

Identification of candidate genes for developmental colour agnosia in a single unique family

Tanja C. Nijboer^{1,2#}, Ellen V.S. Hessel^{3#}, Gijs W. van Haaften^{3#}, Martine J. van Zandvoort², Peter J. van der Spek⁵, Christine Troelstra⁵, Carolien de Kovel³, Bobby P.C. Koeleman³, Bert van der Zwaag³, Eva H. Brilstra³, and J. Peter H. Burbach⁴.

¹UMCU Brain Center and Center of Excellence for Rehabilitation Medicine, University Medical Center Utrecht and De Hoogstraat Rehabilitation, 3584 CX Utrecht, The Netherlands

²Department of Experimental Psychology and Helmholtz Institute, Utrecht University, 3584 CS Utrecht, The Netherlands.

³Department of Biomedical Genetics, University Medical Center Utrecht, 3584 CG Utrecht, The Netherlands.

⁴UMCU Brain Center, Department of Translational Neuroscience, University Medical Center Utrecht, 3584 CG Utrecht, the Netherlands.

⁵Department of Pathology, Erasmus Medical Center Rotterdam, 3015 GD Rotterdam, the Netherlands

These authors contributed equally

Corresponding author: j.p.h.burbach@umcutrecht.nl

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Abstract:

Colour agnosia is a disorder that impairs colour knowledge (naming, recognition) despite intact colour perception. Previously, we have identified the first and only-known family with hereditary developmental colour agnosia. The aim of the current study was to explore genomic regions and candidate genes that potentially cause this trait in this family. For three family members with developmental colour agnosia and three unaffected family members CGH-array analysis and exome sequencing was performed, and linkage analysis was carried out using DominantMapper, resulting in the identification of 19 cosegregating chromosomal regions. Whole exome sequencing resulted in 11 rare coding variants present in all affected family members with developmental colour agnosia and absent in unaffected members. These variants affected genes that have been implicated in neural processes and functions (*CACNA2D4*, *DDX25*, *GRINA*, *MYO15A*), that have a indirect link to brain function or development (*MAML2*, *STAU1*, *TMED3*), and a remaining group lacking brain expression or involved in non-neural traits (*DEPDC7*, *OR1J1*, *OR8D4*, *RABEPK*). Although this is an explorative study, the small set of candidate genes that could serve as a starting point for unravelling mechanisms of higher level cognitive functions and cortical specialization, and disorders therein such as developmental colour agnosia.

Introduction

Genetics of developmental disorders of the human brain has been a main strategy to uncover biological mechanisms of brain development and functioning. In particular, understanding of the development of the cerebral cortex has benefited from the recognition of specific cortical abnormalities and the elucidation of causative gene defects (Ayala, Shu, & Tsai, 2007; Pilz et al., 1998; Sossey-Alaoui et al., 1998). Genetic aberrations in cortical specializations may be a treasure trove to find genetic dominators and mechanisms of cognitive functions (Hu, Chahrour, & Walsh, 2014; Mitchell, 2011).

Different disorders of cortical specialization have been described (reviewed by (Mitchell, 2011): prosopagnosia (i.e. inability to recognise faces/people despite normal intelligence), synesthesia (i.e. stimulation of one sensory or cognitive pathway leads to automatic, involuntary experiences in a second sensory or cognitive pathway, for example, letters or numbers are perceived as inherently coloured), dyslexia (i.e. disability in reading despite normal intelligence), and congenital amusia (i.e. defect in processing pitch, but may also include musical memory and recognition).

In this study we focussed on a very rare disorder of cortical specialization, namely developmental colour agnosia. Disorders of colour processing have been reported at different levels, ranging from wavelength processing deficits up to very specific impairments in object-colour associations (De Vreese, 1991). One of the most intriguing disorders may be colour agnosia. In colour agnosia, people have intact colour perception, yet have severe difficulties in naming, categorising, and recognising colours and form adequate object-colour associations (Nijboer, van Zandvoort, & de Haan, 2006, 2007a, 2007b; van Zandvoort, Nijboer, & de Haan, 2007). In most cases, colour agnosia is acquired after brain damage, either bilateral or left hemisphere lesions (De Renzi, 2000), mostly in the occipitotemporal lobe (De Vreese, 1991).

Over a decade ago we reported the first and only case of developmental colour agnosia (van Zandvoort et al., 2007). Additionally, a familial factor in developmental colour agnosia was described when family members in three generations also showed difficulties in colour naming and recognition (Nijboer, van Zandvoort, et al., 2007a), indicating a genetic origin of the trait. The affected individuals of this family all display normal cognition and intact colour perception at a visuo-sensory level, e.g. matching of equiluminant colours. However, they could not name colours, match colour names to colours, categorise hues into general clusters of colour or point a colour that is mentioned by the examiner (Nijboer, van Der Smagt, van Zandvoort, & De Haan, 2007). This trait likely is a disorder in

cortical specialization (Mitchell, 2011). Such disorders may have a genetic origin resulting in abnormal cortical development.

In order to identify genetic origins and to gain more insight into the developmental mechanisms related to these disorders, we analyzed all participating members of this family by CGH-arrays and whole-exome sequencing. The results reveal a small set of extremely rare variants that are candidate to cause developmental colour agnosia within the limits of this study.

Subjects and Methods

Subjects and DNA isolation

Informed and written consent were obtained from the family members to perform DNA analysis and sequencing, and to publish results in a scientific paper as approved by the Medical Ethical Committee of the University Medical Center Utrecht. Four members were phenotyped earlier (Nijboer, van Zandvoort, et al., 2007a). For this study, two additional family members were tested and diagnosed with or without developmental colour agnosia. Phenotyping has been described by (Nijboer, van Der Smagt, et al., 2007; Nijboer, van Zandvoort, et al., 2007a). In total three affected and three unaffected family members were available for this study (Fig. 1). Genomic DNA of the six family members was isolated from saliva according to standard procedures. The DNA concentration was determined by using Qubit Quant-IT (Invitrogen).

Array-based comparative genomic hybridization and Linkage Analysis

Array-CGH microarray analysis was performed using 105K (Agilent amadid 019015, hg18) microarray slides from Agilent Technologies (Santa Clara, CA) following manufacturer's protocols and a mixed pool of either 50 healthy males or females was used as a reference. Scanned images were analyzed by Feature Extraction software (Agilent Technologies). Data analysis was performed using DNA analytics 4.76 Software from Agilent Technologies using the ADM-2 algorithm. Interpretation of CNV-data was performed as described in the guidelines of (Vermeesch et al., 2007). Copynumber variations and sequence homologies flanking the breakpoints were confirmed using the Database of Genomic Variation (<http://www.projects.tcag.ca/variation/>) and the UCSC Genome Browser (<http://www.genome.ucsc.edu/index>) (Van Binsbergen, Hochstenbach, Giltay, & Swinkels, 2012).

Linkage analysis was performed by using the array CGH chromosome profiles and mapping of the disease loci was calculated with DominantMapper (Carr et al., 2011). This program calculated which

chromosomal regions are present in all affected family members and not present in all non-affected family members based on SNP genotyping data.

Library preparation and whole exome Sequencing

DNA (2ug) was used for the preparation of a barcoded fragment library (Harakalova et al., 2012). This was followed by multiplexed enrichment and sequencing on the SOLiD Next-generation sequencing platform containing the exonic sequences of ~18,000 genes and covering a total of ~37Mb of genomic sequence (Harakalova et al., 2011, 2012). In brief, the DNA was fragmented using the Covaris S2 System (Applied Biosystems), after fragmentation, the DNA was end repaired and phosphorylated at the 5' end using the End-It DNA End-Repair Kit (EpiCenter) and purified with the Agencourt AMPure XP system (Beckman Coulter Genomics). DNA was then ligated to double-stranded truncated adaptors compatible with the SOLiD next-generation sequencing platform using the Quick Ligation Kit (NEB). After purification, each sequencing library was nick translated, barcoded and amplified in a single PCR assay. The intensity of library bands was examined on a 2% agarose gel (Lonza FlashGel System). Amplified library fragments in the range of 175–225 bp in size were selected on a 4% agarose gel and purified using a QIAquick Gel Extraction Kit (Qiagen). Libraries from the family members were pooled in equimolar concentrations and were enriched using the Agilent SureSelect Human All Exon 50Mb Kit (Agilent Technologies). Enriched library pool fragments were amplified using 12 PCR cycles and elongated to a full-length adaptor sequence required for SOLiD sequencing. SOLiD sequencing was performed according to the instructions in the SOLiD 4 manual to produce enough 50-bp reads to obtain sufficient coverage for a single allele in each library.

Variant detection and exome sequencing analysis

Raw sequencing reads were mapped against the human reference genome GRCh37/hg19 using our custom pipeline based on the Burrows-Wheeler Alignment (BWA) algorithm²⁶. Sequence data were submitted to the EMBL-EBI Sequence Read Archive (see URLs). Single-nucleotide variants (SNVs) and small indels (≤ 7 nt) were called by our custom analysis pipeline as described (Nijman et al., 2010) (all scripts are available upon request). The criteria for variant detection were set to enable discovery of heterozygous variants, since we assume dominant inheritance. We required that variants should minimally be supported by two read seeds (first 25 bp, the higher quality portion of a read), and we set the cut-off for coverage to a minimum of ten reads and the cutoff for non-reference allele percentage to 15%. Since it is probably a dominant inherited phenotype, we set the cut-off for non-reference allele percentage to a maximum of 80%. A maximum of five clonal reads (defined as reads with an identical start site) were included in the analysis. Only truncating variants including nonsense frameshifting insertions or deletions, or missense and splice site alterations were

selected. All polymorphisms with an allele frequency < 0.002 in the different databases in EVS database (<http://evs.gs.washington.edu/EVS/>), Exac (<http://exac.broadinstitute.org/>) or GoNL (<http://www.nlgenome.nl/>) were excluded, and variants present in our in-house database > 0,1,2 x present were excluded from analysis; For each variant, the genomic location, amino-acid change, effect on protein function, conservation score and output from prediction programs (Polyphen, Polyphen-2, SIFT and Condel) were compiled (Harakalova et al., 2012). Alamut was used to study the variants in more detail. Prioritization and variant selection is described in Fig. 2.. Identified variants were confirmed with Sanger sequencing.

Gene expression analyses

Comprehensive human brain transcriptome data sets were used to determine the developmental expression profiles of all candidate genes (Visser et al., 2010). The set contains the age-related expression in 39 samples of the dorsolateral prefrontal cortex from individuals ranging from 0.1 to 83 years of age (Gene Logic Inc, Gaithersburg, MD, USA ; (Stubbs et al., 2012; Visser et al., 2010). Raw data of these studies were available online as GEO entry (GSE11882, Figure 3).

For analysis of gene expression single-cell RNAseq datasets provided by the Allen Brain Institute were used in the Allen Brain Map Transcriptomics Explorer (<https://portal.brain-map.org>). For the human brain on datasets Human-M1-10X and Human-Multiple cortical areas-SMART-seq were used (Hodge et al., 2019). For the mouse brain the data set Mouse-Whole mouse brain & hippocampus-10X was used. In addition, data of the mouse cortex (Tasic, Menon, Nguyen, Kim, Jarsky, Yao, Levi, Gray, Sorensen, Dolbeare, Bertagnolli, Goldy, Shapovalova, Parry, Lee, Smith, Bernard, Madisen, Sunkin, Hawrylycz, Koch, Zeng, et al., 2016; Tasic et al., 2018) were mined through the Single Cell Portal (https://singlecell.broadinstitute.org/single_cell) and the Allen Brain Atlas data portal (http://casestudies.brain-map.org/celltax#section_explorea). The Mouse Brain Atlas and Developing Mouse Brain Atlas were used for anatomical analysis of gene expression (<https://portal.brain-map.org>).

Results and Discussion

CGH-array and Linkage Analysis

Segregation patterns of developmental colour agnosia in the family are most consistent with Mendelian dominant inheritance (Fig 1). Array CGH performed on the six family members showed that there are no large chromosomal rearrangements co-segregating with the phenotype of

developmental colour agnosia. Next, we performed linkage analysis on SNP data to identify loci cosegregating with developmental colour agnosia. All family members were included in the linkage analysis assuming an autosomal dominant model of inheritance. The genome wide linkage resulted in the identification of multiple linked chromosomal regions (Table 2). These linked regions were used for the identification of the candidate genes assuming that the causal gene will be present in a linked region.

Exome sequencing and candidate gene selection

We hypothesized that developmental colour agnosia is caused by a single heterozygous rare non-synonymous variant present in all affected family members and absent in unaffected family members. Candidate gene selection was performed by filtering as described in Fig 2. Exome sequencing followed by quality control identified in total in all family members 63795 variants present in exons and canonical splice sites. Of these, 5908 variants were protein truncating with putative loss of function alteration, including nonsense frameshift (indels) or splice site variants, and were *de novo* or rare in the population (0, 1, 2 times present in the in-house database). Of these 5908 variants, 1797 were not present in non-affected family members and 29 were present in all three affected family members. Of these 29 variants we further studied the frequency of these variants, to assume that the variants are rare or *de novo* in a broader population. Based on frequency of presence in the Exome Variant Server (EVS) database (frequency 0:13000-25:13000 (frequency < 0.002)) (n=17) and Variance GoNL database (n=750; Francioli et al., 2014) (n=16), and confirmation of the variant in the linked region by Sanger sequencing (n=11), 11 candidate variants in 11 genes remained (table 2): *GRINA* (glutamate receptor, ionotropic, N-methyl D-aspartate-associated protein 1), *OR1J1* (olfactory receptor, family 1, subfamily J, member 1), *RABEPK* (Rab9 effector protein with Kelch motifs), *DEPDC7* (DEP domain containing 7), *MAML2* (mastermind-like 2 (Drosophila)), *OR8D4* (olfactory receptor, family 8, subfamily D, member 4), *DDX25* (DEAD (Asp-Glu-Ala-Asp) box polypeptide 25), *CACNA2D4* (calcium channel, voltage-dependent, alpha 2/delta subunit 4), *TMED3* (transmembrane emp24 protein transport domain containing 3), *IL16* (interleukin 16), *MYO15A* (myosin XVA), *STAU1* (staufen double-stranded RNA binding protein 1). Additional information including allele frequency in the 1000 Genome Project (Altshuler et al., 2010)(The 1000 Genomes Project Consortium), Variance GoNL (Boomsma et al., 2014; Francioli et al., 2014) and Welllderly cohort (Erikson, Deshpande, Kesavan, & Torkamani, 2015) is presented in Table 3.

Effect of the mutation

Prediction about the effect of the mutations were made by three software packages: Polyphen, Polyphen-2, SIFT and Condel (Table 2). Two variants (*GRINA* and *RABEPK*) were predicted

to be probably damaging (Polyphen) and deleterious (SIFT). Six variants (*OR1J1*, *DEPDC7*, *MAML2*, *DDX25*, *TMED3*, *STAU1*) were predicted as tolerated by SIFT; some of these were benign or possibly damaging by Polyphen (*OR1J1*, *DDX25*, *DEPDC7*, *TMED3*, *STAU1*). *MAML2* was predicted as probably damaging by Polyphen. *OR8D4* was deleterious on SIFT and benign on Polyphen. *MYO15A* was probably damaging by Polyphen and undefined by SIFT. The effects of two variants (*IL16* and *CACNA2D4*) could not be predicted, since these are splice site mutations (Table 3).

Gene Expression during human brain development.

We have attempted to prioritize the 11 candidate genes for a possible role in developmental colour agnosia. For all these genes, we have analyzed (developmental) brain expression (Fig 3). We hypothesised cortical expression during development for a gene linked to developmental colour agnosia and possible increased expression during postnatal brain maturation. For that reason we first studied expression of the 11 candidate genes in the human brain of subjects free of neurological disorders, in the age range of 0.1 to 83 years, represented by the Ingenuity data set. All genes seem to be expressed in the brain, while two genes, *TMED3* and *MAML2*, displayed an increased expression during development (*TMED3*, 0.4-2.4 years old and *MAML2*, 0.1-2.4; Fig 3). The other nine genes do not show a clear expression profile during a specific developmental age window.

Mining of single-cell transcriptomic data of the adult human cortex and mouse brain showed that *MAML2* was the only gene that maintained expression at high levels, albeit restricted to only a limited number of cell types (Fig 4). The other genes displayed a low to very low level of expression (*GRINA*, *RABEPK*, *DDX25*, *CACNA2D4*, *STAU1*) or no expression. In the human M1 motor cortex *MAML2* was expressed by three types of inhibitory interneurons, one type of excitatory neuron, and by non-neuronal cells: astrocytes, an olfactory progenitor neuron and microglia (Fig 4). In this respect it was the only gene that was expressed by non-neuronal cells. In the mouse cortex and hippocampus *Maml2* was also expressed in a very restricted set of neurons (dentate gyrus and CA3) and in non-neuronal cells. *Grina* and *Stau1* displayed wide, but low expression in inhibitory and excitatory neurons (Fig 4).

Possible link of candidate genes to developmental colour agnosia.

In an attempt to prioritize the 11 identified candidate genes for a possible role in colour agnosia, we have analysed gene function, link to visual perception, link to cortical specialization, conservation of the nucleotide or amino acid, frequency worldwide or in the Dutch population (Table 3) and its presence in previously linked regions for other cortical disorders for example prosopagnosia or dyslexia (Mitchell, 2011).

None of the identified genes or its variants was present amongst previously identified genes or loci related to dyslexia, dyscalculi or prosopagnosia. Linked regions were studied on OMIM (www.omim.org/). GO-analysis identified that the *calcium channel, voltage-dependent, alpha 2/delta subunit 4 (Cav1.4)* gene *CACNA2D4* and *MYO15A* were involved in visual perception, which makes *CACNA2D4* and *MYO15A* candidate genes of interest.

CACNA2D4 encodes the $\alpha 2\delta 4$ subunit of the photoreceptor voltage-gated calcium channels (Cav1.4), which modulate glutamate release from the photoreceptors to the bipolar cells. The *CACNA2D4* mutation is a splice-site mutation and its role on the Cav1.4 protein function is unclear. Mutations in the human *CACNA2D4* gene are involved in recessive cone dystrophy (Ba-Abbad et al., 2016; Wycisk et al., 2006a). Mice with a premature stop mutation in *Cacna2d4* show abnormal morphology of ribbon synapses in the rods and cones (Wycisk et al., 2006b). In mice loss of *Cacna2d4* abolished rod synaptogenesis and synaptic transmission and disrupted mGluR6 clustering (Y. Wang et al., 2017). This suggests that *CACNA2D4* is involved in visual perception by the eye, but it does not provide an argument for its involvement in higher order cortical specialisation and visual perception. There was no indication that the family members with developmental colour agnosia suffered from an ophthalmic problem. This may also suggest that the splice-site mutation that we identified is not deleterious. Furthermore, *CACNA2D4* has been implicated in autism spectrum disorders (X. Liao, Liao, & Li, 2020), and a deletion in the *CACNA2D4* gene was found in a patient with bipolar disorder, but it is unknown if this patient suffers with vision problems (Van Den Bossche et al., 2012). Our attempts to trace this patient for developmental colour agnosia testing failed. Studies of other *CACNA* subtypes showed that thermal pain and tactile stimulation in the *Cacna2d3* mutant mice triggered strong crossactivation of brain regions involved in vision, olfaction, and hearing. This study suggested that this knockout model could be a model for synaesthesia (Neely et al., 2010), which is a disorder related to developmental colour agnosia (Rich & Mattingley, 2002).

MYO15A mutations are associated with deafness due to its critical role for the formation of stereocilia in cochlear hair cells (Probst et al., 1998). Recessive mutations lead to autosomal recessive nonsyndromic hearing loss (A. Wang et al., 1998). *MYO15A* mutations are associated with the Usher syndrome having affected hearing (Baux et al., 2017). Visual perception is also affected In Usher syndrome from the age of 10 years onwards. Vision loss is caused by retinopathy which is due to degeneration of retina cells, usually the rods cells.

Since no direct link to visual perception was found for the other genes, their possible function in the (developing) brain was considered for these genes. It has been indicated that glutamate receptor, ionotropic, N-methyl D-aspartate-associated protein 1 (glutamate binding) (*GRINA*) is involved in nervous system development and disease mainly through signalling mechanisms of neuronal survival (Chen, Yang, Lai, Liu, & Zhu, 2020). Upregulation of *GRINA* transcripts was observed in post-mortem brains of patients with major depressive disorder, supporting a role in brain functions (Goswami et al., 2013). *Grina* knock-out (KO) mice do not display an obvious phenotype (Nielsen et al., 2011). This KO study showed that the effects of the *GRINA* gene are probably subtle. Since it is likely that effects on higher order visual perception will not be evident in mice, *GRINA* remains a potentially interesting candidate gene for consideration of a role in developmental colour agnosia. The effect of the mutation is probably damaging or deleterious (SIFT) and the site of mutation is highly conserved for nucleotide sequence and amino acid motif.

DDX25 (*DEAD (Asp-Glu-Ala-Asp) box polypeptide 25*) encodes a member of the large DEAD-box RNA helicases. Members of this protein family share high structural conservation and are generally involved in developmental processes such as gametogenesis, germ cell specification, and stem cell biology. Data on roles of DDX proteins in the brain are scarce. However, recently two members have been implicated in brain processes. Firstly, a homozygous frameshift deletion in *DDX59* (c.185del: p.Phe62fs*13) has been reported in a family presenting with orofacioidigital syndrome phenotype associated with a broad neurological involvement characterized by microcephaly, intellectual disability, epilepsy, and white matter signal abnormalities associated with cortical and subcortical ischemic events (Salpietro et al., 2018). Further research in *Drosophila* showed impaired development of peripheral and central nervous system upon loss-of-function of *DDX59*, supporting a conserved role of this DEAD-box RNA helicase in neurological function (Salpietro et al., 2018). Secondly, mutations in *DDX3X* cause neurodevelopmental abnormalities in females, including intellectual disability, cortical malformations, autism and epilepsy (Lennox et al., 2020). In this study it was shown that in vivo depletion of *Ddx3x* in mice reduced neuronal differentiation and impaired corticogenesis (Lennox et al., 2020).

A role of *DDX25* in the brain has not been studied. So far *DDX25* has been implicated in spermatogenesis and primordial germ cell development (Kavarthapu et al., 2019; Tsai-Morris, Sheng, Lee, Lei, & Dufau, 2004). Polymorphisms in the *DDX25* gene have been found that alter protein function (Tsai-Morris et al., 2007). During our expression analyses of *DDX25* we found significant expression in the brain. The Allen Brain Atlas Data Portal (http://casestudies.brain-map.org/celltax#section_explore) showed expression of *Ddx25* in mouse cerebral cortex, nucleus accumbens, several thalamic and hypothalamic nuclei, locus coeruleus, raphe nucleus, inferior

olivary complex, and absence in globus pallidus, and most of the midbrain and brain stem. Further exploration of mouse cortical gene expression of the data of (Tasic, Menon, Nguyen, Kim, Jarsky, Yao, Levi, Gray, Sorensen, Dolbeare, Bertagnolli, Goldy, Shapovalova, Parry, Lee, Smith, Bernard, Madisen, Sunkin, Hawrylycz, Koch, & Zeng, 2016) through the Single Cell Portal (https://portals.broadinstitute.org/single_cell) showed expression of *Ddx25* in all subtypes of interneurons and pyramidal cells. Transcripts were absent in glial cells (not shown). Mining of more recent data from single-cell transcriptomics (Hodge et al., 2019) showed that expression of *DDX25* in human and mouse cortex is relatively low and is predominating in somatostatin- and parvalbumin-expressing interneurons, and in several types of pyramidal neurons, mostly located in layers 5-6 (dataset human M1; Fig 4A). In the dataset 'Multiple cortical areas' no evident expression was detected. There was no clear age-related change in human brain samples (Fig 3). In mouse single-cell data sets there was very low expression in parvalbumin-expressing interneurons and layer 5 pyramidal neurons (REFs; Fig 4B). Overall, *DDX25* appears a gene with low and restricted expression in the adult cortex. Since it is part of a little-studied family causing severe neurological problems, as in the case of *DDX3A* and *DDX59* mutations, *DDX25* should still be considered a potential candidate gene for involvement in developmental colour agnosia.

Staufen double-stranded RNA binding protein 1 gene *STAU1* carried a possibly damaging (Polyphen) and tolerated (SIFT) mutation in a highly conserved amino acid motif. *Staufen1* is a key protein involved in the local translation of mRNA in the dendrite and is involved in protein synaptic transmission. Cultured neurons from *Stau1* knockout mice showed reduced dendritic trees and developed fewer synapses, but the mice did not show any obvious defects in brain development (Vessey et al., 2008). We note that recently a molecular and functional relationship between *Stau1* and DEAD-box RNA helicases (*Ddx5/17*) was established in a complex with the Kruppel-like factor *Klf4* (Moon et al., 2018). It was shown that this *Klf4/Stau1/Ddx5/7* protein complex regulates neurogenesis-associated mRNAs and plays a role in mammalian corticogenesis. Furthermore, local regulatory networks containing *STAU1* have been identified in cognitive decline in older adults (Tasaki et al., 2018). In accord with a neuronal role, *STAU1* appeared expressed in virtually all types of cortical neurons in mouse and man, but not in non-neuronal cells of the cerebral cortex.

The *Mastermind-like 2 (MAML2)* gene, encoding a transcriptional activator of the Notch signalling pathway (Ribeiro & Wallberg, 2009), has been implicated in cancers, but has no known function in the nervous system. The effect of the found mutation is probably damaging in Polyphen and tolerated in SIFT. The nucleotide and amino acid sequence is moderate and highly conserved, respectively. In this study *MAML2* is notable since its brain expression is increased during

development (age 0.1-0.5 years) in comparison to older ages (Fig 3). Interestingly, it displayed highest expression by several non-neuronal cell types. In the human cortex these were: oligodendrocyte progenitor cell L1-6 PDGFRA COL20A1 (trimmed mean expression level 10.2), several types of astrocytes (level 9.8-10.3) and microglia (level 10.9) (Fig 4). Neuronal expression of MAML2 at appreciable level occurred in three subtype of inhibitory neurons: inhibitory neuron L1 Pax6 Mir 101-1 (level: 8.2), inhibitory neuron L1-2 VIP WNT4 (level: 8.54), inhibitory neuron L1-6 SST NPY (level 8.67). Two excitatory cell types expressed MAML2: L3 LINC00507 and L3 FEZF2 (respectively at level 8.58 and 7.08). In a single- dataset of the mouse cortex and hippocampus, *Maml2* gene was highest expressed in cells of the dentate gyrus and CA region (Fig 4). The Allen Mouse Brain Atlas confirmed the selective high expression of *Maml2* in the dentate gyrus.

The expression of *transmembrane emp24 protein transport domain containing 3 (TMED3)* is noticeable in its increased between the age of 0.3-8 years in the human prefrontal cortex (fig 3). *TMED3* expression is assigned to all interneurons and pyramidal cells in the mouse cortex (Single Cell Portal;(Tasic, Menon, Nguyen, Kim, Jarsky, Yao, Levi, Gray, Sorensen, Dolbeare, Bertagnolli, Goldy, Shapovalova, Parry, Lee, Smith, Bernard, Madisen, Sunkin, Hawrylycz, Koch, & zheng, 2016), but displays very low to no expression in more recent human and mouse datasets. *TMED3* participates in tumor progression, but functions of *TMED3* in the brain have not been described. Interestingly, it has been proposed that in tumors *TMED3* promotes endogenous Wnt-Tcf activity, which is an important pathway in brain development, and *Il11/Stat3* signalling (Duquet et al., 2014; Zheng et al., 2016).development. The effect of the mutation is possibly damaging in Polyphen and tolerated in SIFT and conservation is moderate for nucleotide and highly conserved for amino acid sequence (Table 3).

The effect of the mutation found in the *DEPDC (DEP domain-containing 7)* gene is a possibly damaging one in Polyphen and tolerated in SIFT (Table 3). Brain expression of *DEPDC7* was not detectable. However, it has been suggested that *DEPDC7* DNA hypomethylation may be associated with depression (Córdova-Palomera et al., 2015). The biological function of the protein encoded by the *DEPDC7* (DEP domain-containing 7) gene is poorly understood. The *DEPDC7* protein participates in gene regulation by NF-κB (D'Andrea et al., 2014). It has been shown that it inhibits tumor growth (Liao, Wang, Wang, Li, & Lin, 2017). The gene has been detected amongst nine genes in a microdeletion in a patient with cryptorchidism and azoospermia (Seabra et al., 2014).

No data are available on the function of olfactory receptors *OR1J1 (olfactory receptor, family 1, subfamily J, member 1)* and *OR8D4*. Expression of ORs has been observed in other tissues than olfactory neurons. *OR8D4* has been encountered as a risk gene of systemic lupus erythematosus. The

mutation for *OR1J1* is benign for Polyphen and tolerated for SIFT. The mutation in *OR8D4* is benign in Polyphen and deleterious in SIFT.

RABEPK (*Rab9 effector protein with Kelch motifs*) has not been studied to any extent. It is thought to participate in protein transport in the trans-Golgi network. The effect of the mutations is a probably damaging in Polyphen and deleterious in SIFT and the conservation is high for both sequences. Expression of the *RABEPK* gene is not detectable in single-cell data sets of the human cortex and in very low level in several types of interneurons and pyramidal cells in the mouse cortex and hippocampus (Fig 4).

Taken together, this study is the first attempt to pinpoint genomic regions and genetic variants in developmental colour agnosia, a cognitive trait affecting the understanding of colours. This study was made possible by the finding and characterization of a unique family carrying this trait (Nijboer, van Zandvoort, et al., 2007b; van Zandvoort et al., 2007). We assumed that developmental colour agnosia is an extremely rare trait. Additional subjects were not found since initiation of this study in 2012. Developmental colour agnosia is not a disorder that requires medical or psychological attention, and will hence stay unnoticed, which hampers the estimation of prevalence.

In addition to number of subjects, this study is also limited by the assumption that developmental colour agnosia is inherited in an autosomal dominant manner with complete penetrance, based on the segregation pattern in the family. Furthermore, the exomes were analysed in 2012 and since then the quality of capture and sequencing have improved. In addition, we cannot exclude the involvement of genomic rearrangement or mutations in noncoding parts of the genome which we would not have captured using exome sequencing. Within these limitations this study indicates 11 genes that link to the trait. Amongst these there are candidate that have been implicated in neural processes and functions (*CACNA2D4*, *MYO15A*, *DDX25*, *GRINA*), that have an indirect link to brain function or development (*STAU1*, *MAML2*, *TMED3*), while others remain unlikely due to lack of brain expression, involvement of nonneural traits (*OR1J1*, *OR8D4*, *DEPDC7*, *RABEPK*)

In summary, we identified 19 genomic regions and 11 coding variants in hereditary colour agnosia by exome sequencing. Since replication is not possible yet due to the uniqueness of this identified family we cannot pinpoint the causal variant at this stage. Further research is required to identify the genetic origins of developmental colour agnosia in this family.

Acknowledgements

This project was supported financially by the Neuroscience Center Utrecht as part of the Focus & Massa programme of the Utrecht University.

Author contributions

Conceptualization: MJvZ, TCN, and JPHB; Cohort assembly: TCN, MJvZ, EHB; Experiments: EVSH, GvH, Computational analyses: EVSH, GvH, CdK, BPCK, PvdS, CT, BvdZ; Data interpretation: EVSH, GvH, JPHB; Writing original draft, TCN, EVSH, GvH.; Writing – review & editing: TCH, JPHB, GvH.

References

- Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., ... Peterson, J. L. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073. <https://doi.org/10.1038/nature09534>
- Ayala, R., Shu, T., & Tsai, L.-H. (2007). Trekking across the brain: the journey of neuronal migration. *Cell*, *128*(1), 29–43. <https://doi.org/10.1016/j.cell.2006.12.021>
- Ba-Abbad, R., Arno, G., Carss, K., Stirrups, K., Penkett, C. J., Moore, A. T., ... Holder, G. E. (2016). Mutations in CACNA2D4 cause distinctive retinal dysfunction in humans. *Ophthalmology*, *123*(3), 668–671.e2. <https://doi.org/10.1016/j.ophtha.2015.09.045>
- Baux, D., Vaché, C., Blanchet, C., Willems, M., Baudoin, C., Moclyn, M., ... Roux, A. F. (2017). Combined genetic approaches yield a 48% diagnostic rate in a large cohort of French hearing-impaired patients. *Scientific Reports*, *7*(1). <https://doi.org/10.1038/s41598-017-16846-9>
- Boomsma, D. I., Wijmenga, C., Slagboom, E. P., Swertz, M. A., Karsen, L. C., Abdellaoui, A., ... Van Duijn, C. M. (2014). The Genome of the Netherlands: Design, and project goals. *European Journal of Human Genetics*, *22*(2), 221–227. <https://doi.org/10.1038/ejhg.2013.118>
- Carr, I. M., Johnson, C. A., Markham, A. F., Toomes, C., Bonthron, D. T., & Sheridan, E. G. (2011). DominantMapper: Rule-based analysis of SNP data for rapid mapping of dominant diseases in related nuclear families. *Human Mutation*, *32*(12), 1359–1366. <https://doi.org/10.1002/humu.21597>
- Chen, K., Yang, L. N., Lai, C., Liu, D., & Zhu, L.-Q. (2020). Role of Grina/Nmdara1 in the Central Nervous System Diseases. *Current Neuropharmacology*, *18*(9), 861–867. <https://doi.org/10.2174/1570159x18666200303104235>
- Córdova-Palomera, A., Fatjó-Vilas, M., Palma-Gudiel, H., Blasco-Fontecilla, H., Kebir, O., & Fañanás, L. (2015). Further evidence of DEPDC7 DNA hypomethylation in depression: A study in adult twins. *European Psychiatry*, *30*(6), 715–718. <https://doi.org/10.1016/j.eurpsy.2015.04.001>
- D’Andrea, E. L., Ferravante, A., Scudiero, I., Zotti, T., Reale, C., Pizzulo, M., ... Stilo, R. (2014). The dishevelled, EGL-10 and pleckstrin (DEP) domain-containing protein DEPDC7 binds to CARMA2 and CARMA3 proteins, and regulates NF-κB activation. *PLoS ONE*, *9*(12). <https://doi.org/10.1371/journal.pone.0116062>
- De Renzi, E. (2000). Disorders of visual recognition. *Seminars in Neurology*. Semin Neurol. <https://doi.org/10.1055/s-2000-13181>
- De Vreese, L. P. (1991). Two systems for colour-naming defects: Verbal disconnection vs colour imagery disorder. *Neuropsychologia*, *29*(1), 1–18. [https://doi.org/10.1016/0028-3932\(91\)90090-U](https://doi.org/10.1016/0028-3932(91)90090-U)
- Duquet, A., Melotti, A., Mishra, S., Malerba, M., Seth, C., Conod, A., & Ruiz i Altaba, A. (2014). A novel genome-wide in vivo screen for metastatic suppressors in human colon cancer identifies the positive WNT - TCF pathway modulators TMED 3 and SOX 12 . *EMBO Molecular Medicine*, *6*(7), 882–901. <https://doi.org/10.15252/emmm.201303799>
- Erikson, G. A., Deshpande, N., Kesavan, B. G., & Torkamani, A. (2015). SG-ADVISER CNV: Copy-number variant annotation and interpretation. *Genetics in Medicine*, *17*(9), 714–718. <https://doi.org/10.1038/gim.2014.180>
- Francioli, L. C., Menelaou, A., Pulit, S. L., Van Dijk, F., Palamara, P. F., Elbers, C. C., ... Wijmenga, C. (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nature Genetics*, *46*(8), 818–825. <https://doi.org/10.1038/ng.3021>
- Goswami, D. B., Jernigan, C. S., Chandran, A., Iyo, A. H., May, W. L., Austin, M. C., ... Karolewicz, B. (2013). Gene expression analysis of novel genes in the prefrontal cortex of major depressive disorder subjects. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *43*, 126–133. <https://doi.org/10.1016/j.pnpbp.2012.12.010>
- Harakalova, M., Mokry, M., Hrdlickova, B., Renkens, I., Duran, K., Van Roekel, H., ... Cuppen, E. (2011). Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective

- targeted next-generation sequencing. *Nature Protocols*, 6(12), 1870–1886.
<https://doi.org/10.1038/nprot.2011.396>
- Harakalova, M., Van Harssel, J. J. T., Terhal, P. A., Van Lieshout, S., Duran, K., Renkens, I., ... Cuppen, E. (2012). Dominant missense mutations in ABCC9 cause Cantú syndrome. *Nature Genetics*, 44(7), 793–796. <https://doi.org/10.1038/ng.2324>
- Hodge, R. D., Bakken, T. E., Miller, J. A., Smith, K. A., Barkan, E. R., Graybuck, L. T., ... Lein, E. S. (2019). Conserved cell types with divergent features in human versus mouse cortex. *Nature*, 573(7772), 61–68. <https://doi.org/10.1038/s41586-019-1506-7>
- Hu, W. F., Chahrour, M. H., & Walsh, C. A. (2014). The diverse genetic landscape of neurodevelopmental disorders. *Annual Review of Genomics and Human Genetics*, 15, 195–213. <https://doi.org/10.1146/annurev-genom-090413-025600>
- Just Ribeiro, M., & Wallberg, A. (2009). Transcriptional Mechanisms by the Coregulator MAML1. *Current Protein & Peptide Science*, 10(6), 570–576. <https://doi.org/10.2174/138920309789630543>
- Kavarthapu, R., Anbazhagan, R., Raju, M., Morris, C. H. T., Pickel, J., & Dufau, M. L. (2019). Targeted knock-in mice with a human mutation in GRTH/DDX25 reveals the essential role of phosphorylated GRTH in spermatid development during spermatogenesis. *Human Molecular Genetics*, 28(15), 2561–2572. <https://doi.org/10.1093/hmg/ddz079>
- Lennox, A. L., Hoye, M. L., Jiang, R., Johnson-Kerner, B. L., Suit, L. A., Venkataramanan, S., ... Sherr, E. H. (2020). Pathogenic DDX3X Mutations Impair RNA Metabolism and Neurogenesis during Fetal Cortical Development. *Neuron*, 106(3), 404-420.e8. <https://doi.org/10.1016/j.neuron.2020.01.042>
- Liao, X., Liao, X., & Li, Y. (2020, June 22). Genetic associations between voltage-gated calcium channels and autism spectrum disorder: A systematic review. *Molecular Brain*. BioMed Central Ltd. <https://doi.org/10.1186/s13041-020-00634-0>
- Liao, Z., Wang, X., Wang, X., Li, L., & Lin, D. (2017). DEPDC7 inhibits cell proliferation, migration and invasion in hepatoma cells. *Oncology Letters*, 14(6), 7332–7338. <https://doi.org/10.3892/ol.2017.7128>
- Mitchell, K. J. (2011). Curiouser and curiouser: genetic disorders of cortical specialization. *Current Opinion in Genetics & Development*, 21(3), 271–277. <https://doi.org/10.1016/j.gde.2010.12.003>
- Moon, B. S., Bai, J., Cai, M., Liu, C., Shi, J., & Lu, W. (2018). Kruppel-like factor 4-dependent Staufen1-mediated mRNA decay regulates cortical neurogenesis. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-017-02720-9>
- Neely, G. G., Hess, A., Costigan, M., Keene, A. C., Goulas, S., Langeslag, M., ... Penninger, J. M. (2010). A genome-wide Drosophila screen for heat nociception identifies $\alpha 2\delta 3$ as an evolutionarily conserved pain gene. *Cell*, 143(4), 628–638. <https://doi.org/10.1016/j.cell.2010.09.047>
- Nielsen, J. A., Chambers, M. A., Romm, E., Lee, L. Y. H., Berndt, J. A., & Hudson, L. D. (2011). Mouse transmembrane BAX inhibitor motif 3 (Tmbim3) encodes a 38 kDa transmembrane protein expressed in the central nervous system. *Molecular and Cellular Biochemistry*, 357(1–2), 73–81. <https://doi.org/10.1007/s11010-011-0877-3>
- Nijboer, T. C. W., van Der Smagt, M. J., van Zandvoort, M. J. E., & De Haan, E. H. F. (2007). Colour agnosia impairs the recognition of natural but not of non-natural scenes. *Cognitive Neuropsychology*, 24(2), 152–161. <https://doi.org/10.1080/02643290600989541>
- Nijboer, T. C. W., van Zandvoort, M. J. E., & de Haan, E. H. F. (2006). Covert colour processing in colour agnosia. *Neuropsychologia*, 44(8), 1437–1443. <https://doi.org/10.1016/j.neuropsychologia.2005.12.004>
- Nijboer, T. C. W., van Zandvoort, M. J. E., & de Haan, E. H. F. (2007a). A familial factor in the development of colour agnosia. *Neuropsychologia*, 45(8), 1961–1965. <https://doi.org/10.1016/j.neuropsychologia.2007.01.021>
- Nijboer, T. C. W., van Zandvoort, M. J. E., & de Haan, E. H. F. (2007b). A familial factor in the development of colour agnosia. *Neuropsychologia*, 45(8), 1961–1965. <https://doi.org/10.1016/j.neuropsychologia.2007.01.021>

- Nijman, I. J., Mokry, M., Van Boxtel, R., Toonen, P., De Bruijn, E., & Cuppen, E. (2010). Mutation discovery by targeted genomic enrichment of multiplexed barcoded samples. *Nature Methods*, 7(11), 913–915. <https://doi.org/10.1038/nmeth.1516>
- Pilz, D. T., Matsumoto, N., Minnerath, S., Mills, P., Gleeson, J. G., Allen, K. M., ... Ross, M. E. (1998). LIS1 and XLIS (DCX) mutations cause most classical lissencephaly, but different patterns of malformation. *Human Molecular Genetics*, 7(13), 2029–2037. <https://doi.org/10.1093/hmg/7.13.2029>
- Probst, F. J., Fridell, R. A., Raphael, Y., Saunders, T. L., Wang, A., Liang, Y., ... Camper, S. A. (1998). Correction of deafness in shaker-2 mice by an unconventional myosin in a BAC transgene. *Science*, 280(5368), 1444–1447. <https://doi.org/10.1126/science.280.5368.1444>
- Rich, A. N., & Mattingley, J. B. (2002). Anomalous perception in synaesthesia: A cognitive neuroscience perspective. *Nature Reviews Neuroscience*, 3(1), 43–52. <https://doi.org/10.1038/nrn702>
- Salpietro, V., Efthymiou, S., Manole, A., Maurya, B., Wiethoff, S., Ashokkumar, B., ... Houlden, H. (2018). A loss-of-function homozygous mutation in DDX59 implicates a conserved DEAD-box RNA helicase in nervous system development and function. *Human Mutation*, 39(2), 187–192. <https://doi.org/10.1002/humu.23368>
- Seabra, C. M., Quental, S., Neto, A. P., Carvalho, F., Gonçalves, J., Oliveira, J. P., ... Lopes, A. M. (2014). A novel Alu-mediated microdeletion at 11p13 removes WT1 in a patient with cryptorchidism and azoospermia. *Reproductive BioMedicine Online*, 29(3), 388–391. <https://doi.org/10.1016/j.rbmo.2014.04.017>
- Sossey-Alaoui, K., Hartung, A. J., Guerrini, R., Manchester, D. K., Posar, A., Puche-Mira, A., ... Srivastava, A. K. (1998). Human doublecortin (DCX) and the homologous gene in mouse encode a putative Ca²⁺-dependent signaling protein which is mutated in human X-linked neuronal migration defects. *Human Molecular Genetics*, 7(8), 1327–1332. <https://doi.org/10.1093/hmg/7.8.1327>
- Stubbs, A., McClellan, E. A., Horsman, S., Hiltmann, S. D., Palli, I., Nouwens, S., ... van der Spek, P. J. (2012). Huvariome: a web server resource of whole genome next-generation sequencing allelic frequencies to aid in pathological candidate gene selection. *Journal of Clinical Bioinformatics*, 2(1), 19. <https://doi.org/10.1186/2043-9113-2-19>
- Tasaki, S., Gaiteri, C., Mostafavi, S., Yu, L., Wang, Y., De Jager, P. L., & Bennett, D. A. (2018). Multi-omic directed networks describe features of gene regulation in aged brains and expand the set of genes driving cognitive decline. *Frontiers in Genetics*, 9(AUG). <https://doi.org/10.3389/fgene.2018.00294>
- Tasic, B., Menon, V., Nguyen, T. N., Kim, T. K., Jarsky, T., Yao, Z., ... Zeng, H. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience*, 19(2), 335–346. <https://doi.org/10.1038/nn.4216>
- Tasic, B., Menon, V., Nguyen, T. N. T., Kim, T. T. K., Jarsky, T., Yao, Z., ... Zeng, H. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience*, advance on(January), 1–37. <https://doi.org/10.1038/nn.4216>
- Tasic, B., Yao, Z., Graybiel, L. T., Smith, K. A., Nguyen, T. N., Bertagnolli, D., ... Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature*, 563(7729), 72–78. <https://doi.org/10.1038/s41586-018-0654-5>
- Tsai-Morris, C. H., Koh, E., Sheng, Y., Maeda, Y., Gutti, R., Namiki, M., & Dufau, M. L. (2007). Polymorphism of the GRTH/DDX25 gene in normal and infertile Japanese men: A missense mutation associated with loss of GRTH phosphorylation. *Molecular Human Reproduction*, 13(12), 887–892. <https://doi.org/10.1093/molehr/gam065>
- Tsai-Morris, C. H., Sheng, Y., Lee, E., Lei, K. J., & Dufau, M. L. (2004). Gonadotropin-regulated testicular RNA helicase (GRTH/Ddx25) is essential for spermatid development and completion of spermatogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 101(17), 6373–6378. <https://doi.org/10.1073/pnas.0401855101>
- Van Binsbergen, E., Hochstenbach, R., Giltay, J., & Swinkels, M. (2012). Unstable transmission of a

- familial complex chromosome rearrangement. *American Journal of Medical Genetics, Part A*, 158 A(11), 2888–2893. <https://doi.org/10.1002/ajmg.a.35580>
- Van Den Bossche, M. J., Strazisar, M., De Bruyne, S., Bervoets, C., Lenaerts, A. S., De Zutter, S., ... Del-Favero, J. (2012). Identification of a CACNA2D4 deletion in late onset bipolar disorder patients and implications for the involvement of voltage-dependent calcium channels in psychiatric disorders. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, 159 B(4), 465–475. <https://doi.org/10.1002/ajmg.b.32053>
- van Zandvoort, M. J. E., Nijboer, T. C. W., & de Haan, E. (2007). Developmental colour agnosia. *Cortex*, 43(6), 750–757. [https://doi.org/10.1016/S0010-9452\(08\)70503-3](https://doi.org/10.1016/S0010-9452(08)70503-3)
- Vermeesch, J. R., Fiegler, H., de Leeuw, N., Szuhai, K., Schoumans, J., Ciccone, R., ... Zuffardi, O. (2007). Guidelines for molecular karyotyping in constitutional genetic diagnosis. *European Journal of Human Genetics*, 15(11), 1105–1114. <https://doi.org/10.1038/sj.ejhg.5201896>
- Vessey, J. P., Macchi, P., Stein, J. M., Mikl, M., Hawker, K. N., Vogelsang, P., ... Kiebler, M. A. (2008). A loss of function allele for murine Staufen1 leads to impairment of dendritic Staufen1-RNP delivery and dendritic spine morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), 16374–16379. <https://doi.org/10.1073/pnas.0804583105>
- Visser, W. E., Swagemakers, S. M. A., Ozgur, Z., Schot, R., Verheijen, F. W., van Ijcken, W. F. J., ... Visser, T. J. (2010). Transcriptional profiling of fibroblasts from patients with mutations in MCT8 and comparative analysis with the human brain transcriptome. *Human Molecular Genetics*, 19(21), 4189–4200. <https://doi.org/10.1093/hmg/ddq337>
- Wang, A., Liang, Y., Fridell, R. A., Probst, F. J., Wilcox, E. R., Touchman, J. W., ... Friedman, T. B. (1998). Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. *Science*, 280(5368), 1447–1451. <https://doi.org/10.1126/science.280.5368.1447>
- Wang, Y., Fehlhaber, K. E., Sarria, I., Cao, Y., Ingram, N. T., Guerrero-Given, D., ... Martemyanov, K. A. (2017). The Auxiliary Calcium Channel Subunit $\alpha 2\delta 4$ Is Required for Axonal Elaboration, Synaptic Transmission, and Wiring of Rod Photoreceptors. *Neuron*, 93(6), 1359–1374.e6. <https://doi.org/10.1016/j.neuron.2017.02.021>
- Wycisk, Katharina A., Budde, B., Feil, S., Skosyrski, S., Buzzi, F., Neidhardt, J., ... Berger, W. (2006). Structural and functional abnormalities of retinal ribbon synapses due to Cacna2d4 mutation. *Investigative Ophthalmology and Visual Science*, 47(8), 3523–3530. <https://doi.org/10.1167/iovs.06-0271>
- Wycisk, Katharina Agnes, Zeitz, C., Feil, S., Wittmer, M., Forster, U., Neidhardt, J., ... Berger, W. (2006). Mutation in the auxiliary calcium-channel subunit CACNA2D4 causes autosomal recessive cone dystrophy. *American Journal of Human Genetics*, 79(5), 973–977. <https://doi.org/10.1086/508944>
- Zheng, H., Yang, Y., Han, J., Jiang, W. H., Chen, C., Wang, M. C., ... Zhou, W. P. (2016). TMED3 promotes hepatocellular carcinoma progression via IL-11/STAT3 signaling. *Scientific Reports*, 6. <https://doi.org/10.1038/srep37070>

Figure Legends

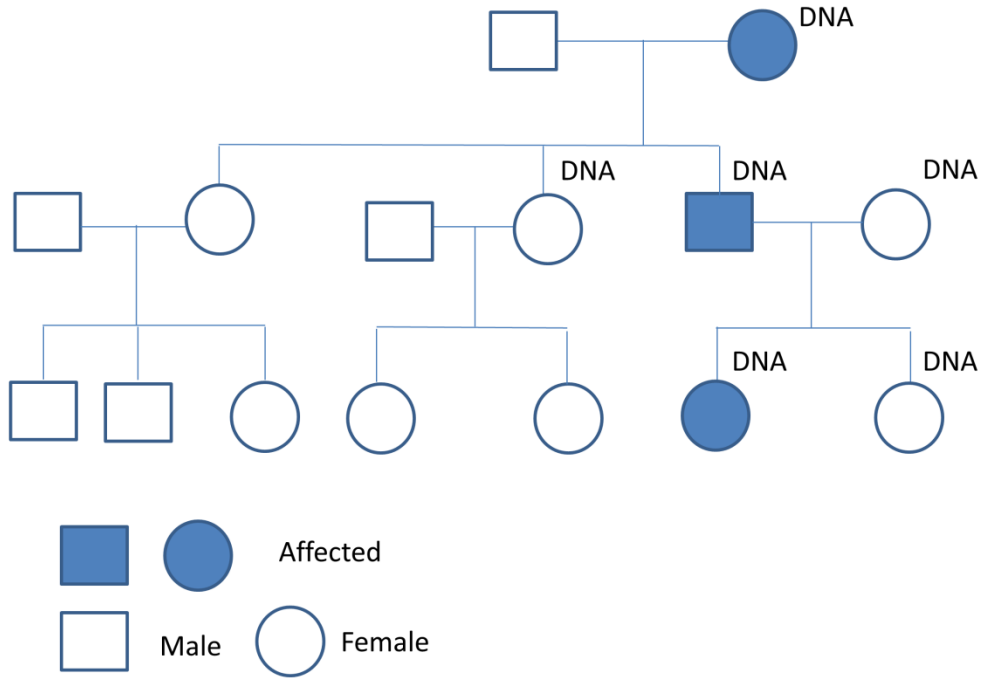


Fig. 1. Family with Developmental Colour Agnosia. DNA means that these family members underwent genetic analysis. Three affected persons are indicated (filled dark blue).

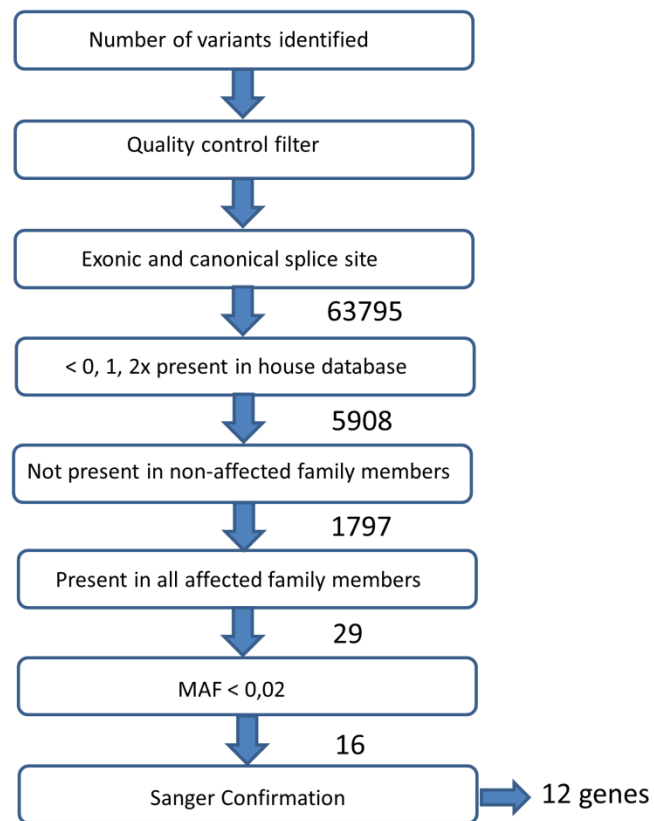


Fig. 2. Selection of variants cosegregating with developmental colour agnosia based on whole exome sequencing. Selection is based on the presence of the mutation in the affected family members and the absence in the unaffected family members.

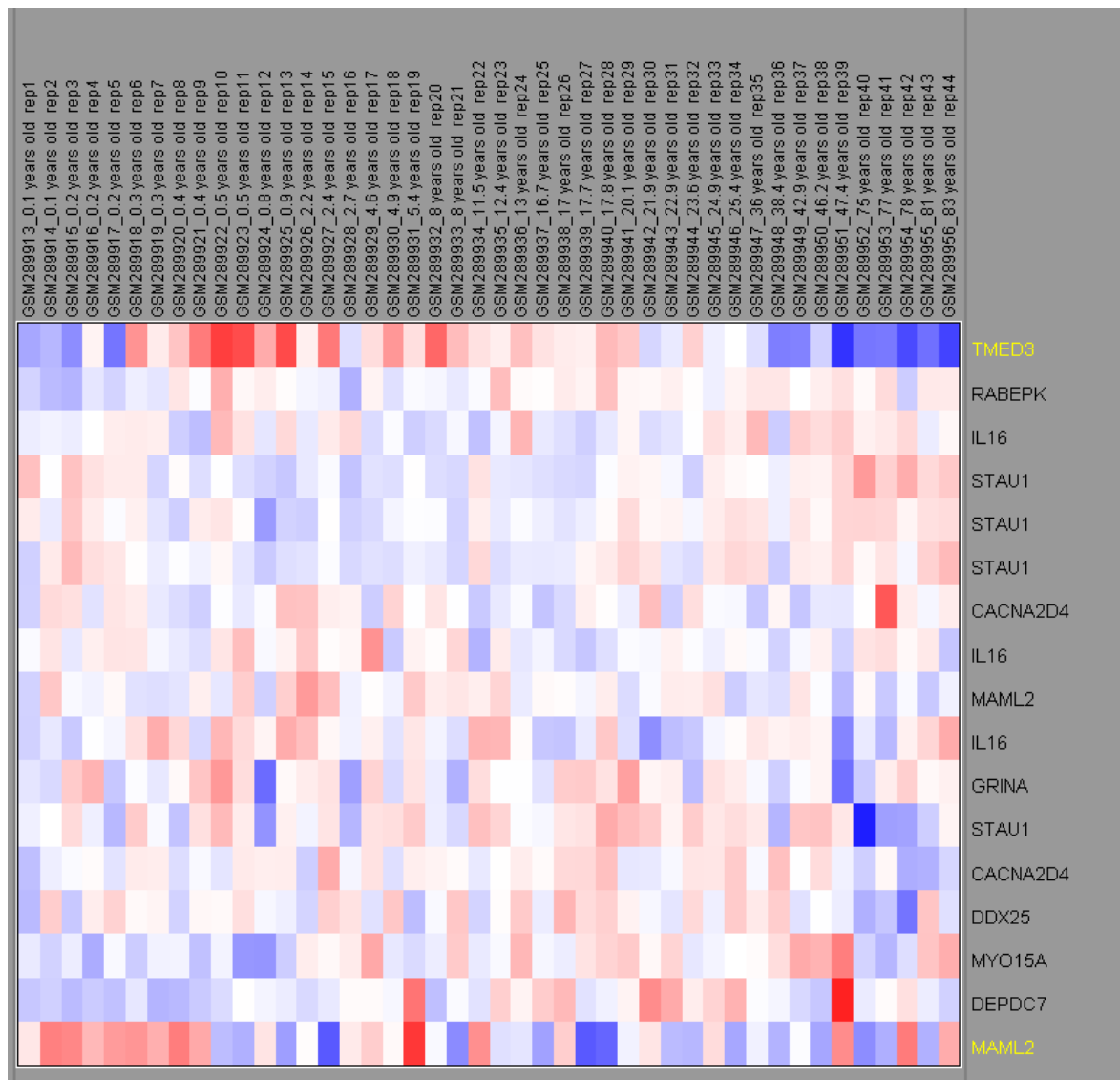
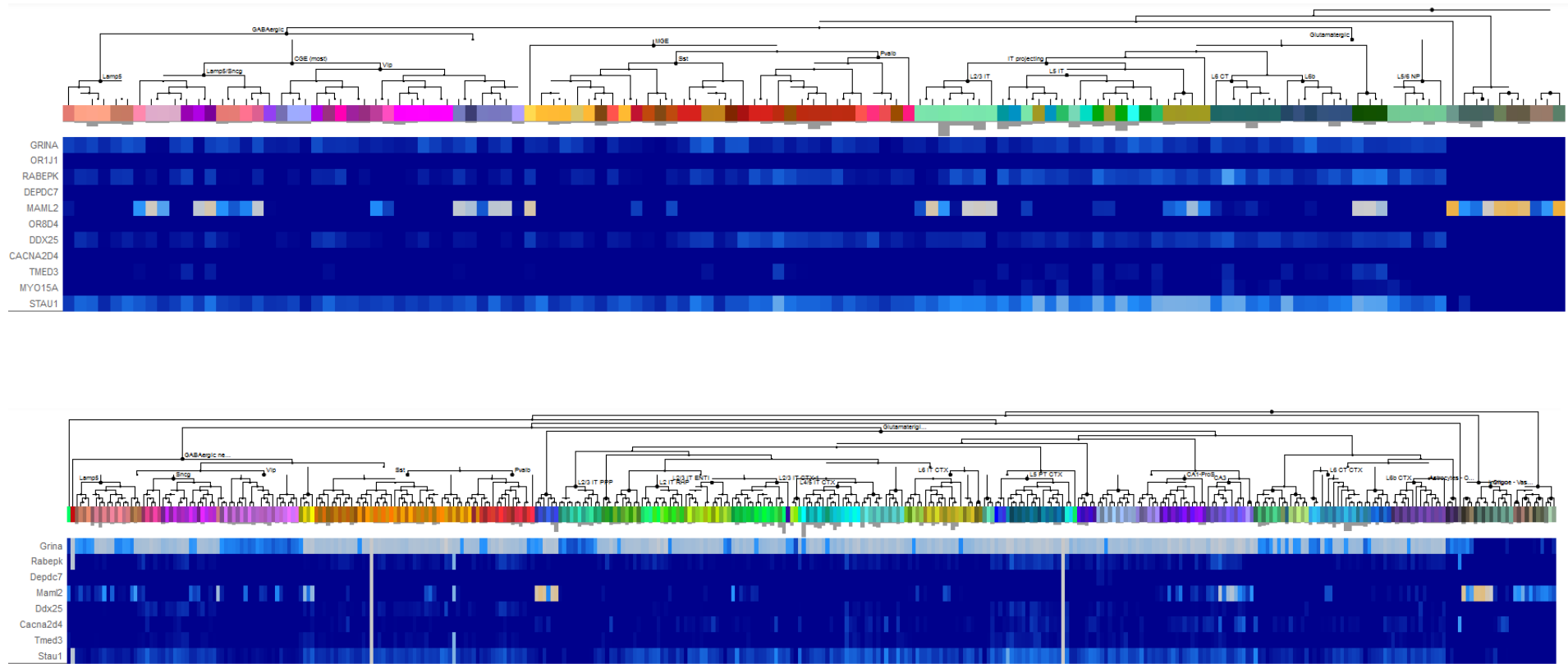


Fig. 3. Expression of the candidate genes in the human brain (prefrontal cortex) during development. The upper lane describes the age of the brain samples and sample ID. The gene names are mentioned at the right. Red box means an up regulation, blue box is a down regulation and white is no change in expression during development.



Mouse brain: whole cortex and hippocampus. Olfactory receptors *OR1J1* and *OR8D4* lack identifiable orthologues in the mouse

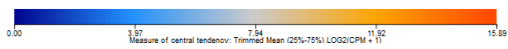


Fig. 4. Expression of the candidate genes in different neuronal and non-neuronal cell types in the human brain (M1 motor cortex) and the mouse cortex and hippocampus. Data were obtained from the single-cell RNA-seq cell types databases of the Allen Brain Institute (<https://portal.brain-map.org/atlas-and-data/rnaseq>). Expression levels are plotted under the neuronal and nonneuronal cell types identified in these samples. Top: human, Bottom: mouse.

Table 1: Linked chromosomal regions to colour agnosia

Chr.	Locus	Start	End	Length
2	2q22.3-35	146360626	221800201	75439575
3	3p11.2-q11.2	87373894	94512868	7138974
3	3q13.3-q25.31	119442572	159239866	39797294
4	4q13.1-4q22.3	60013919	96704454	36690535
5	5p15.33-5p15.2	2346621	10752315	8405694
6	6q26-27	162736476	168554222	5817746
8	8q24.23-24.3	138970102	143270729	4300627
8	8p22-q13.1	14253354	68723609	54470255
9	9q21.32-q34.11	83858558	138149166	54290608
10	10p15.3-p14	135708	8638235	8502527
11	11p14.3-q24	21234023	128206410	106972387
12	12p13-q23.3	661031	67262352	66601321
14	14q12-q31.2	22645683	88752821	66107138
15	15q15.1-q26.1	40509600	94077313	53567713
17	17p12-q25.3	9078327	77698582	68620255
19	19p12-q21	23446279	28836326	5390047
20	20q13.13-q13.33	47050483	60111743	13061260
21	21q11.2-q22.12	14687571	38082812	23395241
22	22q11.2-q13.2	16079545	44595591	28516046

Chromosomal regions linked to developmental colour agnosia (hg19). These haplotypes of these regions are present in all affected individuals and absent in the non-affected individuals.

Table 2: Properties of selected candidate genes

Genome position (hg19)	Nucleotide change	Gene symbol	Ensemble transcript	Consequence	Amino acid change	SNP id	Protein function
8_145066215	G/A	GRINA	ENST00000313269	NON_SYNONYMOUS	R221Q	novel	Glutamate Receptor
9_125239752	C/T	OR1J1	ENST00000259357	NON_SYNONYMOUS	A152T	rs143862742	Olfactory receptor
9_127990266	G/A	RABEPK	ENST00000373538	NON_SYNONYMOUS	G202R	rs142347258	RAB9 effector protein
11_33052944	C/T	DEPDC7	ENST00000241051	NON_SYNONYMOUS	A268V	novel	G protein signalling DEP domain
11_95712527	T/C	MAML2	ENST00000524717	NON_SYNONYMOUS	Q1019R	rs201299551	Transcriptional Coactivator Notch
11_123777296	G/A	OR8D4	ENST00000321355	NON_SYNONYMOUS	R53H	rs201007236	Olfactory receptor
11_125786960	C/G	DDX25	ENST00000263576	NON_SYNONYMOUS	D284E	rs370049057	DEAD box RNA helicase
12_1949897	G/A	CACNA2D4	ENST00000382722	SPLICE_SITE	NA	rs376597005	Calcium Channel
15_79614380	C/T	TMED3	ENST00000299705	NON_SYNONYMOUS	H160Y	novel	Transmembrane protein
17_18043906	C/T	MYO15A	ENST00000205890	NON_SYNONYMOUS	R1763W	rs200146361	Actin based motor molecule
20_47740944	T/C	STAU1	ENST00000371856	NON_SYNONYMOUS	R264G	rs201180807	RNA binding protein

Table 3: Variant frequencies and mutation prediction

Gene_name	Frequencies			Effect of mutation					Conservation (Condel)	
	EVS	GoNL	Exac	Effect	Polyphen_s	Polyphen_p	SIFT_s	SIFT_p	Conserved Nucl.	Conserved AA
<i>GRINA</i>	1/13005	0	0/121000	NS	1	probably damaging	0	deleterious	highly	highly
<i>OR1J1</i>	0	0	1/120930	NS	0,001	benign	0,56	tolerated	weakly	highly
<i>RABEPK</i>	11/12995	2/996	36/120704	NS	1	probably damaging	0	deleterious	highly	highly
<i>DEPDC7</i>	0	0	0/120000	NS	0,896	possibly damaging	0,07	tolerated	moderately	moderately
<i>MAML2</i>	1/12217	0	50/119130	NS	0,998	probably damaging	0,06	tolerated	moderately	highly
<i>OR8D4</i>	6/12996	1/996	102/121352	NS	0	benign	0	deleterious	weakly	weakly
<i>DDX25</i>	1/12537	1/996	3/76840	NS	0	benign	0,49	tolerated	weakly	highly
<i>CACNA2D4</i>	0	0	3/120548	SS	undef	undef	undef	undef	na (splicesite)	na (splicesite)
<i>TMED3</i>	0	0	0/121000	NS	0,773	possibly damaging	0,06	tolerated	moderately	highly
<i>MYO15A</i>	17/12365	4/996	199/119810	NS	0,999	probably damaging	undef	undef	weakly	highly
<i>STAU1</i>	3/13003	0	16/121402	NS	0,911	possibly damaging	0,33	tolerated	weakly	highly

Candidate variants and their frequency in different databases. Polyphen and SIFT predict the effect of the mutation. Effect NS is non-synonymous, SS splice site. Conserved Nucl. shows the conservation of the nucleotides and conserved AA determined the conservation of the AA.