




Future perspectives

on genetic diagnostics





Johan den Dunnen

Human and Clinical Genetics
© JT den Dunnen 


Human & Clinical Genetics

(Leiden University Medical Center)




CMSB

- **Genetic Disease**
neuromuscular disorders
[http:// www.DMD.nl](http://www.DMD.nl)
diagnosis
treatment / therapy




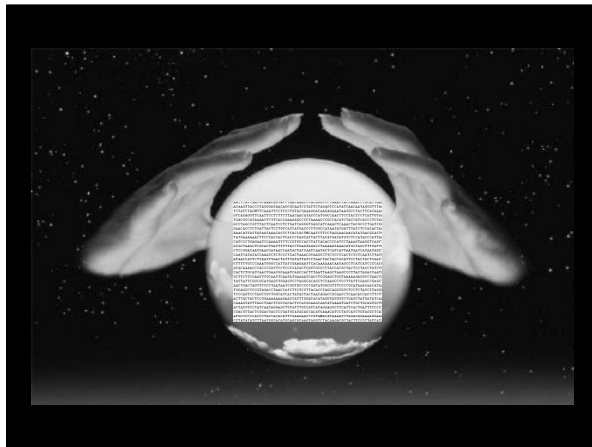
LGTC



CBG
Center for Biomedical Genetics

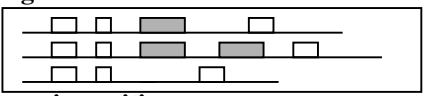

- **Genome Technology**
try and apply
facilitate
Leiden Genome Technology Center
<http:// www.LGTC.nl />

Human and Clinical Genetics
© JT den Dunnen 




Possible variants

- **change in sequence**

ACATCAGGAGAAGATGTC GAGACTTGCCA
ACATCAGGAGAAGATGTT GAGACTTGCCA
ACATCAGGAGAAGATGT GAGACTTGCCA
ACATCAGGAGAAGATGTTCCGAGACTTGCCA
- **change in amount**

- **change in position**



Normal Y chromosome vs Y chromosome with insertion

Human and Clinical Genetics
© JT den Dunnen 

Summary

- **change in sequence**
sequence
- **change in amount (CNV)**
sequence & count
paired-end > breakpoints
- **change in position**
sequence
paired-end > breakpoint

Non-Invasive Prenatal diagnosis
sequence ccf DNA/RNA
paired-end > size

Human and Clinical Genetics
© JT den Dunnen 



LGTC technologies

- **DNA**
genomics
- **RNA**
transcriptomics

Bioinformatics

- **protein**
proteomics
- **metabolite**
metabolomics

Human and Clinical Genetics
© JT den Dunnen

LGTC equipment

sequencing
ABI3730, ABI3730xl
pyrosequence
Solexa

array technology
Affymetrix (incl. 4-colour)
Illumina (incl. BeadXpress)
Agilent scanner
hybstations
arrayer (OmniGrid)
FlexArrayer

PCR
TaqMan 7900HT
LightCycler480
Fluidigm
...

other
LightScanner (Idaho)
BioAnalyzer
NanoDrop
Caliper LC-90
robotics (Tecan, Caliper)
GeneTAC G3

Human and Clinical Genetics
© JT den Dunnen

Melting Curve Analysis

(MCA)

- **principle**
sequence difference > altered melting
mismatch in heteroduplex reduces T_m
- **detection**
DNA + intercalating
fluorescent dye
dsDNA
measure fluorescence
while raising T
denaturation gives
decreasing
fluorescence
- **closed-tube assay**

Human and Clinical Genetics
© JT den Dunnen

MCA applications

- **variant detection**
BRCA1/2, DMD, ...
- **SNP typing**
melt probes
- **gel electrophoresis**
- **allelic imbalances**
CNV confirmation
- **methylation**
- **clone identification**

Human and Clinical Genetics
© JT den Dunnen

Clone selection

©Rolf Vossen

Phage-display

96 clones selected

MCA > determine complexity

- first conclusion
- select for sequencing

more informative than
gel electrophoresis

Human and Clinical Genetics
© JT den Dunnen

Variant detection

©Rolf Vossen
©Jan Harryvan

ApoE
E2 / E3 / E4

Human and Clinical Genetics
© JT den Dunnen

CNV & hrMCA

- **SNP-based required**
homozygous sample AA & BB
test sample A? (AA, A0, AAA)
- **assay amplify**
AA, BB, test sample
1:1 mix AA / BB
1:2 mix AA / BB
1:1 mix BB / test (A?)
other mixes (1:3, 3:1, etc.)
- **result**
test = AA / BB mix > AA / BB
test ≠ AA / BB mix > A0 / BB or AAA / BB

Human and Clinical Genetics © JT den Dunnen

Somatic mosaicism

(high-resolution Melting Curve Analysis)

Human and Clinical Genetics © JT den Dunnen

hrMCA octaploid

(large shift variant)

Human and Clinical Genetics © JT den Dunnen

DNA sequencing

Human and Clinical Genetics © JT den Dunnen

LGTC / Solexa

- **LGTC characteristics**
not a large sequencing center
no Bioinformatics support
result on CD (hard disk)
academic customers
hesitating new users
many different applications
all formats
do it yourself / outsourcing
collaboration ServiceXS

offer alternative technologies

Solexa 1G since Q1 2007

Human and Clinical Genetics © JT den Dunnen

Summary

- **system performs as expected**
with ups & downs
steadily increasing performance
0.5 MB to 2.4 Mb runs
- **important aspects**
sample preparation !
skilled operator
- **data analysis**
software behind on technology
de novo assembly

Sophie Greve

Yavuz Ariyurek

Human and Clinical Genetics

Current projects

- **replace array**
Chromatin-IP
micro RNA profiling & discovery
gene expression profiling
SNP typing
- **SNP-discovery**
- **genome re-sequencing**
- **disease candidate gene / gene region**
- **de novo genome sequencing**
-



Human and Clinical Genetics

© JT den Dunnen LUMC

Expression profiling

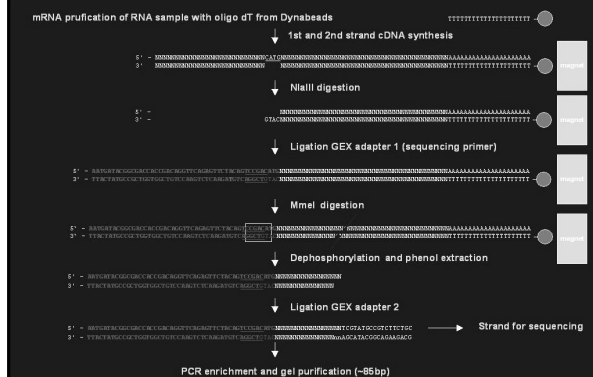
- **model**
wt / transgenic DCLK
constitutive expression δC-doublecortin-like kinase
brain > hippocampus
subtle behavioural abnormalities
- **micro-array analysis**
5 platforms
> only subtle changes
> biological replicates
- **approach**
pools (wt / transgenics)
Solexa / Illumina
individual mice
Leiden (n=4 per group)



Human and Clinical Genetics

© JT den Dunnen LUMC

Solexa – Generation SAGE tags



Results expression

Sequence	Sample 1	Sample 2	Sum	GENE	DESCRIPTION
CATGATAATACATAAAAAA	11484	101445	216933		mitochondrial
CATGTATTAATAAATAGAGCA	41359	31710	73069	Tspan7	Tetraspanin 7
CATGTAACTTAAGATGCTGC	28658	36522	64280	Olfml1	Olfactomedin 1
CATGGATAATAAAACTATT	30593	30787	61370	Hsp48	Heat shock protein 4
CATGGCCTTGGCTTTTGA	22855	27311	50166	Rplp1	Ribosomal protein, large, P1
CATGCATCTTGATGTTTCC	17745	20549	38294	Ckb	Creatine kinase, brain
CATGTGTAATAAATCGGATA	18165	18401	36566	2900019G14Rik	Riken cDNA 2900019G14 gene
CATGCTACCGCTGTAATGG	17787	17577	35364	E130013N25Rik	Riken cDNA E130013N25 gene
CATGCTGAGGAGTGAATAC	16401	19341	35742	Sparrc1	SPARC-like 1 (mast9, hevin)
CATGTAAAAAAGAGGATGGG	15963	17769	33732	Ndr2	N-myc downstream regulated gene 2
CATGTAGCAAAAGAAAAGAC	18053	15461	33514	Cplx2	Complexin 2
CATGTTAATAAAGATATGCT	16952	16598	33550	Mafk1	Morfs family associated protein 1
CATGCAAAAATAAAGCGCA	12881	18090	30971	Eno1	Enolase 1, alpha non-neuron
CATGCTACTAACAGAGCTG	13001	17210	30211	Aldoa	Aldolase 1, A isoform
CATGCAAAATACATTTTGG	13545	16965	30510	Maged1	Melanoma antigen, family D, 1
CATGAGGAGAGACATCTGCT	14897	14380	29277	Eafr11	Eukaryotic translation elongation factor 1, alpha 1
CATGTAGTGAATAAAGGT	16248	13077	29325	Rps8	Ribosomal protein S8

2 samples analysed
mouse brain wt <> transgenic
~ 7.4 M tags per sample
range 36 K (0.5%) - 1



Human and Clinical Genetics

© JT den Dunnen LUMC

Result expression

Tag sequence	Gene-ID	sample A	sample B	Ratio	P-value
AATTTCCTCTCTCT	Msp1	2165	38	57.59	0
TATATTTTATTTAGAG	Cd771464	1071	7	154.82	0
ACAATGGGGATTTTTT	Pfig5	887	2783	0.32	0
AGAATAATATGAATT		849	383	2.24	0.04
TAAATGTTGCTAGCCTC		690	2	348.64	0
TCCACCTGCTTTCTTA	Mpp21	606	1458	0.42	0
GGGAGCGAAAAGTTAA	Id2	380	2	192	0
AGTGTGACGTGACCGGG	Mask3p3	323	1	326.39	0
AAAGACTGAGCTCATC	Ab1	315	1	318.31	0
ATGATAATGGACTAGCC					
ATTTGTACTATACAG	CAAAATACATTATTA	8	54	0.15	0
AGGAAGCCTCTCAAG	GGCTCGGCTCTTATGA	7	61	0.12	0
TCCACCGCAAAATA	TTCCAGCCTTACTATGT	7	36	0.2	0.03
GAATTCCTCATTGATT	CAAAATAGCAAGSGTGG	6	31	0.2	0.05
CCTGATGCTACAGAAA	AAAAATTTGTTTGTCT	5	32	0.16	0.02
AAAGATGTTAATTA	TATAGTATGTTGATTA	4	53	0.12	0.01
AAATTCCTCTCTCT	CAATTAAGTGAAGAC	4	26	0.16	0.04
CAAGTAAATTTCTTA		3	32	0.09	0.01
ACCCCTAAGAGACGA		3	24	0.13	0.04
CATCCTTGATTTAGCC		2	40	0.05	0
TCCACCTGTAAGGATA		2	20	0.1	0.05
CCTCTCCAGAAAGAG		1	22	0.05	0.01
CCATCTCAGSAGACTA		1	17	0.06	0.05
GGATTTGGCTGTATGA		0	29	0	0
CTCGTTTGTGATGA		0	23	0	0
CAGCGCACACAGAGCG	CJ230439	0	14	0	0.05



Human and Clinical Genetics

© JT den Dunnen LUMC

Expression profiling

- **easily interpretable results**
numbers
no cross-hybridization & background
- **differential expression**
4 orders of magnitude, more genes significant
- **additional to arrays**
47% genes differential polyA addition
51% genes anti-sense transcription
- **lower variation**
fewer replicates required
inter-lab data compare much better
- **array comparison**
limited overlap
all indicate disturbed GABA-ergic signaling
MPSS higher fold changes



Human and Clinical Genetics

© JT den Dunnen LUMC



Identification of SNPs In Turkey using massive parallel sequencing on the Solexa sequencing platform

Martien Groenen
Animal Breeding & Genomics Centre

ANIMAL SCIENCES GROUP
WAGENINGEN UR

Animal Breeding & Genomics Centre

Solexa sequencing: strategy

- Mix DNA from 6 individuals from 2 breeds
- DNA digested with *Sau3A*
- Separate on agarose gel
- Isolate 2000-4000 bp fraction
- Random shearing of fragments
- Isolate 200-250 bp fragments

ANIMAL SCIENCES GROUP
WAGENINGEN UR

Animal Breeding & Genomics Centre

Turkey / chicken

ANIMAL SCIENCES GROUP
WAGENINGEN UR

Animal Breeding & Genomics Centre

Turkey / chicken

ANIMAL SCIENCES GROUP
WAGENINGEN UR

Animal Breeding & Genomics Centre

SNP identification

GGAT19 5474 gtttctctgagtggttagcccaaggactgcaagt-gatctctc-gg

46773742 35 Sau3A

29142882 35 SNP

9132386 35 MAF ~ 0.2

50889210 34 >40,000 candidate SNPs

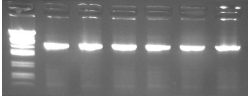
73393306 35
62913761 35
60250883 35
52125147 31
50086129 35
49263872 35
43190606 35
40722688 35
40434882 35
36035668 35
38994691 35
34687370 35
33758413 35
32688469 35
27353056 35


ANIMAL SCIENCES GROUP
WAGENINGEN UR

Animal Breeding & Genomics Centre

Deep sequencing

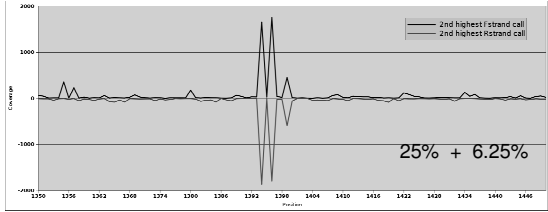
- **viral pool**
identify low-% variants
- **sensitivity test**
3.1 Kb fragment, plasmid cloned
PCR product
high-fidelity polymerase
gel-purified
fragmented
- **spike known variants**
25% (2), 6.25%, 1.56% (2),
0.42%, 0.10% (2)




© JT den Dunnen 

Deep sequencing

(low quality sequence run)

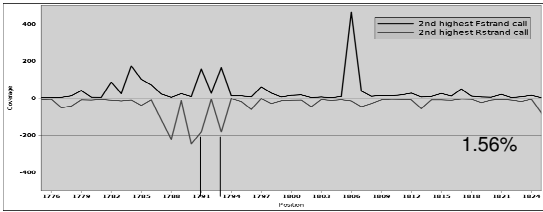


true variant: F + R
1394A>C + 1396A>C
1399C>T


© JT den Dunnen 

Deep sequencing

(low quality sequence run)

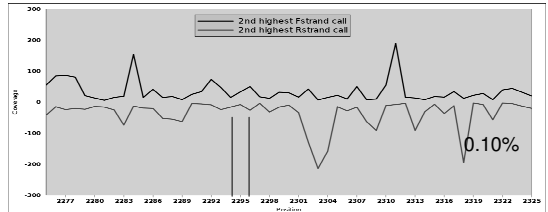


true variant: F + R
1791A>G + 1793C>T


© JT den Dunnen 

Deep sequencing

(low quality sequence run)



true variant: F + R
2294C>T + 2296T>C


© JT den Dunnen 

Targeted sequencing

(complexity reduction)


- **chromosome sorting**
- **gel separation**
Pulsed-Field Gel-electrophoresis
- **megabase regions**
cover by long-range PCR
1 Mb > 100 x 10Kb fragments
multiplex PCR
- **smaller regions**
pool samples
add sequence tag
- **array hyb-selection**

compare
controls <> cases

© JT den Dunnen 

Candidate gene

- **compare controls / cases**
- **long-range PCR** (8.5 Kb)
pool 176 DNA samples, then IrPCR
no reproducible result
very sensitive to DNA quality
- **pool 176 IrPCR products**
PCR, hrMCA check (yield, purity)
reproducible result
SNP freq. 1-2%
all variants detected (1/176) + more...
error rate 0.3 - 0.5%
(= 1/176 chromosomes)
redundancy > useful
- **pool tagged (Ir)PCR products**
...under investigation

© JT den Dunnen 

Genome re-sequencing

- **bacteria**
disease-related strains
large collections
drug resistance, virulence, infectivity, ...
diagnostic typing
genes involved
- **viruses**
- **moulds**
- ... (human patients)

Future ahead

Soon we will be able to sequence a complete human genome

A human genome

(www.LUMC.nl)

- **by academic hospital**
not a large genome center,
nor a company (sequence technology)
- **Marjolein Kriek**
PhD, clinical geneticist (i.t.)
first from LUMC,
Leiden,
Nederland,
Europe
female



A human genome

- **why ?**
show it is possible
technical, computational, analytical
to learn
technology, data floods, analysis
attractive project to tackle
- **results**
technically - no problem
computationally - at our limits
analytically - not possible
as expected
- >> **to be applied in patients**
resolve cause genetic disease

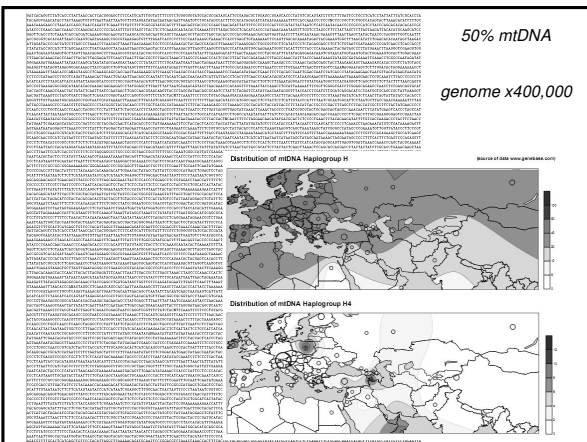


Analyse a genome



draw DNA-based conclusions

1. a female > no Y-chromosome sequences




Analyse a genome

- **red hair...**

suggested by all

literature
MC1R gene
no LSDB

many variants
 %change red hair

Human and Clinical Genetics 


Analyse a genome

- **blue eyes...**

suggested by all

literature
blue iris color
HERC2, OCA2
no LSDBs

many variants
 %chance iris colour
 analyse SNP profile

Human and Clinical Genetics 

Analyse a genome

The American Journal of Human Genetics 82, 411–423, February 2008 411

ARTICLE

Three Genome-wide Association Studies and a Linkage Analysis Identify *HERC2* as a Human Iris Color Gene


Manfred Kayser,^{1*} Fan Liu,² A. Cecilia J.W. Janssens,³ Fernando Rivadeneira,^{4,5} Oscar Lao,⁴ Erik van Duijn,⁶ Mark Vermeulen,⁶ Pascal Arp,⁴ Mila M. Ijzerman,⁴ Wilfried E.J. van Ijcken,⁴ Johan T. den Dunnen,⁶ Simon Heath,¹⁰ Diana Zelenika,¹⁰ Dominiek D.G. Despreet,^{4,7} Caroline C.W. Klaver,^{8,9} Johannes R. Vingeting,^{8,7} Paulus T.V.M. de Jong,^{8,11} Albert Hofman,⁸ Yurii S. Aulchenko,⁸ Andre G. Uitterlinden,^{8,9} Ben A. Oostra,⁷ and Cornelia M. van Duijn⁶

Abstract Iris color is one of the first traits for which Mendelian segregation was established. To date, the genetics of iris color is still not fully understood and is of interest, particularly in view of forensic applications. In three independent genome-wide association (GWA) studies of a total of 1406 persons and a genome-wide linkage study of 1292 relatives, all from the Netherlands, we found that the 15q13.1 region is the predominant region involved in human iris color. There were no other regions showing consistent genome-wide evidence for association and linkage to its color. Single nucleotide polymorphisms (SNPs) in the *HERC2* gene and, to a lesser extent, in the neighboring *OCA2* gene were independently associated to iris color variations. *OCA2* has been implicated in iris color previously. A replication study within two populations confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

Introduction Iris color is one of the first traits for which Mendelian segregation was established. To date, the genetics of iris color is still not fully understood and is of interest, particularly in view of forensic applications. In three independent genome-wide association (GWA) studies of a total of 1406 persons and a genome-wide linkage study of 1292 relatives, all from the Netherlands, we found that the 15q13.1 region is the predominant region involved in human iris color. There were no other regions showing consistent genome-wide evidence for association and linkage to its color. Single nucleotide polymorphisms (SNPs) in the *HERC2* gene and, to a lesser extent, in the neighboring *OCA2* gene were independently associated to iris color variations. *OCA2* has been implicated in iris color previously. A replication study within two populations confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

RESULTS We first performed a genome-wide linkage study of 1292 relatives from the Netherlands. We found a significant linkage signal on chromosome 15q13.1 (LOD score = 4.37, $P = 1.1 \times 10^{-5}$). This region contains the *HERC2* gene and the neighboring *OCA2* gene. We then performed a genome-wide association study (GWA) of 1406 unrelated individuals from the Netherlands. We found a significant association signal on chromosome 15q13.1 (lead SNP: rs11855019, $P = 1.1 \times 10^{-8}$). This region contains the *HERC2* gene and the neighboring *OCA2* gene. We then performed a replication study within two populations (Flemish and Dutch) and confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

CONCLUSIONS We found that the 15q13.1 region is the predominant region involved in human iris color. There were no other regions showing consistent genome-wide evidence for association and linkage to its color. Single nucleotide polymorphisms (SNPs) in the *HERC2* gene and, to a lesser extent, in the neighboring *OCA2* gene were independently associated to iris color variations. *OCA2* has been implicated in iris color previously. A replication study within two populations confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

Human and Clinical Genetics 

Analyse a genome

NCBI ENTREZ SNP

Search SNP for rs11855019

dbSNP BUILD 129

The following term was not found: rs11855019[All Fields]
 See Details: No items found

NCBI ENTREZ SNP

Search SNP for rs916977


dbSNP BUILD 129

Display: Graphic Summary | Show: 20 | Sort by: | Send to: |

Allele: 1 | ClinVar/LSDB Submissions: 0 | Human: 1 | Mouse: 0 |

rs916977 [Homo sapiens]

CCCCCTCCAGCCCTTGGCCAGCCCTCTCTTCTAATGCTGCTAACTCCATCCCATCTG

Human and Clinical Genetics 

Single Nucleotide Polymorphism

NCBI ENTREZ SNP

Search SNP on NCBI Reference Assembly

dbSNP BUILD 129

Reference SNP (refSNP) Cluster Report: rs4778138

rsSNP ID: rs4778138

Organism: human (*Homo sapiens*)

Molecule Type: Genomic

Created/Updated in build: 11/1/02


Map to Genome Build: 35.3

SNP Details are organized in the following sections:

Submitter records for this RefSNP Cluster

The submission rs17539455 has the longest flanking sequence of all cluster members and was used to annotate sequence

NCBI Assay ID	Probe/Assay ID	Validation Status	Orientation	Strand	Allele
rs5224556	WU_SSAHASNP01_01020119_739798				A/G
rs17539455	NCBI_HAFMAP/NCBI_HG00300402.chr15:147,010200_16_701311				A/G
rs42694779	ARIS_07_07_07_07_07				A/G
rs665511	ILLUMINA/HumanRefSeq001.1.rs11855019				A/G
rs665606	ILLUMINA/HumanRefSeq55001.1.rs11855019				A/G
rs7127697	ILLUMINA/HumanRefSeq55019.0.rs11855019				A/G
rs92171488	PERL/GENP/CP04220463				A/G
rs70401012	ILLUMINA/HumanRefSeq30092.0.rs4778138				A/G
rs10782704	ILLUMINA/HumanRefSeq55019.0.rs4778138				A/G
rs7491477	ILLUMINA/ILMN_Human_ILM144778138				A/G

Human and Clinical Genetics 

Analyse a genome

The American Journal of Human Genetics 82, 411–423, February 2008 411

ARTICLE

Three Genome-wide Association Studies and a Linkage Analysis Identify *HERC2* as a Human Iris Color Gene


Manfred Kayser,^{1*} Fan Liu,² A. Cecilia J.W. Janssens,³ Fernando Rivadeneira,^{4,5} Oscar Lao,⁴ Erik van Duijn,⁶ Mark Vermeulen,⁶ Pascal Arp,⁴ Mila M. Ijzerman,⁴ Wilfried E.J. van Ijcken,⁴ Johan T. den Dunnen,⁶ Simon Heath,¹⁰ Diana Zelenika,¹⁰ Dominiek D.G. Despreet,^{4,7} Caroline C.W. Klaver,^{8,9} Johannes R. Vingeting,^{8,7} Paulus T.V.M. de Jong,^{8,11} Albert Hofman,⁸ Yurii S. Aulchenko,⁸ Andre G. Uitterlinden,^{8,9} Ben A. Oostra,⁷ and Cornelia M. van Duijn⁶

Abstract Iris color is one of the first traits for which Mendelian segregation was established. To date, the genetics of iris color is still not fully understood and is of interest, particularly in view of forensic applications. In three independent genome-wide association (GWA) studies of a total of 1406 persons and a genome-wide linkage study of 1292 relatives, all from the Netherlands, we found that the 15q13.1 region is the predominant region involved in human iris color. There were no other regions showing consistent genome-wide evidence for association and linkage to its color. Single nucleotide polymorphisms (SNPs) in the *HERC2* gene and, to a lesser extent, in the neighboring *OCA2* gene were independently associated to iris color variations. *OCA2* has been implicated in iris color previously. A replication study within two populations confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

Introduction Iris color is one of the first traits for which Mendelian segregation was established. To date, the genetics of iris color is still not fully understood and is of interest, particularly in view of forensic applications. In three independent genome-wide association (GWA) studies of a total of 1406 persons and a genome-wide linkage study of 1292 relatives, all from the Netherlands, we found that the 15q13.1 region is the predominant region involved in human iris color. There were no other regions showing consistent genome-wide evidence for association and linkage to its color. Single nucleotide polymorphisms (SNPs) in the *HERC2* gene and, to a lesser extent, in the neighboring *OCA2* gene were independently associated to iris color variations. *OCA2* has been implicated in iris color previously. A replication study within two populations confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

RESULTS We first performed a genome-wide linkage study of 1292 relatives from the Netherlands. We found a significant linkage signal on chromosome 15q13.1 (LOD score = 4.37, $P = 1.1 \times 10^{-5}$). This region contains the *HERC2* gene and the neighboring *OCA2* gene. We then performed a genome-wide association study (GWA) of 1406 unrelated individuals from the Netherlands. We found a significant association signal on chromosome 15q13.1 (lead SNP: rs11855019, $P = 1.1 \times 10^{-8}$). This region contains the *HERC2* gene and the neighboring *OCA2* gene. We then performed a replication study within two populations (Flemish and Dutch) and confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

CONCLUSIONS We found that the 15q13.1 region is the predominant region involved in human iris color. There were no other regions showing consistent genome-wide evidence for association and linkage to its color. Single nucleotide polymorphisms (SNPs) in the *HERC2* gene and, to a lesser extent, in the neighboring *OCA2* gene were independently associated to iris color variations. *OCA2* has been implicated in iris color previously. A replication study within two populations confirmed that the *HERC2* gene is a new and significant determinant of human iris color variation. In addition to *OCA2*, furthermore, *HERC2* rs916977 showed a clinal allele distribution across 23 European populations, which was significantly correlated to iris color variation. We suggest that genetic variants regulating the expression of the *OCA2* gene exist in the *HERC2* gene or, alternatively, within the 11.2 kb of sequence between *OCA2* and *HERC2*, and that most iris color variations in Europeans is explained by these two genes. Testing markers in the *HERC2*-*OCA2* region may be useful in forensic applications to predict eye color phenotypes of unknown persons of European genetic origin.

Human and Clinical Genetics 

A human genome

...what do all these variants mean ??

- few tools available
- data scattered over the web
*databases, incl. LSDBs
many formats
little phenotype information*
- most data in drawers
- ...tools are there
gene variant databases (LSDBs)
- change in attitude required



Human and Clinical Genetics

© JT den Dunnen L U M C



Johan vd Leij

Future ahead

Soon we will be able to sequence a complete human genome,

but if we can not make sense out of the variants detected, as to whether they are "pathogenic or not", this information is useless and misinterpretation....



Human and Clinical Genetics

© JT den Dunnen L U M C

Gene variant databases

Submit all the changes you have, NOW

(without errors)



Human and Clinical Genetics

© JT den Dunnen L U M C

Summary

Sequencing has the future

*As clinical lab,
do not buy a system yet,
use that of your colleague,
but start saving money,
in 3-5 years you need it...*



Human and Clinical Genetics

© JT den Dunnen L U M C