC_8848); and control (*lacZ*) RNAi P{GD936}v51446) from the Vienna *Drosophila* Resource Center (RRID:SCR_013805). All experiments were conducted using male flies.

Locomotor and lifespan assays

The startle induced negative geotaxis (climbing) assay was performed using a counter-current apparatus. Briefly, 20-23 males were placed into the first chamber, tapped to the bottom, and given 10 s to climb a 10 cm distance. This procedure was repeated five times (five chambers), and the number of flies that has remained into each chamber counted. The weighted performance of several group of flies for each genotype was normalized to the maximum possible score and expressed as *Climbing index* [44].

For lifespan experiments, flies were grown under identical conditions at low-density. Progeny were collected under very light anaesthesia and kept in tubes of approximately 20 males each, around 50-100 in total. Flies were transferred every 2-3 days to fresh medium and the number of dead flies recorded. Percent survival was calculated at the end of the experiment after correcting for any accidental loss.

Immunohistochemistry and sample preparation

Drosophila brains were dissected from aged flies and immunostained as described previously [45]. Adult brains were dissected in PBS and fixed in 4% formaldehyde for 30 min on ice, permeabilized in 0.3% Triton X-100 for 3 times 20 min, and blocked with 0.3% Triton X-100 plus 4% goat serum in PBS for 4 h at RT. Tissues were incubated with anti-tyrosine hydroxylase (Immunostar Inc. #22491), diluted in 0.3% Triton X-100 plus 4% goat serum in PBS for 72 h at 4°C, then rinsed 3 times 20 min with 0.3% Triton X-100 in PBS, and incubated with the appropriate fluorescent secondary antibodies overnight at 4°C. The tissues were washed 2 times in PBS and mounted on slides using Prolong

Diamond Antifade mounting medium (Thermo Fisher Scientific). Brains were imaged with a Zeiss LMS 880 confocal. Tyrosine hydroxylase-positive neurons were counted under blinded conditions.

Statistical Analysis

Intensity measurements from imaging experiments were normalised to SCR O/A for each experiment and presented as a percentage. N numbers are shown in figure legends and refer to the number of independent, replicate experiments. Within each experiment, the mean values of every condition were calculated from a minimum of 3 technical replicates. Intensity measurements from Western blot experiments were normalised to PINK1 O/A. GraphPad Prism 6 (La Jolla, California, USA) was used for statistical analyses and graph production. Data were subjected to either one-way or two-way ANOVA with Dunnett's post-hoc analysis for multiple comparisons, unless otherwise stated. All error bars indicate mean ± standard deviation (SD) from replicate experiments.

Acknowledgements.

This work was supported in part by the UK Medical Research Council (MRC) funding to the Dementia Platform UK (MR/M02492X/1), MRC core funding to the High-Content Biology Platform at the MRC-UCL LMCB university unit (MC_U12266B) and MRC MBU (MC_UU_00015/6), and by UCL Translational Research Office administered seed funds. MS, EA, CM and DT are funded by MRC MR/N026004/1. DM is supported by an MRC CASE studentship (MR/P016677/1). AM, MB and PW are funded by ARUK (ARUK-2018DDI-UCL). MR was supported by the UK MRC through the award of Tenure-track Clinician Scientist Fellowship (MR/N008324/1). This work was supported in part by the Intramural Research Programs of the National Institute on Aging (NIA). We also acknowledge the support of the NIHR BRC award to University College London Hospitals, UCL. Finally, the authors would like to thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to these resources. A full list of contributing groups can be found at https://gnomad.broadinstitute.org/about.

Author Contributions.

HPF, PL, JH, AW and PW conceived the idea. MS, DM, EA, AM, DT, MB, PW, JH, AW, MR, PL and HPF designed the experiments. MS, DM, EA, AM, NW, NW, KDS, SG, DZ, AP, DT, KP, CM, CB and HPF carried out analysis and experiments. MS, DM, EA, AM, PW, CM, AW, MR, PL and HPF wrote the manuscript, with input from all co-authors. HPF, PL and MR supervised the project.

Competing Interests: The authors declare that they have no conflict of interest

Correspondence and requests for materials should be addressed to h.plun-favreau@ucl.ac.uk.

REFERENCES

1. Hardy, J., Lewis, P., Revesz, T., Lees, A. & Paisan-Ruiz, C. The genetics of Parkinson's syndromes: a critical review. *Curr. Opin. Genet. Dev.* **19**, 254–265 (2009).
2. Connolly, B. S. & Lang, A. E. Pharmacological treatment of Parkinson disease: A review. *JAMA - J. Am. Med. Assoc.* **311**, 1670–1683 (2014).
3. Chang, D. *et al.* A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* (2017). doi:10.1038/ng.3955
4. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **18**, 1091–1102 (2019).
5. International Parkinson Disease Genomics Consortium, D. *et al.* Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* (2011). doi:10.1016/S0140-6736(10)62345-8
6. Narendra, D., Tanaka, A., Suen, D.-F. & Youle, R. J. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *J. Cell Biol.* **183**, 795–803 (2008).
7. McWilliams, T. G. & Muqit, M. M. PINK1 and Parkin: emerging themes in mitochondrial homeostasis. *Curr. Opin. Cell Biol.* **45**, 83–91 (2017).
8. Narendra, D. P. *et al.* PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. *PLoS Biol.* **8**, e1000298 (2010).
9. Kazlauskaite, A. *et al.* Parkin is activated by PINK1-dependent phosphorylation of ubiquitin at Ser65. *Biochem. J.* **460**, 127–39 (2014).
10. Shiba-Fukushima, K. *et al.* Phosphorylation of Mitochondrial Polyubiquitin by PINK1 Promotes Parkin Mitochondrial Tethering. *PLoS Genet.* **10**, e1004861 (2014).
11. Lazarou, M. *et al.* The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature* (2015). doi:10.1038/nature14893
12. Plotegher, N. & Duchen, M. R. Crosstalk between Lysosomes and Mitochondria in Parkinson's Disease. *Front. Cell Dev. Biol.* (2017). doi:10.3389/fcell.2017.00110
13. Chatterjee, A. *et al.* MOF Acetyl Transferase Regulates Transcription and Respiration in Mitochondria. *Cell* **167**, 722-738.e23 (2016).
14. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* (2014). doi:10.1371/journal.pgen.1004383
15. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosci.* (2014). doi:10.1038/nn.3801

16. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R. & Lo, E. The Genotype-Tissue Expression (GTEx) project : Nature Genetics : Nature Publishing Group. *Nature* (2013).

17. Kia, D. A. *et al.* Integration of eQTL and Parkinson ' s disease GWAS data implicates 11 disease genes. (2019).

18. Ferrari, R. *et al.* Stratification of candidate genes for Parkinson's disease using weighted protein-protein interaction network analysis. *BMC Genomics* (2018). doi:10.1186/s12864-018-4804-9

19. Nalls, M. A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.* (2014). doi:10.1038/ng.3043

20. Robak, L. A. *et al.* Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. *Brain* (2017). doi:10.1093/brain/awx285

21. Hou, X. *et al.* Age- and disease-dependent increase of the mitophagy marker phospho-ubiquitin in normal aging and Lewy body disease. *Autophagy* (2018). doi:10.1080/15548627.2018.1461294

22. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* (2016). doi:10.1038/ng.3506

23. Simon, R. P., Robaa, D., Alhalabi, Z., Sippl, W. & Jung, M. KATching-Up on Small Molecule Modulators of Lysine Acetyltransferases. *J. Med. Chem.* **59**, 1249–1270 (2016).

24. Sheikh, B. N. & Akhtar, A. The many lives of KATs — detectors, integrators and modulators of the cellular environment. *Nat. Rev. Genet.* **20**, 7–23 (2019).

25. Bertolin, G. *et al.* The TOMM machinery is a molecular switch in PINK1 and PARK2/PARKIN-dependent mitochondrial clearance. *Autophagy* **9**, 1801–1817 (2013).

26. Raja, S. J. *et al.* The nonspecific lethal complex is a transcriptional regulator in Drosophila. *Mol. Cell* **38**, 827–841 (2010).

27. Wray, S. & Lewis, P. A. A tangled web - tau and sporadic Parkinson's disease. *Frontiers in Psychiatry* (2010). doi:10.3389/fpsyt.2010.00150

28. Stefansson, H. *et al.* A common inversion under selection in Europeans. *Nat. Genet.* (2005). doi:10.1038/ng1508

29. Zody, M. C. *et al.* Evolutionary toggling of the MAPT 17q21.31 inversion region. *Nat. Genet.* (2008). doi:10.1038/ng.193

30. Pittman, A. M. *et al.* Linkage disequilibrium fine mapping and haplotype association analysis of the tau gene in progressive supranuclear palsy and corticobasal degeneration. *J. Med. Genet.* (2005). doi:10.1136/jmg.2005.031377

31. Hutton, M. *et al.* Association of missense and 5'-splice-site mutations in tau with the

inherited dementia FTDP-17. *Nature* (1998). doi:10.1038/31508

32. Guelfi, S. *et al.* Regulatory sites for known and novel splicing in human basal ganglia are enriched for disease-relevant information. *bioRxiv* (2019). doi:10.1101/591156

33. Fung, H. C. *et al.* Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol.* (2006). doi:10.1016/S1474-4422(06)70578-6

34. Dong, X. *et al.* Enhancers active in dopamine neurons are a primary link between genetic variation and neuropsychiatric disease. *Nat. Neurosci.* (2018). doi:10.1038/s41593-018-0223-0

35. Lai, Y. *et al.* Phosphoproteomic screening identifies Rab GTP ases as novel downstream targets of PINK 1 . *EMBO J.* (2015). doi:10.15252/embj.201591593

36. Sheikh, B. N., Guhathakurta, S. & Akhtar, A. The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis . *EMBO Rep.* (2019). doi:10.15252/embr.201847630

37. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* (2016). doi:10.1038/cr.2016.39

38. Valente, E. M. *et al.* Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Sci. (New York, NY)* **304**, 1158–1160 (2004).

39. Kitada, T. *et al.* Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* (1998). doi:10.1038/33416

40. Chu, C. T. Mechanisms of selective autophagy and mitophagy: Implications for neurodegenerative diseases. *Neurobiology of Disease* (2019). doi:10.1016/j.nbd.2018.07.015

41. Pruim, R. J. *et al.* LocusZoom: Regional visualization of genome-wide association scan results. in *Bioinformatics* (2011). doi:10.1093/bioinformatics/btq419

42. Ardley, H. C. *et al.* Inhibition of proteasomal activity causes inclusion formation in neuronal and non-neuronal cells overexpressing Parkin. *Mol. Biol. Cell* **14**, 4541–4556 (2003).

43. Soutar, M. P. M. *et al.* AKT signalling selectively regulates PINK1 mitophagy in SHSY5Y cells and human iPSC-derived neurons. 1–11 (2018). doi:10.1038/s41598-018-26949-6

44. Greene, J. C. *et al.* Mitochondrial pathology and apoptotic muscle degeneration in Drosophila parkin mutants. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4078–4083 (2003).

45. Whitworth, A. J. *et al.* Increased glutathione S-transferase activity rescues dopaminergic neuron loss in a Drosophila model of Parkinson's disease. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8024–8029 (2005).

46. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* (2016).

doi:10.1038/nature19057

FIGURE LEGENDS

Figure 1 - Selection of the genes for the PINK1-dependent mitophagy screening.

A. The heat-map represents increasing evidence for gene prioritization (white, light blue, and dark blue: one, two, and three evidences, respectively). ColB = coloc analysis using Braineac, ColG = coloc analysis using GTEx, WPPINA = weighted protein interaction network; GWAS = genes prioritised in PD-GWAS [3].

B. Genes prioritised by means of multiple prediction techniques, grouped based on the types of evidences.

C. Venn diagram highlighting the three genes prioritised by means of three prediction techniques.

Figure 2 – High content mitophagy screen identifies KAT8 as a modulator of pUb(Ser65) levels.

A. pUb(Ser65) Z-scores of one representative mitophagy screen plate.

B. Overview of the PD GWAS genetic signal at the *KAT8* locus.

Figure 3 - KAT8 knockdown decreases pUb(Ser65).

A. Representative images of pUb(Ser65) following treatment of SCR, PINK1 and KAT8 siRNA KD POE SH-SY5Y with 1 µM O/A for 3 h. Insets show the nuclei for the same fields. Scale bar: 20 µm.

B. Quantification of pUb(Ser65) levels in A (n=3, two-way ANOVA with Dunnett's correction).

C. Representative IB of mitochondrial fractions from SCR, PINK1 and KAT8 KD POE SH-SY5Y treated with 1 µM O/A for 1.5 or 3 h.

D. Quantification of pUb(Ser65) in C (n=5, one-way ANOVA with Dunnett's correction).

E. pUb(Ser65) Z-scores of one representative KAT screen plate. See Supplementary Table 4 for the complete list of the genes screened.

Data are shown as mean ± SD.

Figure 4 – Knockdown of the mitochondrial components of the NSL complex reduces pUb(Ser65) levels.

A. Schematic representation of the NSL complex.

B, C. Quantification of pUb(Ser65) following treatment of SCR, PINK1 or NSL components siRNA KD POE SH-SY5Y cells with 1 µM O/A for 1.5 h (B) or 3 h (C). Data are shown as mean ± SD; n=6, one-way ANOVA with Dunnett's correction.

Figure 5 - KANSL1 knockdown decreases pUb(Ser65).

A. Representative images of pUb(Ser65) following treatment of SCR, PINK1 and KANSL1 KD POE SH-SY5Y cells with 1 µM O/A for 3 h. Insets show the nuclei for the same fields. Scale bar: 20 µm.

B. Quantification of pUb(Ser65) in A (n=3, two-way ANOVA with Dunnett's correction).

C. Representative IB of mitochondrial fractions from SCR, PINK1 and KANSL1 KD POE SH-SY5Y treated with 1 µM O/A for 1.5 or 3 h.

D. Quantification of pUb(Ser65) in C (n=5, one-way ANOVA with Dunnett's correction).

Data are shown as mean ± SD.

Figure 6 - KANSL1 and KAT8 knockdown decrease pUb(Ser65).

A. Representative images of pUb(Ser65) (green) following treatment of SCR, PINK1, KAT8 and KANSL1 KD POE SH-SY5Y cells with 1 µM O/A for 0-7 h. Insets show the nuclei for the same fields. Scale bar: 20 µm.

B. Quantification of pUb(Ser65) in A (n=6, two-way ANOVA with Dunnett's correction).

Data are shown as mean ± SD.

Figure 7 - KANSL1 and KAT8 knockdown decrease mitochondrial clearance.

A. Representative images of PMPCB (green) and nuclei (blue) following treatment of SCR, PINK1, KAT8 and KANSL1 KD POE SH-SY5Y cells with 1 µM O/A for 0, 3, 6, 9 h. Scale bar: 10 µm.

B. Quantification of the number of PMPCB spots in A (n=5, two-way ANOVA with Dunnett's correction).

Data are shown as mean ± SD.

Figure 8 - ASE sites in KANSL1 in LD with the H1/H2 SNP.

A. ASEs derived from putamen and substantia nigra in high linkage disequilibrium with the H1/H2 tagging SNP, rs12185268 and their position along the *KANSL1* gene. The missense variants track displays the variants annotated as missense by gnomAD v2.1.1 [46]. The valid track displays the heterozygous sites (orange = missense) with an average read depth greater than 15 reads across all samples, which were examined for ASE. The topmost track displays the FDR-corrected minimum -log10 p-value across samples for the sites that show an ASE in at least one sample.

B. Conservation of the KANSL1 protein across species. The four coding variants in the *KANSL1* gene are in high LD (r2 >0.8) with the H1/H2 haplotypes.

Figure 9. High content mitophagy screening of the ORFs on the 17q21 locus identifies only KANSL1 as a modulator of pUb(Ser65) levels.

pUb(Ser65) Z-scores of one representative 17q21 locus screen plate. See Supplementary Table 8 for the complete list of the genes screened.

EXTENDED FIGURE LEGENDS

Extended Data Figure 1. High Content siRNA Screen for modulators of pUb(Ser65).

A. Workflow of the high content screen for O/A-induced pUb(Ser65) levels.

B. Fold decrease in TOM20 levels following 1.5 and 3 h treatment with 0.1, 1 and 10 µM O/A, compared to DMSO control.

C. Representative images of TOM20 and pUb(Ser65) following 3 h treatment of SCR KD POE SH-SY5Y cells with 10 µM O/A. Scale bar: 20 µm.

D. Quantification of the co-localization in **C** as % of TOM20-positive pUb(Ser65) spots. Graph shows all replicates of non-transfected, SCR, PINK1 and PLK1 KD for 3 independent experiments.

E. Representative images of pUb(Ser65) following treatment of SCR and PINK1 KD POE SH-SY5Y cells with 10 µM O/A for 3 h. Scale bar: 20 µm.

F. Quantification of pUb(Ser65) in **E** (n=6, two-way ANOVA with Tukey's multiple comparisons test).

G. Representative analysis of integrated intensity of pUb(Ser65)and TOM20 for a single HCS plate.

H. pUb(Ser65) Z-scores of the two other replicate screen plates.

Data are shown as mean ± SD.

Extended Data Figure 2. Image processing workflow of the high content screen for O/A induced pUb(Ser65).

Images were loaded as maximum projections. The Hoechst 33342 channel was used to find the nuclei and border nuclei were excluded. The cytoplasm was then found on the calculated Hoechst + TOM20-568 image. The pUb(Ser65)-488 spots were identified within the whole cell and their intensity measured. The outputs of the analysis were the number of nuclei selected and the mean integrated pUb(Ser65) intensity, calculated as the area of the cell covered by pUb(Ser65) spots x corrected intensity of the spots. Analysis parameters for each building block of the Columbus workflow are detailed in the boxes.

Extended Data Figure 3. KAT8 knockdown has no effect on cell viability.

A. Representative images of nuclei following treatment of SCR, PINK1 and PLK1 siRNA KD POE SH-SY5Y cells with 10 µM O/A for 3 h. Scale bar: 20 µm.

B. Quantification of the number of nuclei in A (n=6, two-way ANOVA with Tukey's multiple comparisons test).

C. Z-scores of a representative screen plate showing that KAT8 or PINK1 siRNA KD don't affect cell viability, on the contrary to PLK-1 KD.

Data are shown as mean ± SD.

Extended Data Figure 4. KAT8 eQTLs colocalise with SNPs associated with PD risk

The x-axis displays the physical position on chromosome 16 in megabases. The minus log p-values are plotted for every SNP present in both the PD GWAS [3] and *KAT8* eQTLs derived from the GTEx V7 caudate data. The p-values for the PD GWAS are plotted in yellow and p-values for *KAT8* eQTLs are plotted in blue.

Extended Data Figure 5. KAT8 knockdown decreases pUb(Ser65) levels.

A. Representative images of pUb(Ser65) following treatment of SCR, PINK1 and KAT8 siRNA KD POE SH-SY5Y with 10 µM O/A for 3 h. Insets show nuclear staining for the same fields. Scale bar: 20 µm.

B. Quantification of pUb(Ser65) levels in A (n=3, two-way ANOVA with Dunnett's correction).

C. Representative IB of mitochondrial fractions from SCR, PINK1 and KAT8 KD POE SH-SY5Y treated with 10 µM O/A for 1.5 or 3 h.

D, E. Quantification of pUb(Ser65) (**D**) and KAT8 levels (**E**) in C (n=5, one-way ANOVA with Dunnett's correction).

Data are shown as mean ± SD.

Extended Data Figure 6. Neuronal loss of *mof* or *nsl1* causes locomotor deficit, shortened lifespan and neurodegeneration.

A, B. Climbing ability of pan-neuronal (*nSyb-GAL4*) driven knockdown of *mof* (**A**) or *nsl1* (**B**) measured at the indicated age of adults, compared to control RNAi (A: Kruskal-Wallis test, with Dunn's post-hoc multiple comparisons; B: Mann-Whitney test).

C, D. Lifespan of ***mof* (C) or *nsl1* (D)** pan-neuronal knockdown (*nSyb-GAL4*) compared to control RNAi (Log-rank (Mantel-Cox) test).

E, F. Quantification of dopaminergic neurons (PPL1 cluster) after pan-neuronal or dopaminergic (DA) neuron (*TH-GAL4*) driven depletion of *mof* (**E**), nsl1 (**F**), or control RNAi. Representative images of PPL1 neurons (as bounded by the box) under depletion conditions are shown. Flies were aged 30 days, except for pan-neuronal *nsl1* kd which are 16-days-old. Scale bar: 20 µm; Mann-Whitney test.

For all tests, n numbers are indicated in the graphs; p<0.0001 = ****; p<0.001 = ***.

Extended Data Figure 7. Overview of the PD GWAS genetic signal at the *MAPT* locus.

A. *MAPT* primary GWAS signal.

B. *MAPT* conditional GWAS signal.

Extended Data Figure 8. ASE sites in *MAPT* in LD with the H1/H2 SNP.

ASEs derived from putamen and substantia nigra that are in LD with the H1/H2 tagging SNP, rs12185268 and their position along the *MAPT* gene. The missense variants track displays the variants annotated as missense by gnomAD v2.1.1 [46]. The valid track displays the heterozygous sites (orange = missense) with an average read depth greater than 15 reads across all samples, in LD with H1/H2, which were examined for ASE. The topmost track displays the –log10 scale for the minimum FDR across samples for the sites that show an ASE in at least one sample.

A

Gene	ColB/G	WPPINA	GWAS
CAB39L	X		
CCNT2			X
CD38	X		X
CTSB	X	X	X
DDRGK1			X
DGKQ			X
GALC	X		X
GBA		X	X
GPNMB	X		X
HSD3B7	X		
INPP5F		X	
KAT8	X	X	X
KLHL7	X		X
LRRK2		X	X
LSM7	X		X
MAPT		X	X
NCKIPSD	X	X	X
NSF		X	
NUCKS1	X		X
NUPL2	X		X
PDLIM2	X		X
PM20D1	X		
RAB7L1	X	X	
SH3GL2		X	X
SLC41A1	X		X
SNCA			X
SPPL2B	X		
STK39			X
VAMP4	X		
WDR6	X		
ZNF646			X

B

Prediction Technique	Genes
ColB/G; GWAS; WPPINA	*CTSB* *KAT8* *NCKIPSD*
ColB/G; GWAS	*CD38* *GALC* *GPNMB* *KLHL7* *LSM7* *NUCKS1* *NUPL2* *PDLIM2* *SLC41A1*
ColB/G; WPPINA	*RAB7L1*
GWAS; WPPINA	*GBA* *LRRK2* *MAPT* *SH3GL2*

C

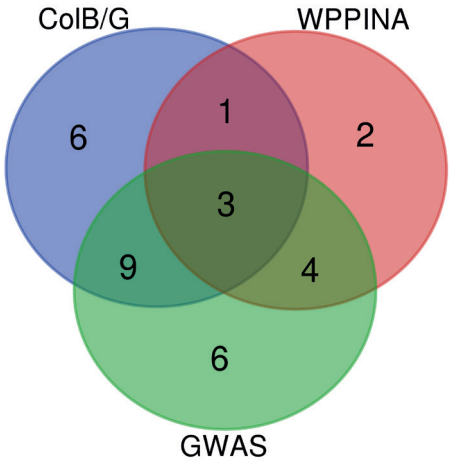


Figure 1 - Selection of the genes for the PINK1-dependent mitophagy screening.

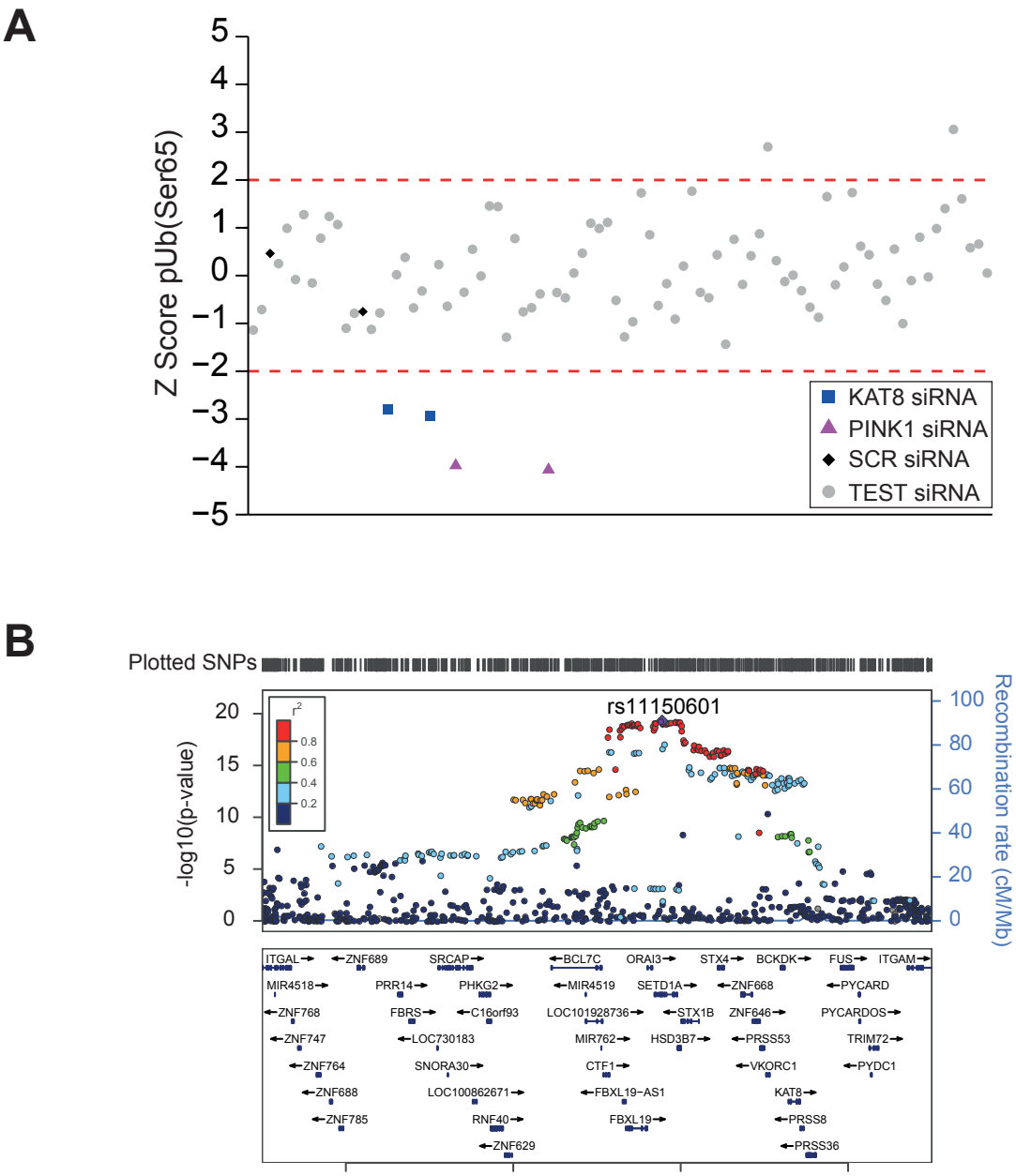


Figure 2 – High content mitophagy screen identifies KAT8 as a modulator of pUb(Ser65) levels.

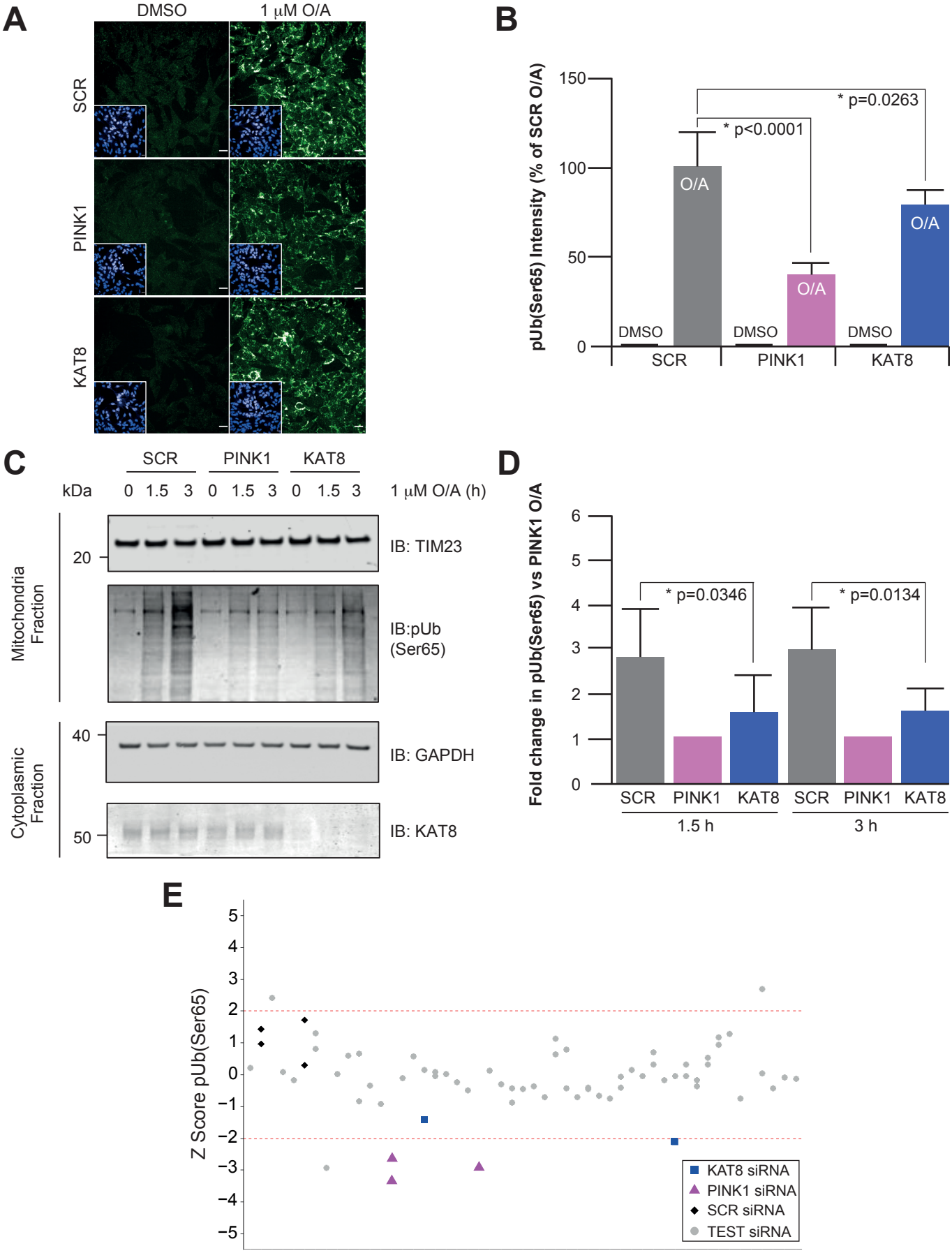


Figure 3 - KAT8 knockdown decreases pUb(Ser65).

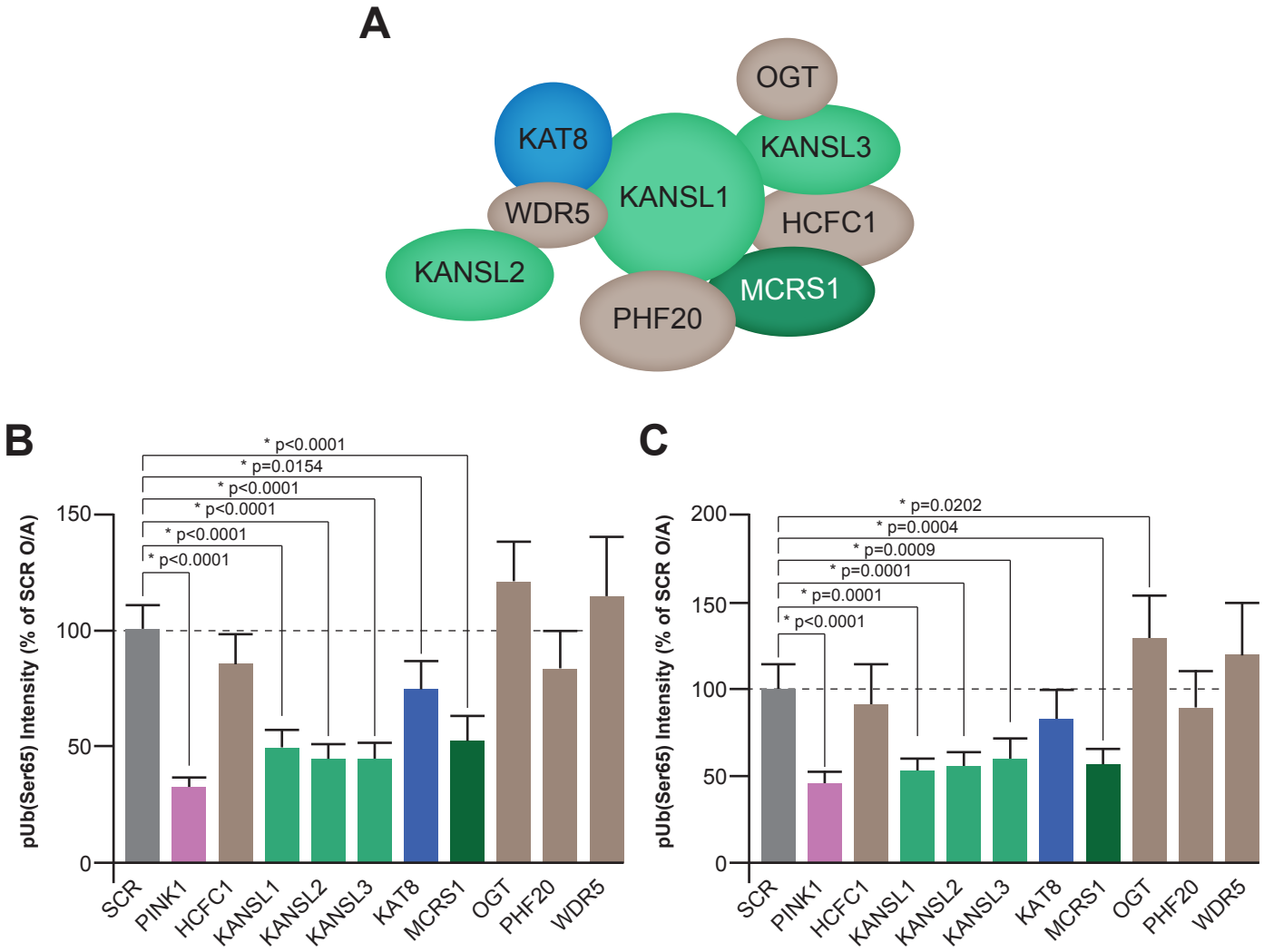


Figure 4 – Knockdown of the mitochondrial components of the NSL complex reduces pUb(Ser65) levels.

.

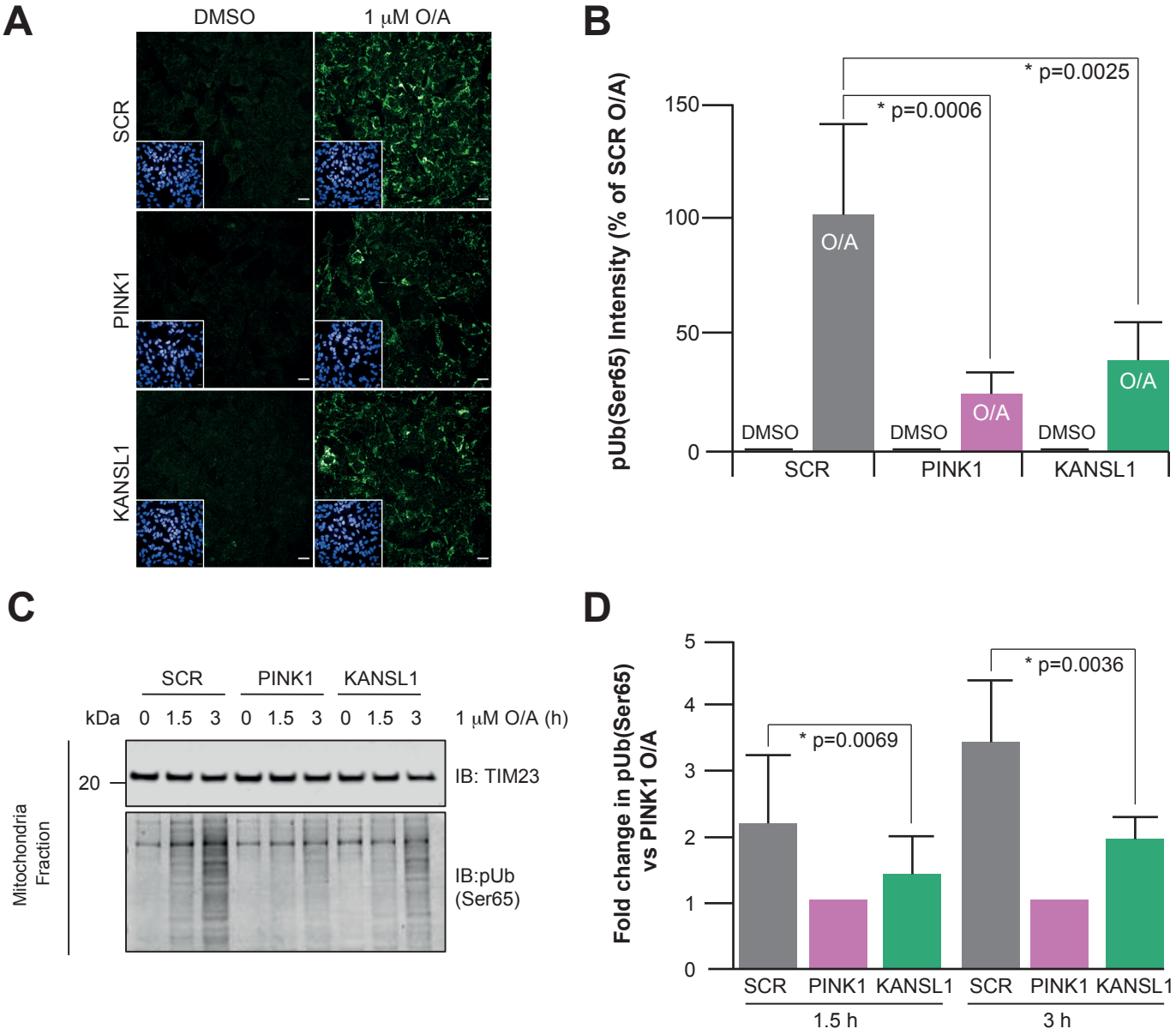


Figure 5 - KANSL1 knockdown decreases pUb(Ser65).

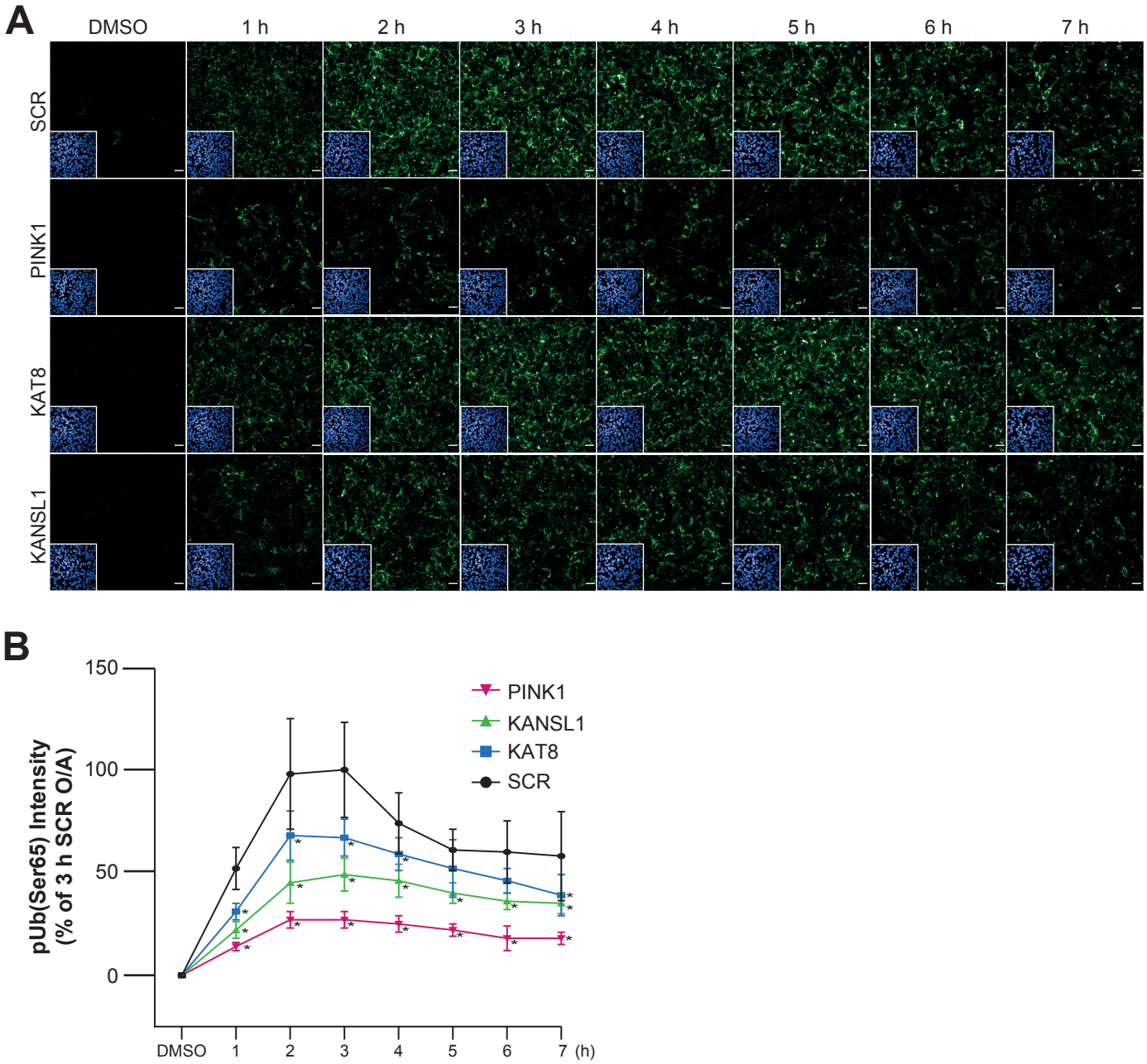


Figure 6 - KANSL1 and KAT8 knockdown decrease pUb(Ser65).

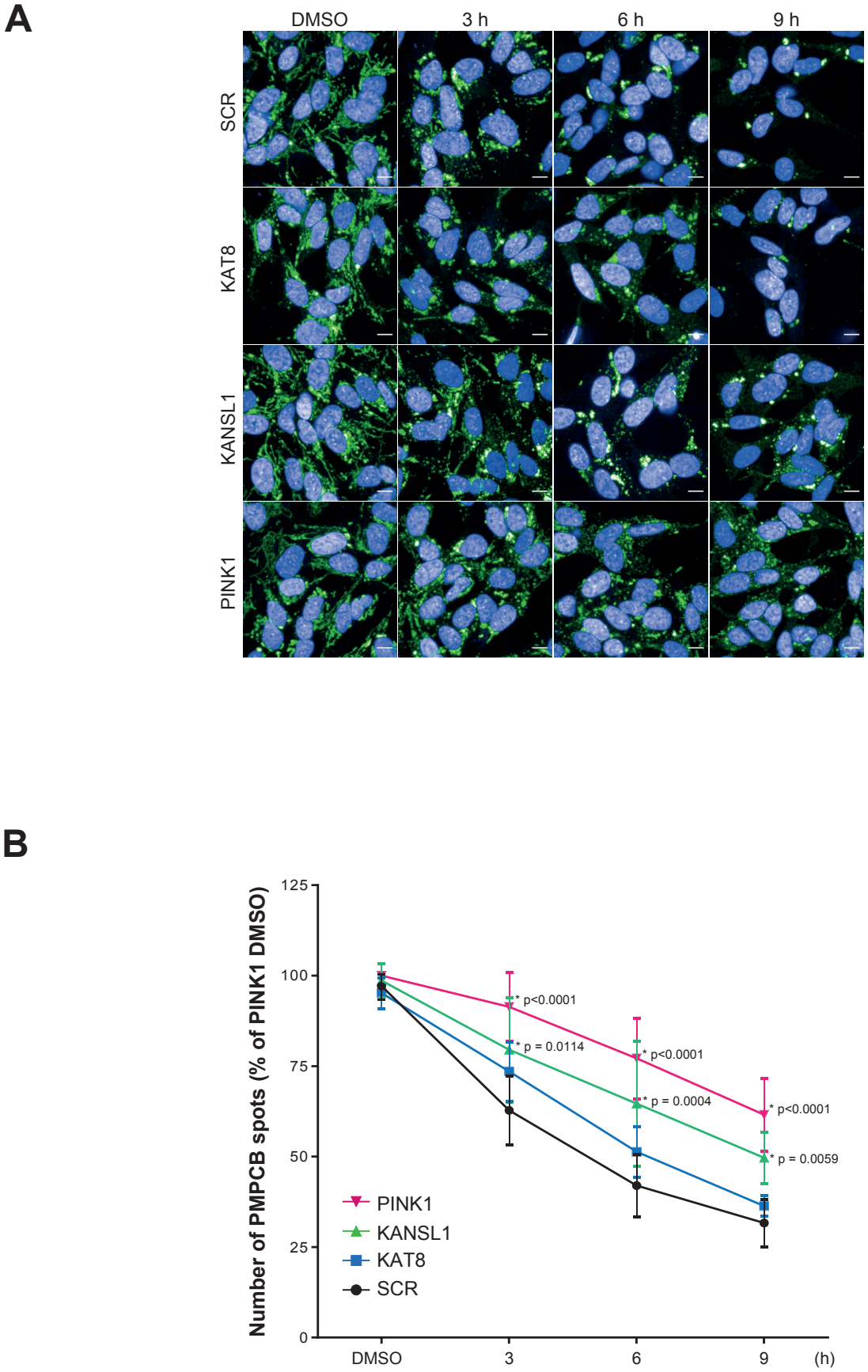
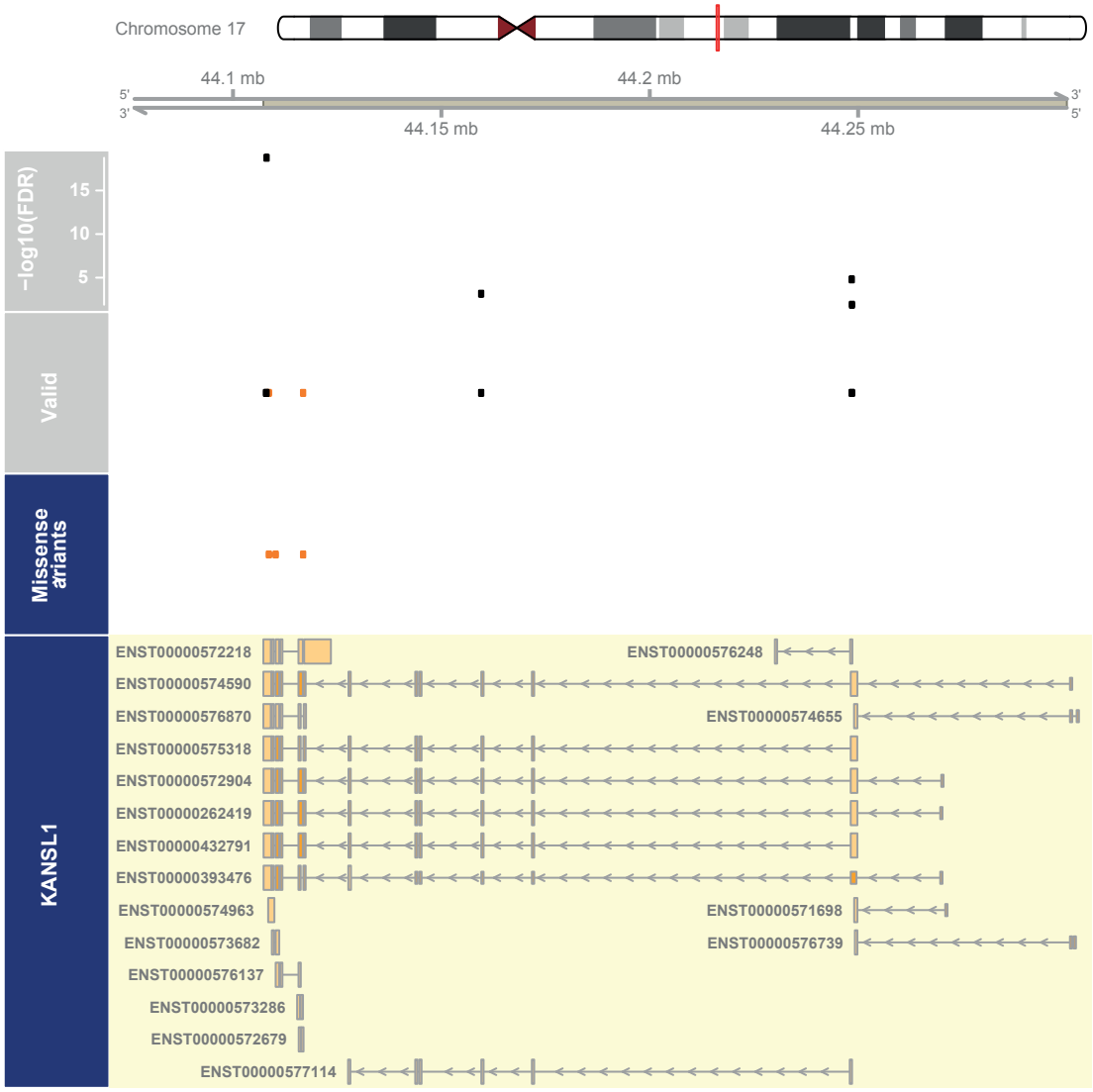
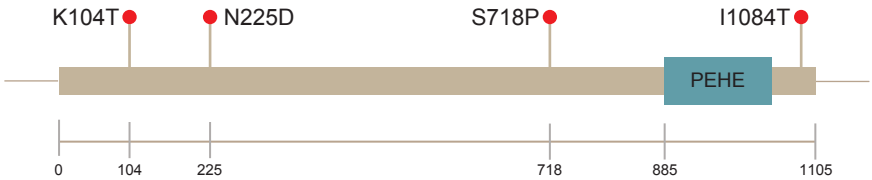


Figure 7 – KANSL1 and KAT8 knockdown decrease mitochondrial clearance.

A



B



	Haplotype	K104T	N225D	S718P	I1084T	*% hap*	
Homo sapiens	H1	K	N	S	I	*0.79*	***PD associated***
	H2	T	D	P	T	*0.19*	
	H2	K	N	P	T	*0.015*	
Pan troglodytes	H2	K	N	P	T		
Pongo pygmaeus	H2	K	N	P	T		
Macaca mulatta	n/a	K	N	P	T		
Mus musculus	n/a	K	N	P	V		

Figure 8 - ASE sites in KANSL1 in LD with the H1/H2 SNP.

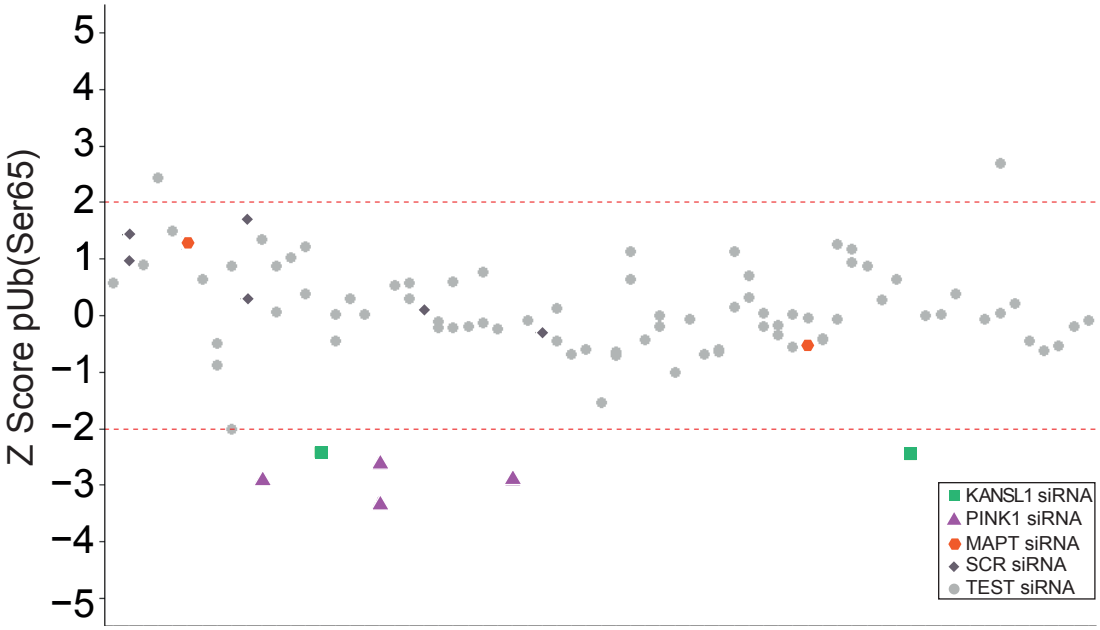
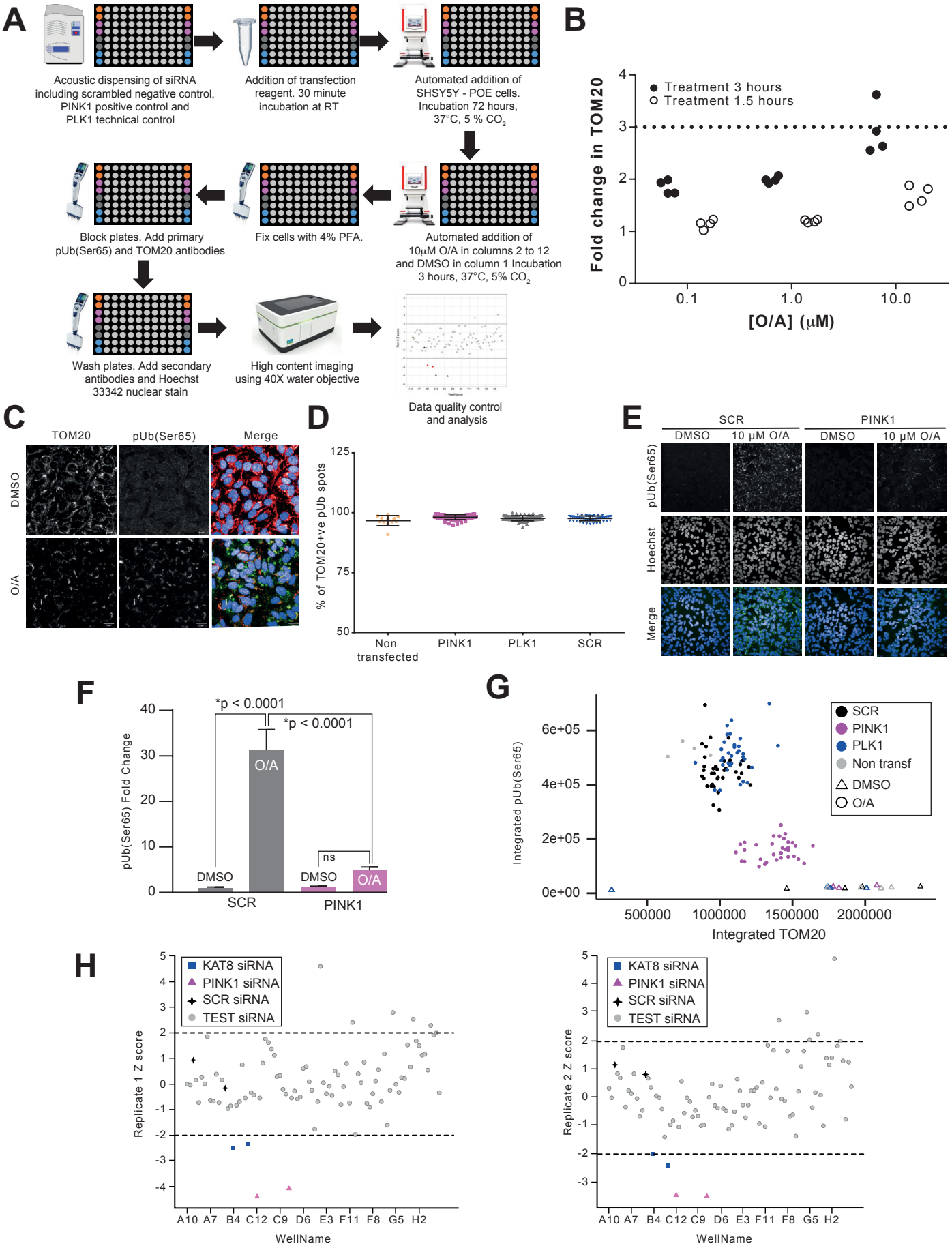
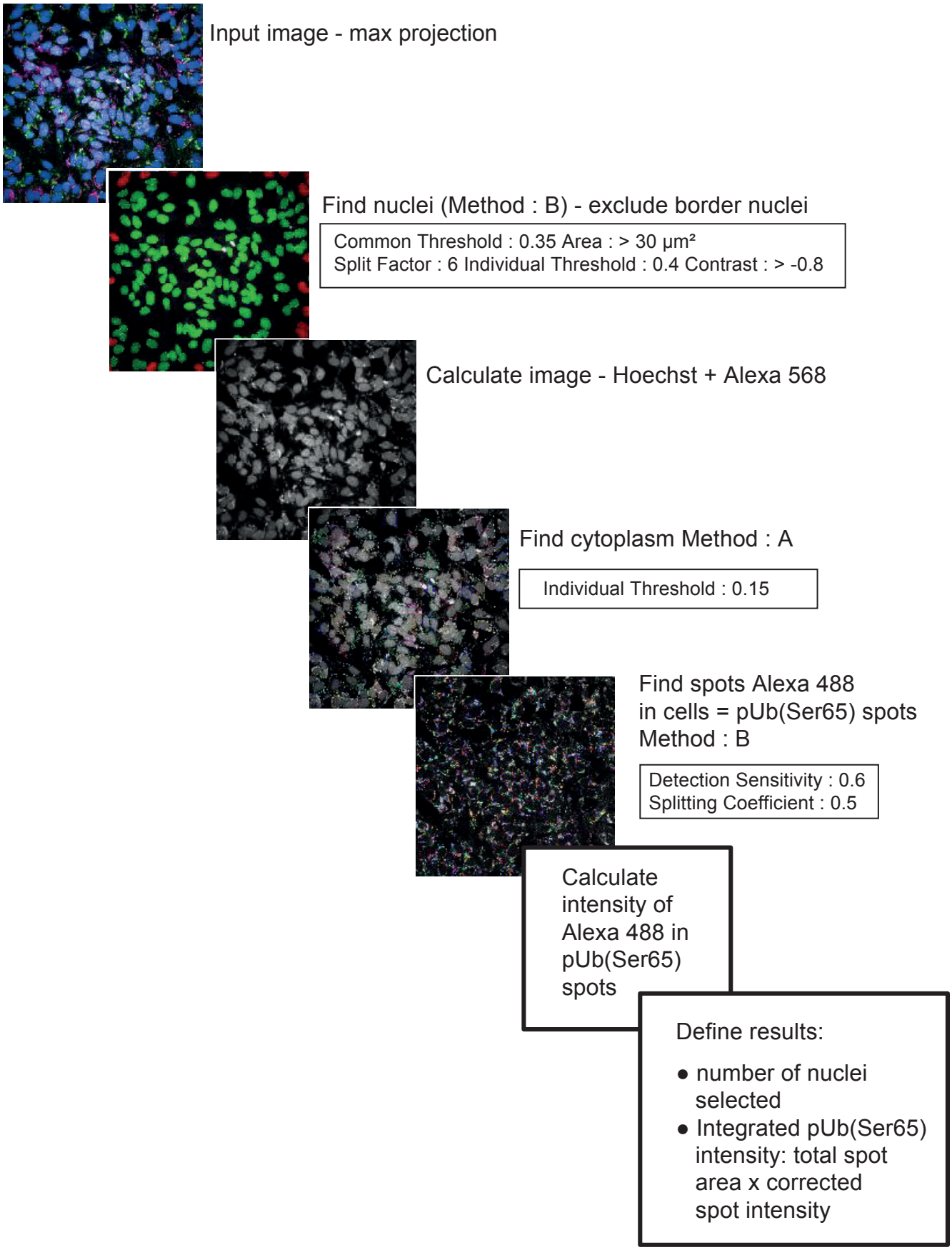


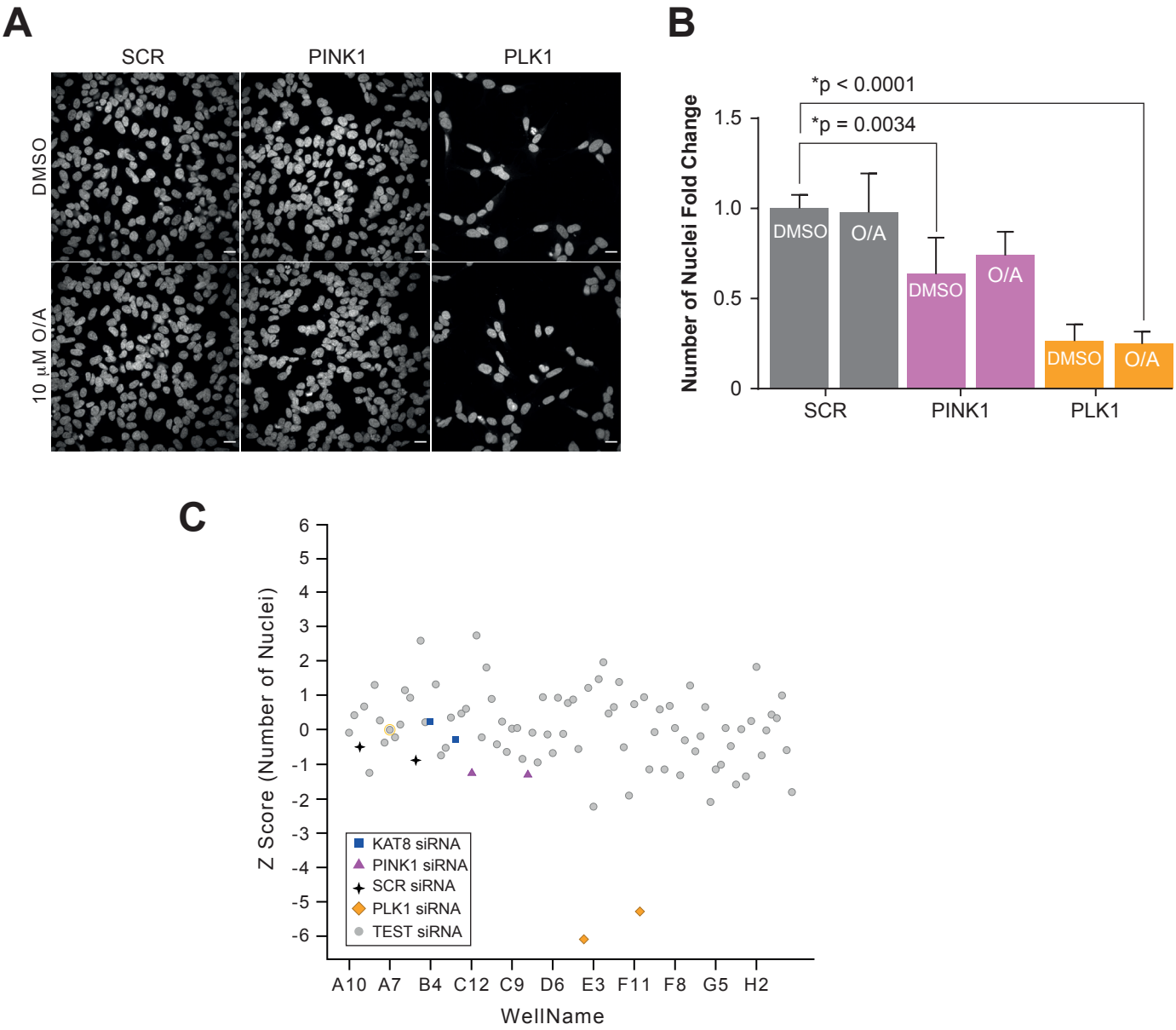
Figure 9 – High content mitophagy screening of the ORFs on the the 17q21 locus identifies only KANSL1 as a modulator of pUb(Ser65) levels.



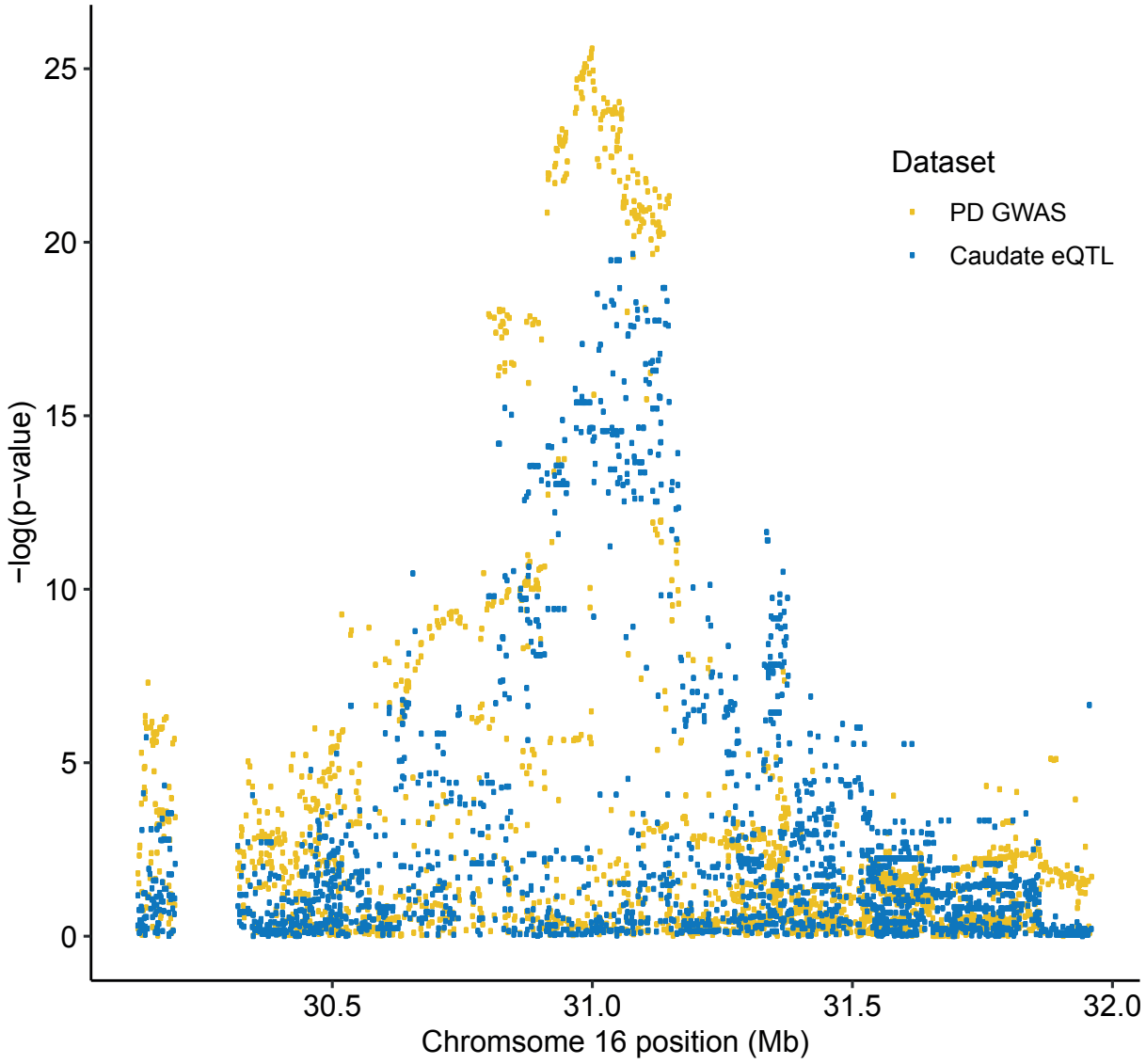
Extended Data Figure 1. High Content siRNA Screen for modulators of pUb(Ser65).



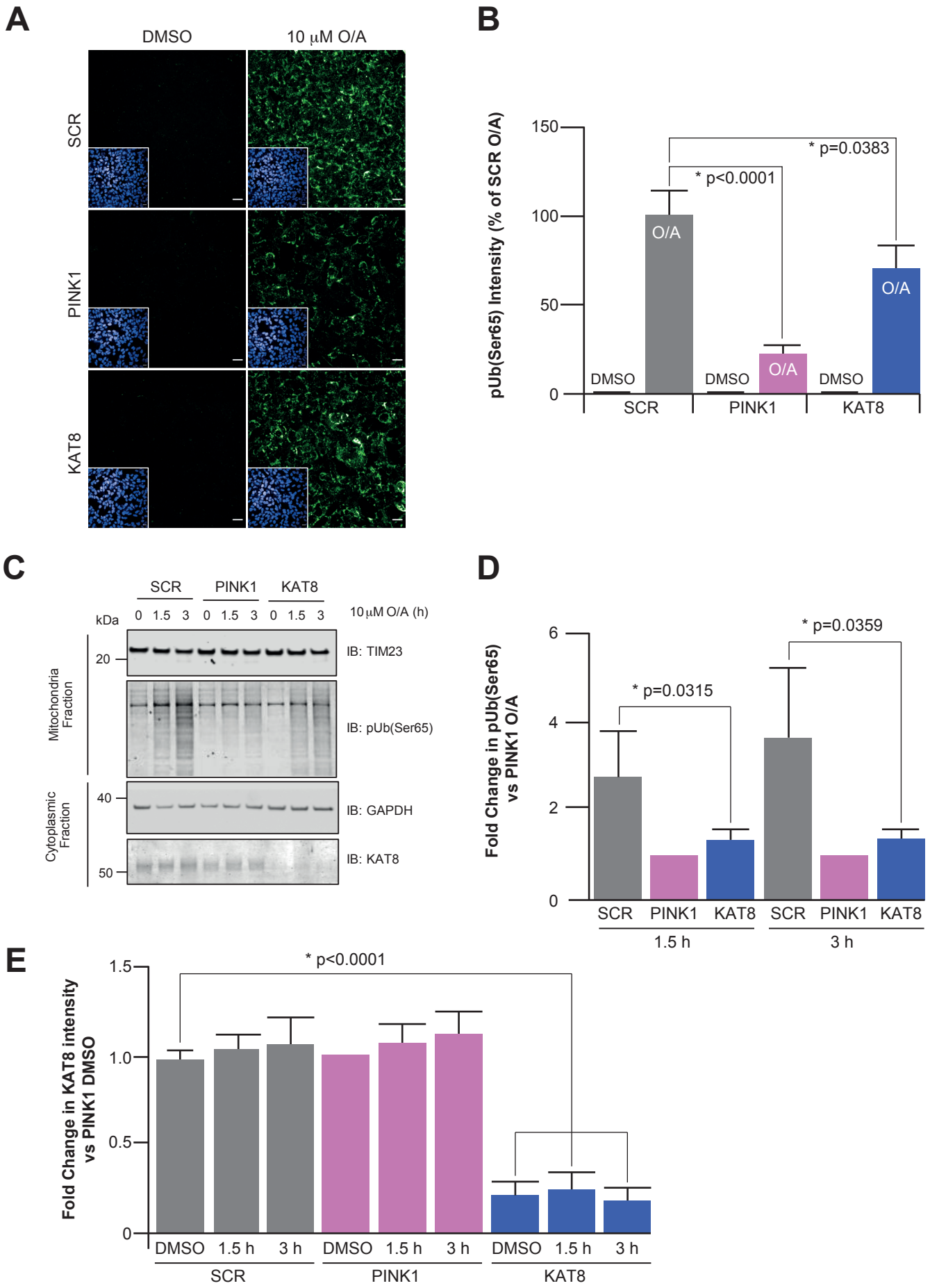
Extended Data Figure 2. Image processing workflow of the high content screen for O/A induced pUb(Ser65).



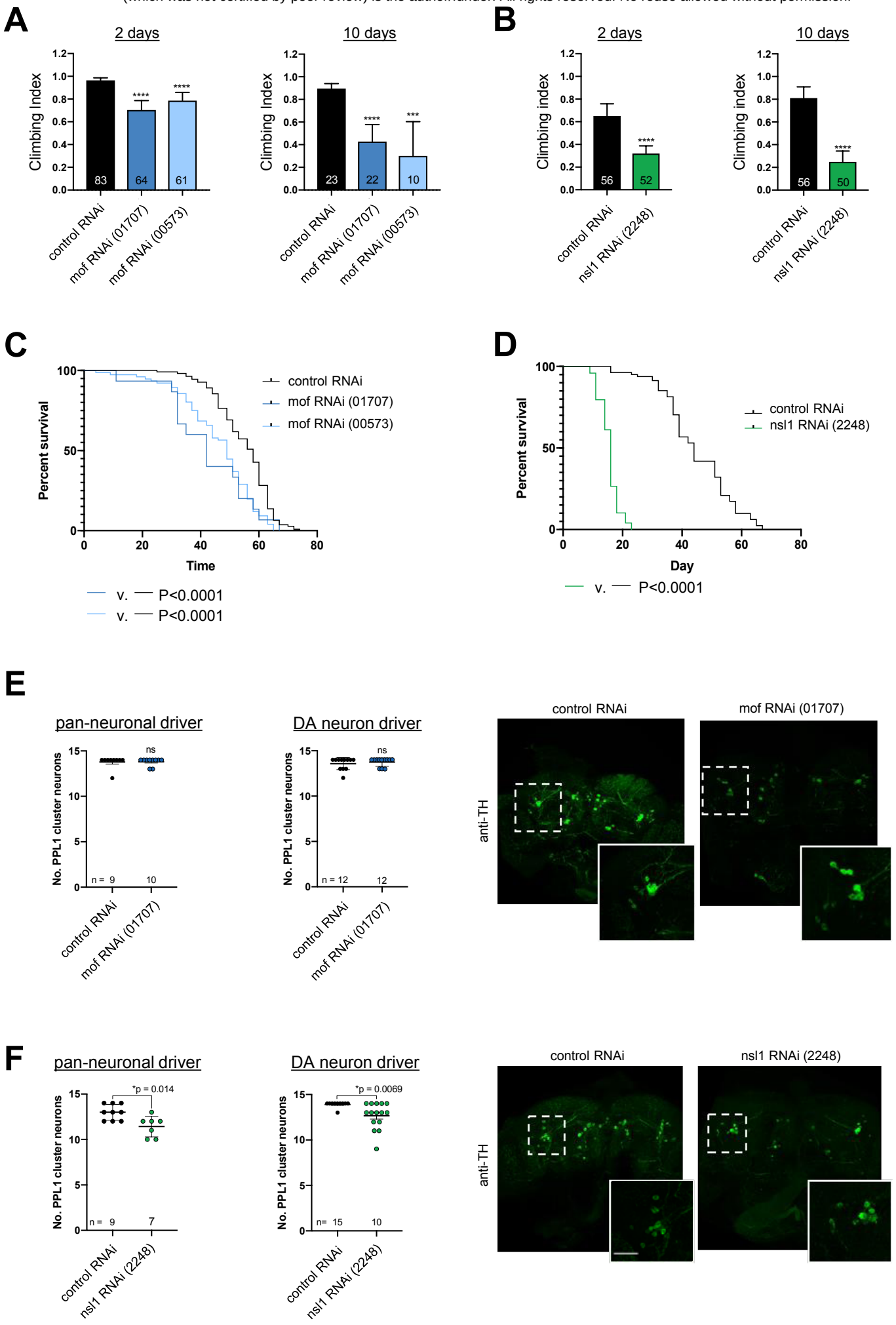
Extended Data Figure 3. KAT8 knockdown has no effect on cell viability.



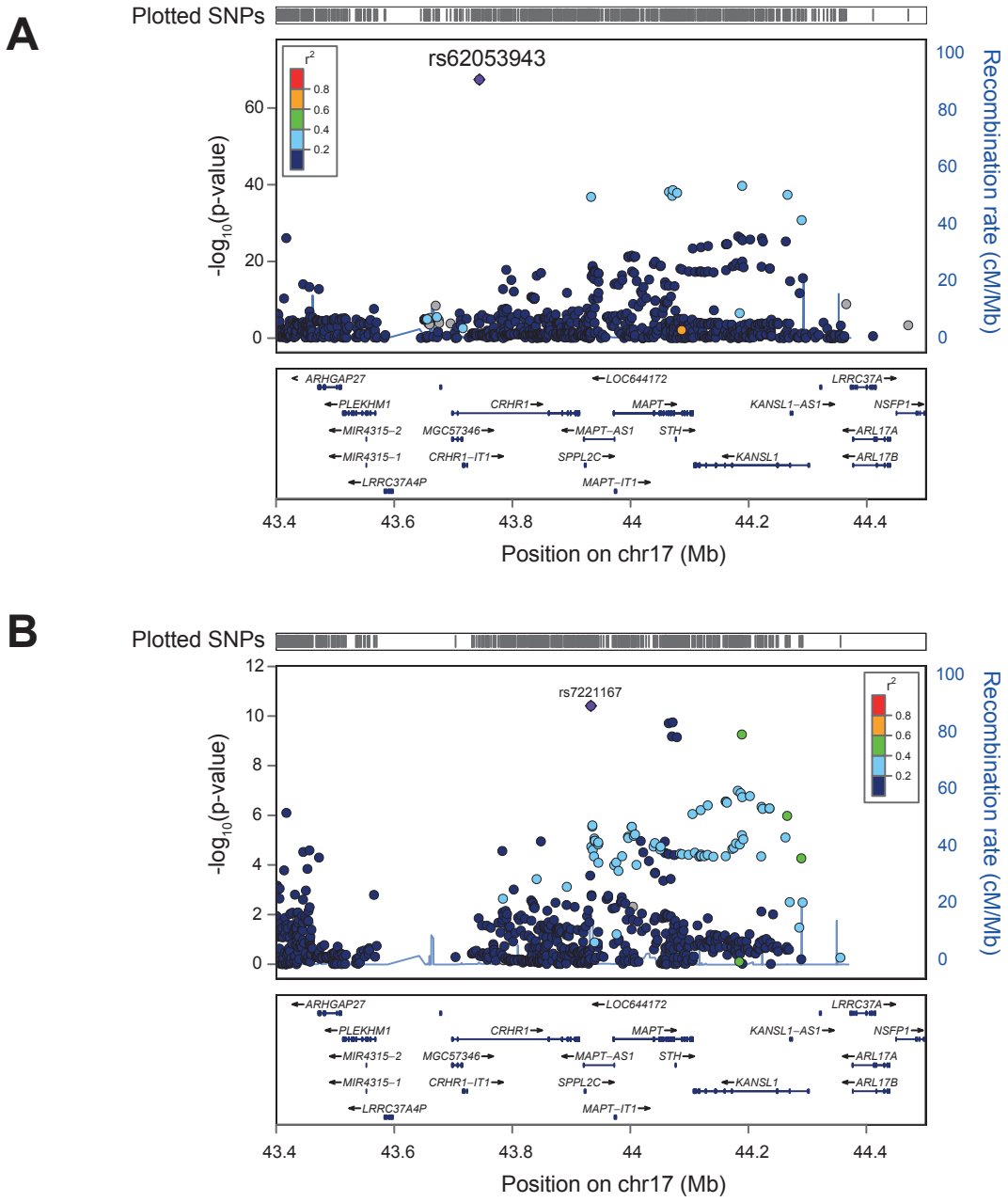
Extended Data Figure 4. KAT8 eQTLs colocalise with SNPs associated with PD risk.



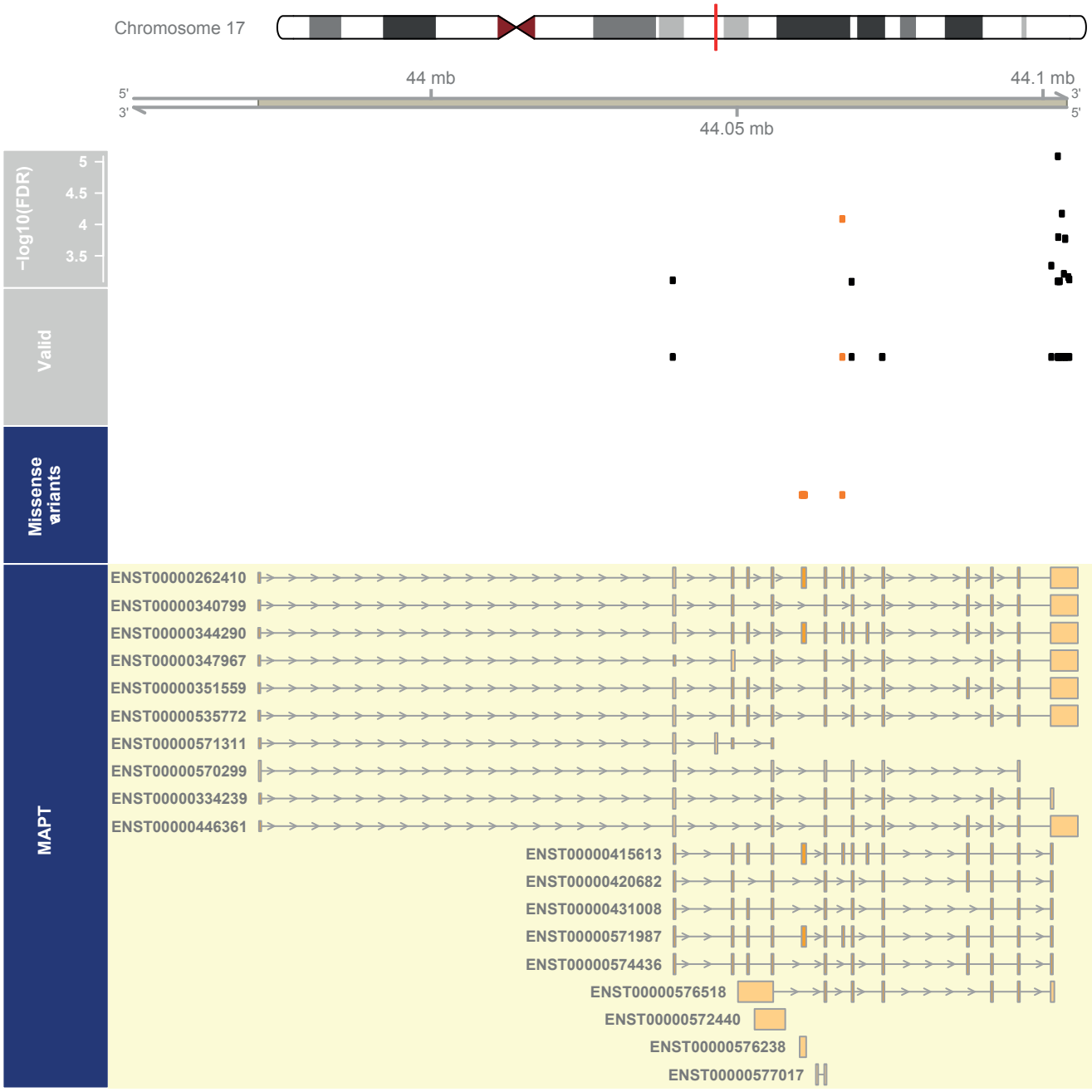
Extended Data Figure 5. KAT8 knockdown decreases pUb(Ser65) levels.



Extended Data Figure 6. Neuronal loss of *mof* or *nsl1* causes locomotor deficit, shortened lifespan and neurodegeneration.



Extended Data Figure 7. Overview of the PD GWAS genetic signal at the *MAPT* locus.



Extended Data Figure 8. ASE sites in *MAPT* in LD with the H1/H2 SNP.

Gene	ColB	ColG	PPI	GWAS	MPD	MLS
ATP13A2					X	X
CCNT2				X		
CD38	X	X		X		
CTSB	X	X	X			
DDRGK1				X		
DGKQ				X		
DJ1					X	
DNAJC13					X	
FBXO7					X	
GALC	X			X		X
GBA			X		X	X
GPNMB	X	X		X		
HSD3B7		X				
IDUA						X
INPP5F			X			
KAT8		X	X	X		
KLHL7		X		X		
LRRK2			X	X	X	
LSM7	X	X		X		
MAPT			X	X		
NCKIPSD	X	X	X	X		
NEK1	X					
NSF			X			
NUCKS1		X		X		
NUPL2	X	X		X		
PDLIM2		X		X		
PM20D1	X					
PRKN					X	
RAB7L1	X	X	X			
SH3GL2			X			
SLC41A1		X		X		
SNCA					X	
SPPL2B	X					
STK39				X		
VAMP4	X	X				
VPS35					X	
WDR6	X	X				
ZNF646				X		

Supplementary Table 1. Complete list of the 38 genes screened in the high content screen. ColB = coloc analysis using Braineac, ColG = coloc analysis using GTEx, WPPINA = weighted protein-protein interaction network; GWAS = genes prioritised in PD-GWAS (Chang *et al*, 2017), MPD = Mendelian genes associated with PD, MLS = Mendelian genes associated with lysosomal storage disorders

eQTL_ dataset	GTEx
sequencing_ method	RNA-seq
gene_ symbol	KAT8
tissue	Brain_Caudate_ basal_ganglia
braineac_ probe_id	NA
nsnps	2499
PPH0	4.89E-10
PPH1	1.76E-07
PPH2	6.76E-04
PPH3	0.242293929
PPH4	0.75702972
PD_top_snp	chr16:31000809
coloc_top_ snp	chr16:31048079
coloc_SNP_ PPH4	0.08637809
coloc_eQTL_ effect_allele	T
coloc_eQTL_ other_allele	C
coloc_eQTL_ beta	-0.358743
coloc_eQTL_ SE	0.0563187
coloc_eQTL_ Freq1	NA
coloc_eQTL_ p_val	3.49E-09
coloc_PD_Al1	C
coloc_PD_Al2	T
coloc_PD_ beta	-0.0739
coloc_PD_ SE	0.0115
coloc_PD_ Freq1	0.5972
coloc_PD_ p_val	1.38E-10

Supplementary Table 2. Results of the Colocalization analysis for KAT8.
PD_top_snp = lead SNP in the PD GWAS, coloc_top snp = Most likely SNP responsible for the colocalization signal, coloc_SNP_PPH4 = posterior probability of coloc top SNP being the true SNP responsible for the colocalization signal.

ID	KAT8	KAT8	KAT8	KAT8	KAT8	KAT8
CHR	16	16	16	16	16	16
P0	31127075	31127075	31127075	31127075	31128984	31127075
P1	31142714	31142714	31142714	31142714	31142714	31142714
HSQ	0.144	0.191	0.121	0.156	0.0641	0.207
BEST.GWAS.ID	rs9938550	rs9938550	rs9938550	rs9938550	rs2305880	rs9938550
BEST.GWAS.Z	-6.82	-6.82	-6.82	-6.82	-6.84	-6.82
EQTL.ID	rs8046707	rs1549293	rs2855475	rs12597511	rs749767	rs4527034
EQTL.R2	0.13718	0.129562	0.227501	0.19827	0.074695	0.00998
EQTL.Z	-4.26	-4.12	-4.58	-4.61	-6.05	-4.23
EQTL.GWAS.Z	6.28455	6.13675	6.1538	6.178	6.06897	6.1724
NSNP	240	240	240	240	248	240
NWGT	3	4	3	1	1	9
MODEL	lasso	lasso	lasso	lasso	lasso	lasso
MODELCV.R2	0.103	0.181407	0.17535	0.17222	0.07213	0.04964
MODELCV.PV	0.000889	1.37E-05	2.91E-05	2.56E-05	3.95E-09	0.0139
TWAS.Z	-6.6932	-6.2068	-6.1876	-6.178	-6.069	-6.04195
TWAS.P	2.18E-11	5.41E-10	6.11E-10	6.49E-10	1.29E-09	1.52E-09
FDR	2.09E-09	4.84E-08	7.15E-08	6.79E-08	4.37E-07	2.36E-07
Region	GTEx.Brain _Cortex	GTEx.Brain_ Nucleus_accumbens _basal_ganglia	GTEx.Brain_ Cerebellar_ Hemisphere	GTEx.Brain_ Frontal_Cortex _BA9	CMC.BRAIN .RNASEQ	GTEx.Brain_ Cerebellum

Supplementary Table 3. Results of the TWAS analysis for *KAT8*.

KAT name	Alternative name(s)
KAT1	HAT1
KAT2A	GCN5
KAT2B	PCAF
KAT3A	CREBBP, CBP
KAT3B	EP300
KAT4	TAF1, TFII250
KAT5	TIP60
KAT6A	MOZ, MYST3
KAT6B	MORF, MYST4
KAT7	HBO1, MYST2
KAT8	MOF, MYST1
KAT9	ELP3
KAT12	GTF3C4, TFIIIC90
KAT13A	NCOA1, SRC1
KAT13B	NCOA3, SRC3, ACTR
KAT13C	NCOA2
KAT13D	CLOCK
	ACAT1
	ATAT1
	ATF2
	BLOC1S1 (GCN5L1)
	NAT10

Supplementary Table 4. Complete list of the 22 KATs screened in the high content screen. (See Fig 3E)

Within each row, compare columns (simple effects within rows)

Number of families	8
Number of comparisons per family	3

Dunnett's multiple comparisons test	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adjusted P Value
DMSO					
SCR vs. KAT8	0	-14.19 to 14.19	No	ns	>0.9999
SCR vs. KANSL1	0	-14.19 to 14.19	No	ns	>0.9999
SCR vs. PINK1	0	-14.19 to 14.19	No	ns	>0.9999
1 h					
SCR vs. KAT8	21	6.807 to 35.19	Yes	**	0.0017
SCR vs. KANSL1	30	15.81 to 44.19	Yes	****	<0.0001
SCR vs. PINK1	38	23.81 to 52.19	Yes	****	<0.0001
2 h					
SCR vs. KAT8	30	15.81 to 44.19	Yes	****	<0.0001
SCR vs. KANSL1	53	38.81 to 67.19	Yes	****	<0.0001
SCR vs. PINK1	71	56.81 to 85.19	Yes	****	<0.0001
3 h					
SCR vs. KAT8	33	18.81 to 47.19	Yes	****	<0.0001
SCR vs. KANSL1	51	36.81 to 65.19	Yes	****	<0.0001
SCR vs. PINK1	73	58.81 to 87.19	Yes	****	<0.0001
4 h					
SCR vs. KAT8	15	0.8065 to 29.19	Yes	*	0.0355
SCR vs. KANSL1	28	13.81 to 42.19	Yes	****	<0.0001
SCR vs. PINK1	49	34.81 to 63.19	Yes	****	<0.0001
5 h					
SCR vs. KAT8	9	-5.193 to 23.19	No	ns	0.307
SCR vs. KANSL1	21	6.807 to 35.19	Yes	**	0.0017
SCR vs. PINK1	39	24.81 to 53.19	Yes	****	<0.0001
6 h					
SCR vs. KAT8	14	-0.1935 to 28.19	No	ns	0.0542
SCR vs. KANSL1	24	9.807 to 38.19	Yes	***	0.0003
SCR vs. PINK1	42	27.81 to 56.19	Yes	****	<0.0001
7 h					
SCR vs. KAT8	19	4.807 to 33.19	Yes	**	0.0051
SCR vs. KANSL1	23	8.807 to 37.19	Yes	***	0.0005
SCR vs. PINK1	40	25.81 to 54.19	Yes	****	<0.0001

Supplementary Table 5. p-values for Figure 6B.

hetSNP	Sample ID	Read count	Allele	Ens.Alt	Individual ID	Tissue
17:44108355	A653_043	47	A	A	004_06	SNIG
17:44108355	A653_043	20	G	A	004_06	SNIG
17:44108355	A653_441	69	A	A	032_09	PUTM
17:44108355	A653_441	0	G	A	032_09	PUTM
17:44159849	A653_031	2	C	C	024_09	PUTM
17:44159849	A653_031	23	T	C	024_09	PUTM
17:44159849	A653_719	58	C	C	035_09	SNIG
17:44159849	A653_719	26	T	C	035_09	SNIG
17:44248769	A653_093	36	C	C	004_08	SNIG
17:44248769	A653_093	4	T	C	004_08	SNIG
17:44248769	A653_184	30	C	C	013_09	PUTM
17:44248769	A653_184	7	T	C	013_09	PUTM
17:44248769	A653_326	26	C	C	017_09	PUTM
17:44248769	A653_326	3	T	C	017_09	PUTM
17:44248769	A653_617	22	C	C	029_09	SNIG
17:44248769	A653_617	0	T	C	029_09	SNIG
17:44248769	A653_679	23	C	C	030_06	PUTM
17:44248769	A653_679	3	T	C	030_06	PUTM
17:44248769	A653_753	43	C	C	029_09	PUTM
17:44248769	A653_753	10	T	C	029_09	PUTM
17:44248769	A653_794	24	C	C	036_09	SNIG
17:44248769	A653_794	4	T	C	036_09	SNIG
17:44248814	A653_043	10	A	A	004_06	SNIG
17:44248814	A653_043	0	G	A	004_06	SNIG
17:44248814	A653_326	19	A	A	017_09	PUTM
17:44248814	A653_326	3	G	A	017_09	PUTM
17:44248814	A653_617	12	A	A	029_09	SNIG
17:44248814	A653_617	0	G	A	029_09	SNIG
17:44248814	A653_679	19	A	A	030_06	PUTM
17:44248814	A653_679	2	G	A	030_06	PUTM
17:44248814	A653_753	25	A	A	029_09	PUTM
17:44248814	A653_753	5	G	A	029_09	PUTM
17:44248814	A653_950	28	A	A	006_10	SNIG
17:44248814	A653_950	8	G	A	006_10	SNIG

Supplementary Table 6. Summary of the results of ASE analysis across the *KANSL1* gene.

Column name	**Description**
hetSNP	Heterozygous SNP with average read depth >15 across samples
Avg.Reads.all.samples	Average read depth across all samples
min.FDR	Minimum false discovery rate across the samples
Allele1	Allele 1 for the hetSNP
Allele2	Allele 2 for the hetSNP
ens.alt	Ensembl alternate allele
symbol	Gene name obtained using the variant effect predictor tool (VEP)
most.severe.consequence	Most severe of all observed consequence types reported for the hetSNP (VE
ASE	Indicates ASE = 'Y' if min.FDR < 0.05, 'N' if min.FDR >=0.05

The following column details were from obtained from LDProxy (https://ldlink.nci.nih.gov/)

RSID	Identifier for the hetSNP
Alleles	SNP alleles
MAF	Minor allele frequency
Dprime	Indicator of allelic segregation for two genetic variants
R2	Measure of correlation of alleles for two genetic variants.

hetSNP	Sample ID	Read count	Allele	Ens.alt	Individual ID	Tissue
17:44067400	A653_031	18	C	C	024_09	PUTM
17:44067400	A653_031	3	T	C	024_09	PUTM
17:44102689	A653_031	5	C	C	024_09	PUTM
17:44102689	A653_031	21	G	C	024_09	PUTM
17:44103616	A653_031	23	C	T	024_09	PUTM
17:44103616	A653_031	0	T	T	024_09	PUTM
17:44103825	A653_031	11	C	C	024_09	PUTM
17:44103825	A653_031	31	T	C	024_09	PUTM
17:44104509	A653_043	43	C	C	004_06	SNIG
17:44104509	A653_043	82	T	C	004_06	SNIG
17:44102604	A653_056	85	C	C	015_07	PUTM
17:44102604	A653_056	138	T	C	015_07	PUTM
17:44104343	A653_093	84	A	C	004_08	SNIG
17:44104343	A653_093	34	C	C	004_08	SNIG
17:44102689	A653_171	37	C	C	040_08	SNIG
17:44102689	A653_171	87	G	C	040_08	SNIG
17:44039691	A653_177	118	A	G	015_07	SNIG
17:44039691	A653_177	184	G	G	015_07	SNIG
17:44104343	A653_177	35	C	C	015_07	SNIG
17:44104343	A653_177	74	A	C	015_07	SNIG
17:44102689	A653_205	30	C	C	040_08	PUTM
17:44102689	A653_205	64	G	C	040_08	PUTM
17:44102689	A653_225	36	C	C	004_08	PUTM
17:44102689	A653_225	70	G	C	004_08	PUTM
17:44104343	A653_243	37	C	C	021_09	SNIG
17:44104343	A653_243	71	A	C	021_09	SNIG
17:44102689	A653_283	47	C	C	017_09	SNIG
17:44102689	A653_283	90	G	C	017_09	SNIG
17:44102865	A653_283	79	A	C	017_09	SNIG
17:44102865	A653_283	41	C	C	017_09	SNIG
17:44104343	A653_283	42	C	C	017_09	SNIG
17:44104343	A653_283	84	A	C	017_09	SNIG
17:44104509	A653_283	85	T	C	017_09	SNIG
17:44104509	A653_283	43	C	C	017_09	SNIG
17:44067400	A653_288	50	C	C	038_08	PUTM
17:44067400	A653_288	12	T	C	038_08	PUTM
17:44104509	A653_326	73	T	C	017_09	PUTM
17:44104509	A653_326	33	C	C	017_09	PUTM
17:44101563	A653_441	34	C	C	032_09	PUTM

hetSNP	Sample ID	Read count	Allele	Ens.alt	Individual ID	Tissue
17:44101563	A653_441	0	T	C	032_09	PUTM
17:44102638	A653_441	0	A	G	032_09	PUTM
17:44102638	A653_441	197	G	G	032_09	PUTM
17:44102689	A653_441	77	C	C	032_09	PUTM
17:44102689	A653_441	0	G	C	032_09	PUTM
17:44103296	A653_441	112	C	C	032_09	PUTM
17:44103296	A653_441	0	T	C	032_09	PUTM
17:44103825	A653_441	75	C	C	032_09	PUTM
17:44103825	A653_441	0	T	C	032_09	PUTM
17:44103826	A653_441	74	A	A	032_09	PUTM
17:44103826	A653_441	0	G	A	032_09	PUTM
17:44067400	A653_627	9	T	C	033_09	PUTM
17:44067400	A653_627	30	C	C	033_09	PUTM
17:44068924	A653_627	39	A	A	033_09	PUTM
17:44068924	A653_627	72	G	A	033_09	PUTM
17:44102933	A653_627	73	C	C	033_09	PUTM
17:44102933	A653_627	118	T	C	033_09	PUTM
17:44104343	A653_627	68	A	C	033_09	PUTM
17:44104343	A653_627	35	C	C	033_09	PUTM
17:44104509	A653_627	92	T	C	033_09	PUTM
17:44104509	A653_627	53	C	C	033_09	PUTM
17:44102689	A653_643	24	C	C	017_08	PUTM
17:44102689	A653_643	54	G	C	017_08	PUTM
17:44104509	A653_679	11	C	C	030_06	PUTM
17:44104509	A653_679	40	T	C	030_06	PUTM
17:44067400	A653_719	13	T	C	035_09	SNIG
17:44067400	A653_719	51	C	C	035_09	SNIG
17:44102689	A653_719	45	C	C	035_09	SNIG
17:44102689	A653_719	84	G	C	035_09	SNIG
17:44067400	A653_738	23	C	C	036_09	PUTM
17:44067400	A653_738	3	T	C	036_09	PUTM
17:44067400	A653_753	24	C	C	029_09	PUTM
17:44067400	A653_753	6	T	C	029_09	PUTM
17:44101563	A653_763	0	C	C	034_08	PUTM
17:44101563	A653_763	33	T	C	034_08	PUTM
17:44103296	A653_763	0	C	C	034_08	PUTM
17:44103296	A653_763	91	T	C	034_08	PUTM
17:44068924	A653_794	39	A	A	036_09	SNIG
17:44068924	A653_794	73	G	A	036_09	SNIG

Column name	Description
hetSNP	Heterozygous SNP with average read depth >15 across samples
Avg.Reads.all.samples	Average read depth across all samples
min.FDR	Minimum false discovery rate across the samples
Allele1	Allele 1 for the hetSNP
Allele2	Allele 2 for the hetSNP
ens.alt	Ensembl alternate allele
symbol	Gene name obtained using the variant effect predictor tool (VEP)
most.severe.consequen	Most severe of all observed consequence types reported for the hetSNP (VEP)
ASE	Indicates ASE = 'Y' if min.FDR < 0.05, 'N' if min.FDR >=0.05

The following column details were from obtained from LDProxy (https://ldlink.nci.nih.gov/)

RSID	Identifier for the hetSNP
Alleles	SNP alleles
MAF	Minor allele frequency
Dprime	Indicator of allelic segregation for two genetic variants
R2	Measure of correlation of alleles for two genetic variants.

17q21 ORFs	
ACBD4	HEXIM1
ADAM11	HIGD1B
ARHGAP27	KANSL1
ARL17A	KIF18B
ARL17B	LRRC37A
C1QL1	LRRC37A2
CCDC103	MAP3K14
CDC27	MAPT
CRHR1	MYL4
DBF4B	NMT1
DCAKD	NSF
EFTUD2	PLEKHM1
FMNL1	RPRML
GFAP	SPPL2C
GJC1	STH
GOSR2	WNT3

Supplementary Table 8. Complete list of the 32 ORFs in the 17q21 locus screened in the high content screen. (See Fig 9)