Interpretation of Genomic Copy Number Variants Using DECIPHER

Manuel Corpas,[1] Eugene Bragin,[1] Stephen Clayton,[1] Paul Bevan,[1] and Helen V. Firth[1,2]

[1]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom
[2]Cambridge University Hospitals Foundation Trust, Addenbrooke's Hospital, Cambridge, United Kingdom

ABSTRACT

Many patients suffering from developmental disorders have submicroscopic deletions or duplications affecting the copy number of dosage-sensitive genes or disrupting normal gene expression. Many of these changes are novel or extremely rare, making clinical interpretation problematic and genotype/phenotype correlations difficult. Identification of patients sharing a genomic rearrangement and having phenotypes in common increases certainty in the diagnosis and allows characterization of new syndromes. The DECIPHER database is an online repository of genotype and phenotype data whose chief objective is to facilitate the association of genomic variation with phenotype to enable the clinical interpretation of copy number variation (CNV). This unit shows how DECIPHER can be used to (1) search for consented patients sharing a defined chromosomal location, (2) navigate regions of interest using in-house visualization tools and the Ensembl genome browser, (3) analyze affected genes and prioritize them according to their likelihood of haploinsufficiency, (4) upload patient aberrations and phenotypes, and (5) create printouts at different levels of detail. By following this protocol, clinicians and researchers alike will be able to learn how to characterize their patients' chromosomal imbalances using DECIPHER. *Curr. Protoc. Hum. Genet.* 72:8.14.1-8.14.17. © 2012 by John Wiley & Sons, Inc.

Keywords: copy number variation • clinical genetics • array CGH • genotype • phenotype • developmental disorders • bioinformatics

INTRODUCTION

Array comparative genomic hybridization (array CGH) has revolutionized the capacity to identify microdeletions and microduplications, significantly improving the prospect of diagnosis for patients with syndromic and non-syndromic mental retardation (Stankiewicz and Beaudet, 2007). As next generation sequencing gradually transfers from a research to a clinical setting, further improvement in the ability to diagnose rare genetic disorders is anticipated. The DECIPHER database is a resource specifically designed for collaboration among clinicians by collecting potentially pathogenic CNVs and phenotypes and displaying them in the context of the human genome reference sequence and annotation data (Firth et al., 2009). The DECIPHER Consortium was initiated in 2004 as a community of academic centers of Clinical Genetics who submit consented, anonymized genotype and phenotype data from patients with rare genomic disorders for sharing with other clinicians and researchers. DECIPHER has now become an established international resource with >11,300 patients and 250 contributing centers in more than 30 countries (October 2011). Among its most widely utilized tools are: direct searching for variants, phenotypes, and syndromes overlapping with an index case; prioritization of genes likely to be responsible for the phenotype; and visualization of variant data in graphs and major genome browsers (Flicek et al., 2010; Sanborn et al., 2010).

Current Protocols in Human Genetics 8.14.1-8.14.17, January 2012
Published online January 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/0471142905.hg0814s72

Consistency in patient phenotype descriptions is ensured using a controlled vocabulary, namely the London Dysmorphology and Neurogenetics Database (LDNB) (Winter and Baraitser, 1987; Baraitser et al., 1989). LDNB offers a hierarchically organized set of 1700 phenotypic terms with a strong neurological component. Many of these terms can be reliably mapped to the Human Phenotype Ontology (HPO) (Robinson et al., 2008).

How DECIPHER works

An international community of academic clinical genetics units contribute to the DECIPHER database. Each contributing center has a minimum of one senior clinical geneticist responsible for local governance of the project and for phenotype entry, and a molecular cytogeneticist responsible for genotype data entry. Access to the database is provided by password-protected accounts. To share anonymized patient data with other centers, informed consent is required to allow genotype and phenotype data to become visible via the DECIPHER search interface, Ensembl (*UNIT 6.11*), and other genome browsers, e.g., UCSC Genome Browser (*UNIT 18.6*). Data shared anonymously in this way allows access by other users for analysis and facilitates discovery of new microdeletion/microduplication syndromes. Moreover, a collection of 59 well-annotated established microdeletion/microduplication syndromes is available for reference with links to further information and related resources. DECIPHER public users (i.e., users with no login account) may search, browse, and access anonymized aberrations and phenotype terms from consented patients in a variety of ways (see below).

The most widely used functionalities of DECIPHER is described in this unit, first introducing public access to the database and then the facilities available to a logged-in consortium member.

BASIC PROTOCOL 1

SEARCHING DECIPHER BY LOCATION AS A PUBLIC USER

Nadia is a 4-year-old girl, the second child of healthy unrelated parents. At the age of 4 months, psychomotor delay, brachycephaly, and severe muscular hypotonia were noted by a pediatrician. Her developmental milestones were delayed and when evaluated in the clinical genetics department she presented severe learning disability and developmental delay. She has an apparently normal female karyotype, 46,XX. A genomic array analysis was performed and a deletion was found in chromosome 5, start position 85,030,173 and end position 89,050,512 (GRCH37/hg19 genome assembly).

To establish whether this region is the one causing the observed phenotype, use the DECIPHER database to confirm the diagnosis and determine whether there are similar patients in the database displaying the same symptoms.

View DECIPHER homepage

1. Access the DECIPHER Website (*http://decipher.sanger.ac.uk/*), which shows the homepage with a range of overview information via active vertical tabs. By default, these vertical tabs show a map of the world with the location of center members of the DECIPHER consortium (see Fig. 8.14.1). Click on any of the pins in the map to show the name of the center to which it refers.

   This is useful to learn the location of the current contributing DECIPHER centers and whether an established center already exists nearby.

2. Click on News to learn about the latest developments in DECIPHER.

3. Click on Overview to read some general information about DECIPHER.

4. Click on Patients to learn about the statistics of the project. On July 30, 2011, there were a total of 10,408 patients (5104 consented), 59 syndromes, and 233 participating centers.

Figure 8.14.1 Vertical tabs on DECIPHER homepage show by default an interactive map with the location of all members of the consortium.

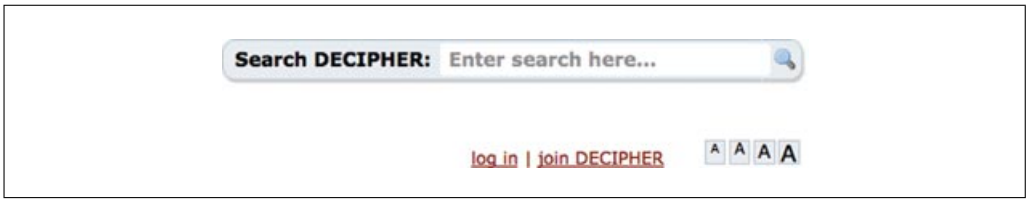


Figure 8.14.2 DECIPHER search box. This input box is visible on every page of DECIPHER's interface. Public users can search by phenotype, chromosomal location, gene name, or band.

5. Click on DDD UK to obtain information about the UK Deciphering Developmental Diseases Project (Firth and Wright, 2011).

Search desired chromosome

6. Locate the search box at the top right corner of the screen (Fig. 8.14.2).
7. In the search box, paste 5:85030173-89050512 and hit return.

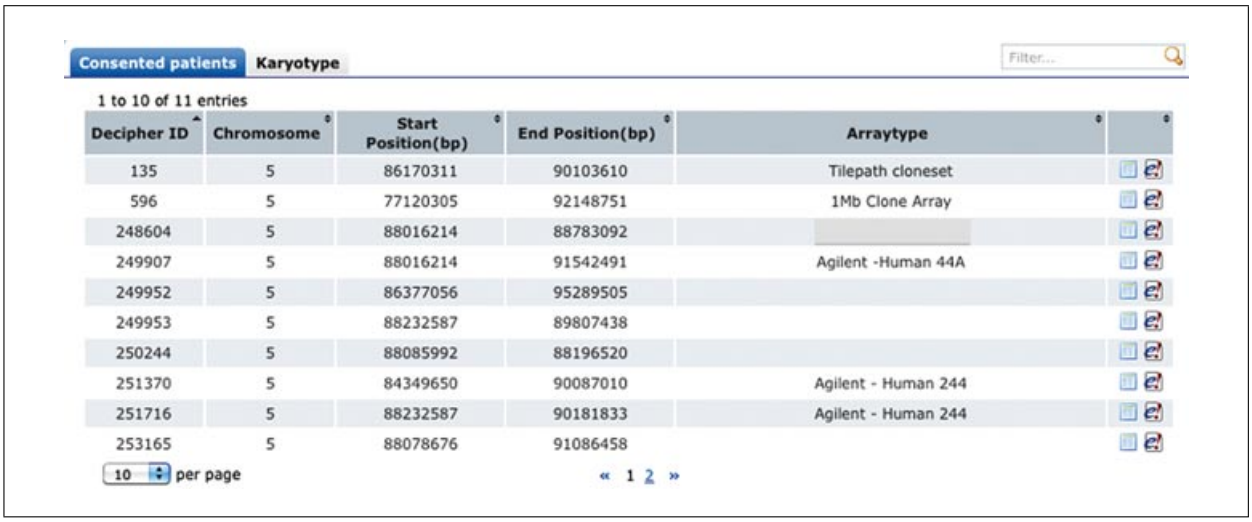
   When doing a search as a public user all consented patients and syndromes in DECIPHER are searched. Any of the patients overlapping this region in chromosome 5 will be returned in the results hit list.

8. Examine the results. At the time of this writing, this query returned 11 consented patients and 0 syndromes overlapping the searched region (Fig. 8.14.3).
9. Click on the *e!* icon to display in Ensembl the region in the top row, patient 135.
10. Once the Ensembl page is loaded, to make the DECIPHER track visible, click on the Configure This Page (Fig. 8.14.4) of the Ensembl redirected page.

    The current version of Ensembl provides DECIPHER data in GRCh37 coordinates. NCBI36 genome coordinate data are still available via the Ensembl archive on a static form, containing DECIPHER data deposited before February 16, 2011. To access DECIPHER data in NCBI36 coordinates, activate the DECIPHER track in the Ensembl archive page.

View DECIPHER data

11. View the pop-up window that appears (Fig. 8.14.5) when clicking on Configure This Page. Place the cursor on the top right search box and start typing "decipher". The DECIPHER data source appears with an empty checkbox.
12. Click on the empty checkbox and select Normal.



Consented patients | Karyotype | Filter...

1 to 10 of 11 entries

Decipher ID	Chromosome	Start Position(bp)	End Position(bp)	Arraytype
135	5	86170311	90103610	Tilepath cloneset
596	5	77120305	92148751	1Mb Clone Array
248604	5	88016214	88783092	
249907	5	88016214	91542491	Agilent -Human 44A
249952	5	86377056	95289505	
249953	5	88232587	89807438	
250244	5	88085992	88196520	
251370	5	84349650	90087010	Agilent - Human 244
251716	5	88232587	90181833	Agilent - Human 244
253165	5	88078676	91086458	

10 per page « 1 2 »

Figure 8.14.3 Results retrieved from searching DECIPHER by location. The typed location was 5:85030173-89050512 and 11 consented patients and no syndromes are found to overlap with region.

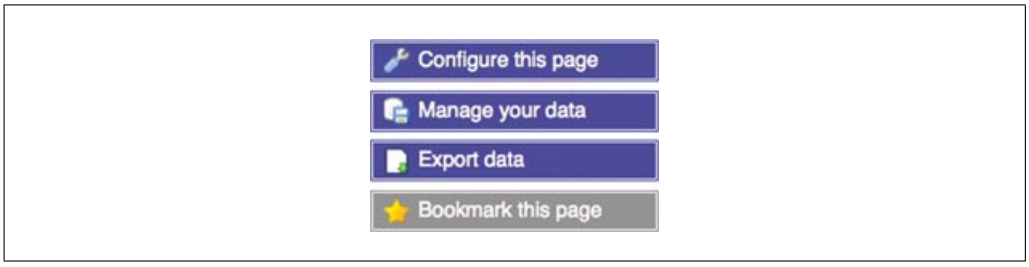


Figure 8.14.4 Click on Configure this page to add the DECIPHER track to Ensembl.

13. Click on the top right tick to save the new configuration. Wait until the data are loaded and look for the DECIPHER track in the window (Fig. 8.14.6).

14. Examine all the tracks. Note that red bars denote losses and blue bars denote gains. Because Nadia has a loss, click on the top red bar to find a patient with a loss in that region.

    When any DECIPHER bar in Ensembl is clicked, a pop-up window appears with the start and end position and classification of the genomic aberration in the affected patient (e.g., de novo or inherited).

15. Compare the start and end positions of the loss detected in Nadia with the start and end position for patient 135.

16. Click on the link of the pop-up window to view the patient report page (Fig. 8.14.7).

    The patient report is composed of two rows of tabs in the page. By default the aberration tab and the first aberration on the list of affected regions are shown. For this aberration, related information about its location, interval copy number, and links to UCSC and Ensembl genome browsers are provided.

Gather more information

17. To gather more information about this patient and links to further resources, click on the Overview tab.

    The information in the Overview tab provides some general background information about the patient.

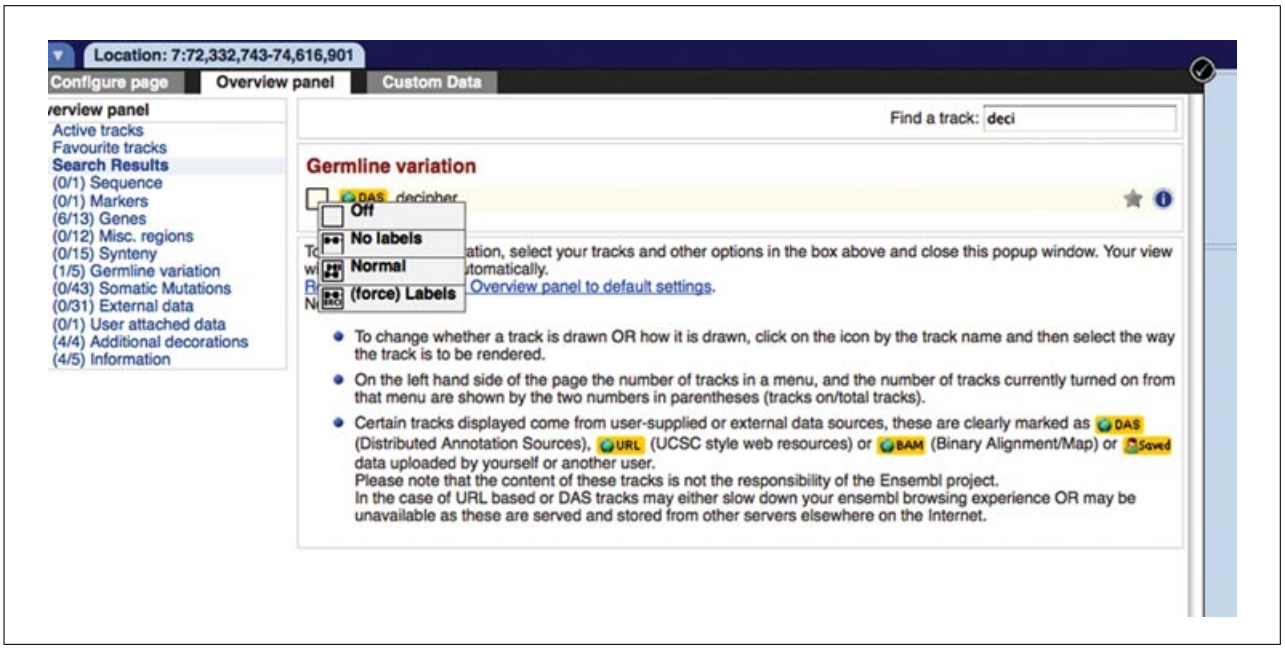


Figure 8.14.5 Interface in Ensembl to add the DECIPHER source. In the top right corner, type "DECIPHER" and immediately the DECIPHER data source appears. Click on the empty checkbox and a pop-up window appears. Click on Normal and the click on the top right tick to save your configuration.

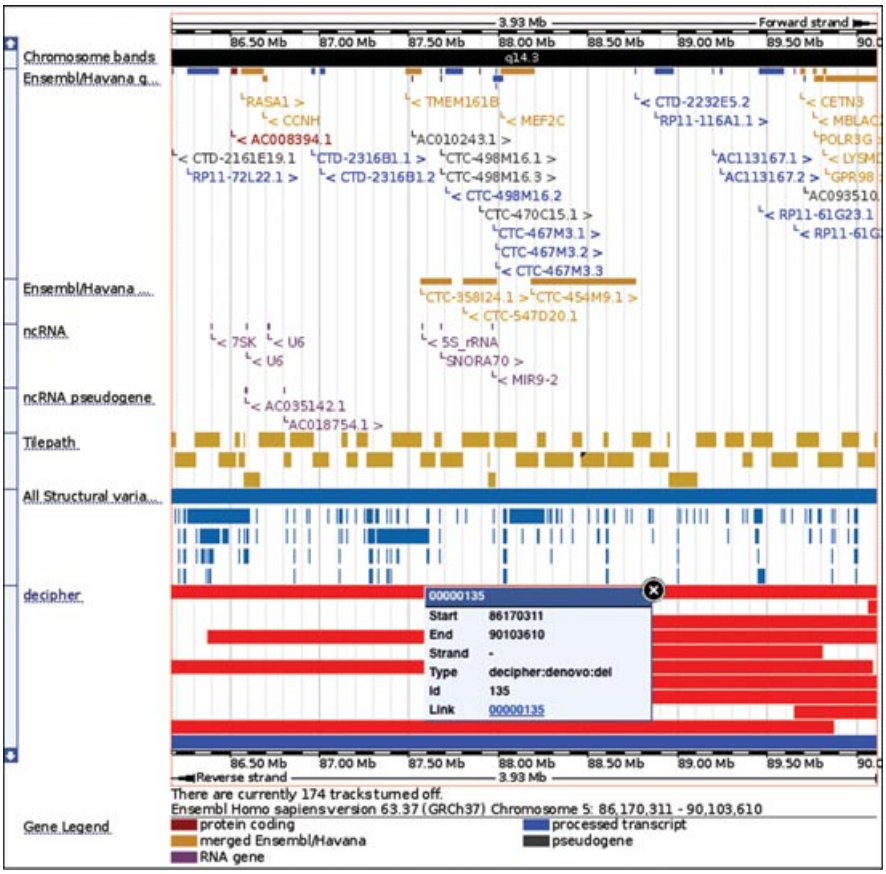


Figure 8.14.6 DECIPHER track as shown in Ensembl for the 5:85030173-89050512 region. Red and blue correspond to losses and gains, respectively. Any bar in the Ensembl track contains links back to their DECIPHER entry. Only consented patient data are displayed. In this example, the feature bar for DECIPHER patient 135 is clicked and shows a pop-up window with more detailed information about the aberration and a link back to the DECIPHER database.

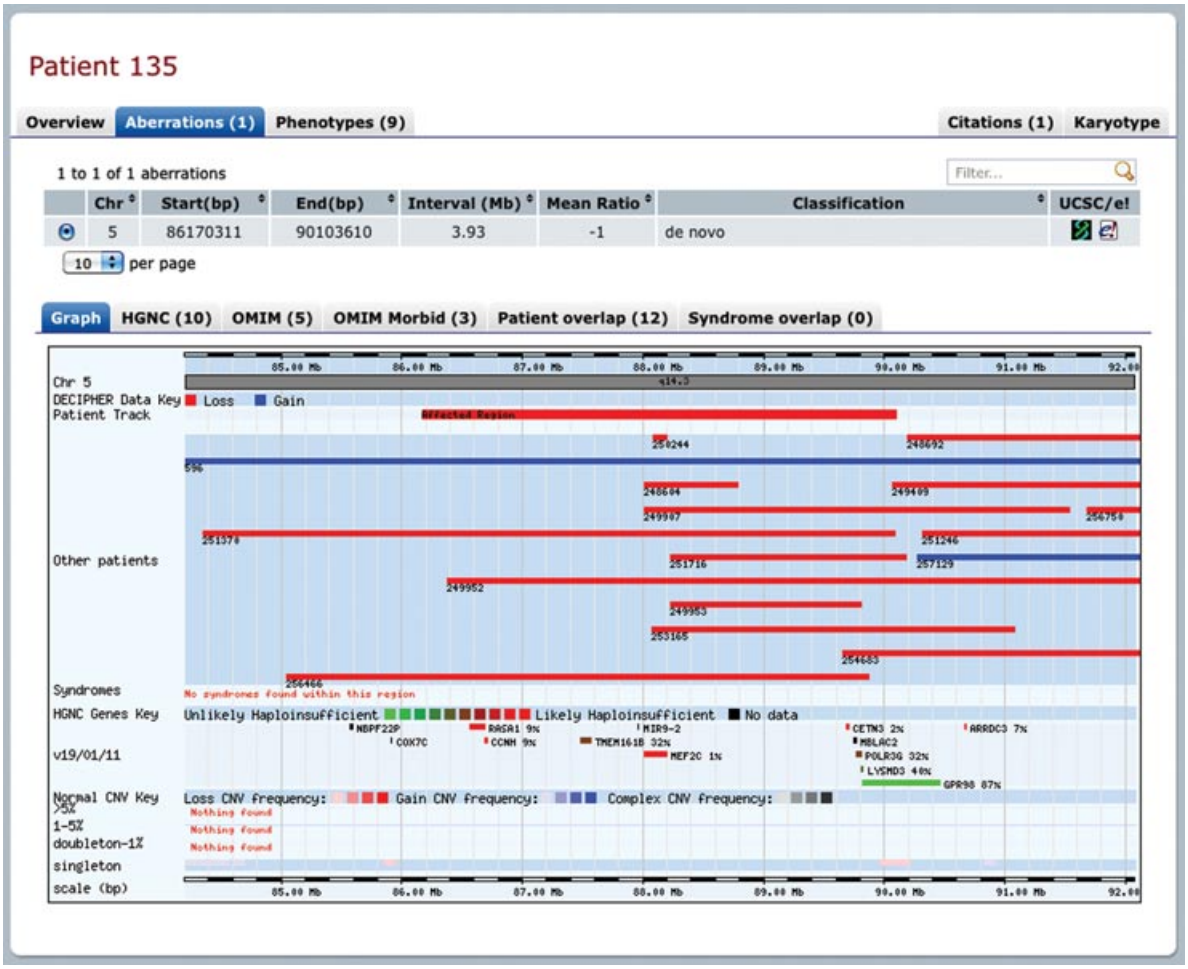


Figure 8.14.7 Typical patient report. A series of tabs are provided describing the deposited information for this patient in DECIPHER. The top tabs refer to the patient and the second level tabs refer to each aberration. Each patient has an Overview, Aberration, Phenotypes, Citations, and Karyotype tab. Each aberration contains a graph, a list of HGNC genes, OMIM genes, OMIM morbid genes, Patient overlap, and a Syndrome overlap tab.

18. Compare observed phenotypes in Nadia with the ones reported in the phenotype tab.

19. Look at the Citations tab for a list of relevant publications.

    A publication is shown referring to a novel microdeletion syndrome involving 5q14.3-q15 (Engels et al., 2009).

20. Click again on the Aberrations tab to find more information about the affected genes and overlapping patients.

    By default, the Graph sub-tab is selected, showing overlapping consented patients, syndromes, and genes (Fig. 8.14.7). Relative locations of overlapping genes are also provided. Those genes where haploinsufficiency predictions are available appear colored in a scale from green to red, from least to most probability of haploinsufficiency (Huang et al., 2010).

    A number of tracks are available within the Graph provided. At the top, a scale bar contains the relative location within the zoomed-in chromosome. Below that, bands are also drawn and aberrations for consented overlapping patients are shown. Underneath it, a track shows overlapping syndromes. All overlapping Ensembl genes with their HGNC approved names (Bruford et al., 2008) are then shown. These genes are presented as colored bars according to their haploinsufficiency score mapped to the GRCh37 human

assembly. The band at the bottom of the graph shows a CNV consensus track that combines a series of normal CNV datasets observed in the 1000 genomes project (Mills et al., 2011) and other published sources (Conrad et al., 2009). Four bands of frequency in control populations are shown in the CNV consensus track; >5%, 1% to 5%, doubletons to 1%, and singletons. Colors of CNV bars are composed of different shades of blue for gains, red for losses, and gray for complex (regions reported to be both gained or lost in individuals).

21. Click on the HGNC tab to obtain a report of all genes overlapping patient 135, in this case, ten genes in total. Click on any of the gene table headers to sort rows by that criterion.

    The gene nomenclature is provided by HGNC and the coordinates based on Ensembl mappings. Each gene name is linked to its corresponding entry in Ensembl.

22. Click once on the %HI (haploinsufficiency scores) to sort genes according to their predicted likelihood of being haploinsufficient.

    Figure 8.14.8 shows genes sorted by their haploinsufficiency scores. A good starting point is to consider those genes whose scores are in the top 10% of being likely to be haploinsufficient as possible candidates responsible for the phenotype. Haploinsufficiency scores can be of great help as a way to prioritize genes for further study.

23. Find out the predicted most haploinsufficient genes (those with the top percentile score or the most red in the list). Among the most haploinsufficient genes, identify those that appear to have an entry in OMIM and OMIM morbid.

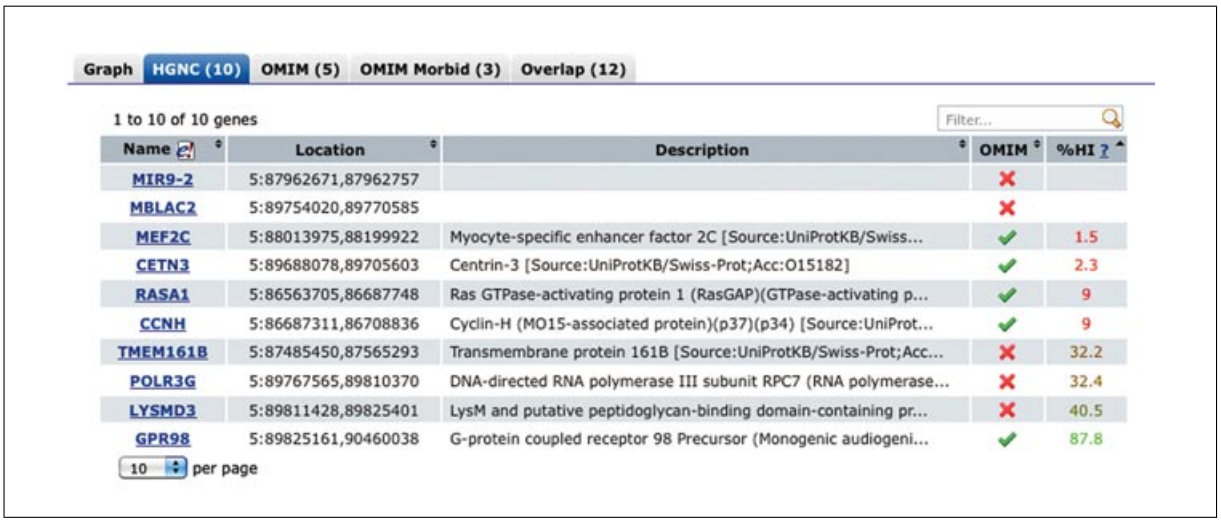


Figure 8.14.8 A total of ten genes have been found to be contained in the deleted region for patient 135. Genes are ordered according to their haploinsufficiency score (%HI). Those with no score come first and then the most likely to be haploinsufficient. The most haploinsufficient gene is predicted to be MEF2C.

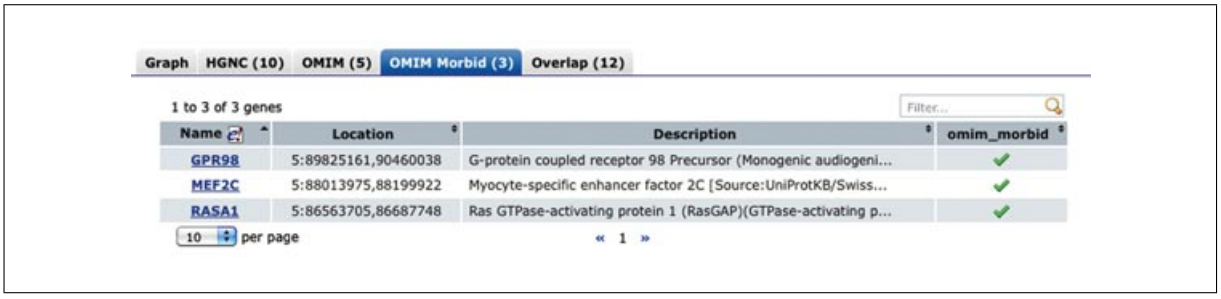


Figure 8.14.9 GPR98, MEF2C, and RASA1 are identified as disease causing by OMIM Morbid among the list of overlapping genes in the 5:86170311-90103610 region.

For patient 135 a total of four genes are predicted as likely haploinsufficient: MEF2C, CETN3, RASA1, and CCNH. Note that a number of them appear blank, i.e., no haploinsufficiency score is available for those genes. In DECIPHER, OMIM and OMIM morbid genes are a subset of the official HGNC gene list. OMIM morbid, in particular, is the most useful from a clinical perspective because it is a catalogue of known disease genes. When clicking on any gene name, a new window appears with its corresponding entry in its source database.

GPR98, MEF2C, and RASA1 appear to be the only genes among the ten affected in patient 135 that have been listed as disease-causing genes in OMIM morbid (Fig. 8.14.9). MEF2C has recently been shown to be the critical gene underlying the 5q14.3q15 microdeletion syndrome (Jaillard et al., 2010; Le Meur et al., 2010; Zweier et al., 2010) giving rise to severe mental retardation, seizure disorders, and a common set of facial features.

BASIC PROTOCOL 2

SEARCHING DECIPHER FOR A PHENOTYPE AS A PUBLIC USER

For Nadia's case, the phenotype brachycephaly is a defining morphological feature, which is also matched by patient 135. To check whether there are other consented patients in the DECIPHER database that have this phenotype in the region 5q14, the following protocol may be carried out.

1. Go to the DECIPHER home page (*https://decipher.sanger.ac.uk*).
2. Type "brachycephaly" in the top right search box and hit enter.
3. Examine in the result table the phenotypes matched and the number of aberrations (Fig. 8.14.10).
4. Click on the Karyotype tab to see the distribution of patients who have brachycephaly in their list of annotated phenotypes (Fig. 8.14.11).

   A total of 78 aberrations are drawn in the ideogram, denoted in blue (gain) and red (losses). Below it a table listing all aberrations is provided with a filter box.

5. Locate the Aberration list table below the ideogram and type "5q14" in the filter box. This reduces the list of aberrations to two. These are the only two aberrations whose patients have brachycephaly as an annotated phenotype in this region.

   Any of the listed aberrations in the Aberration table can be clicked for access to the corresponding patient report.

6. The top aberration corresponds to patient 135, which we have seen above. Click on the second aberration (arr 5q14.3(86,304,862-89,692,876)del) to reach the report for another affected patient (patient 249546) with brachycephaly in the same 5q14 region.
7. Examine the graph for patient 249546.
8. Find the OMIM Morbid tab and compare it with the phenotypes of Nadia and patient 135.
9. Compare matched genes between patients 135 and 249546.

   Patient 249546 matches two of the three genes (MEF2C and RASA1) that were reported in the OMIM Morbid tab for patient 135.

10. Check in the phenotype tab that *brachycephaly* is among its reported phenotypes.
11. Explore other tabs to find more information about patient 249546.

    For historical reasons, DECIPHER uses mean ratio or copy number to define whether an aberration is a loss (mean ratio <0 or copy number <2) or a gain (mean ratio >0 or copy number >2).

12. Click on the Patient Overlap tab.

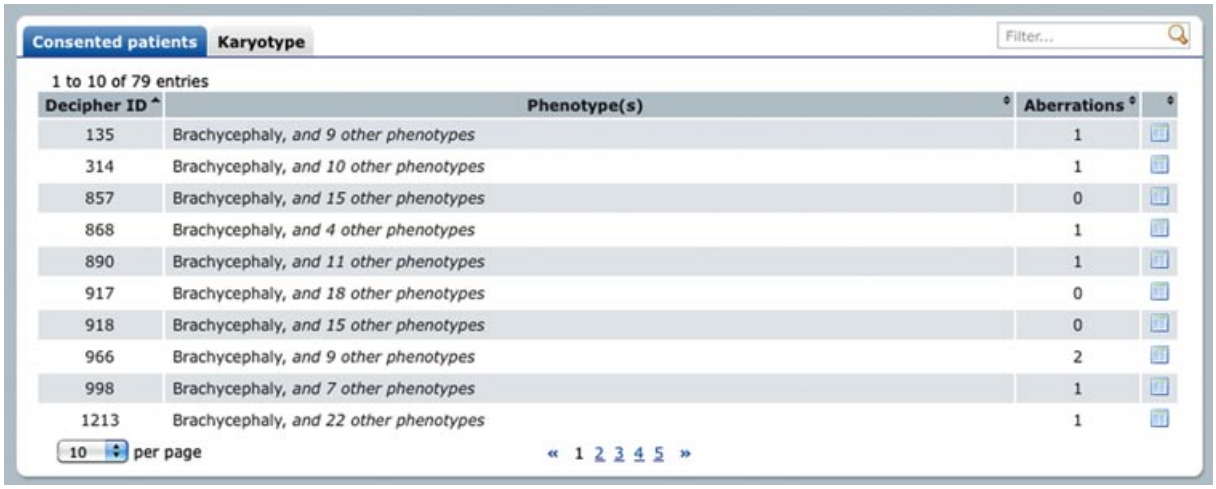


Figure 8.14.10 Search results for patients matching brachycephaly. A total of 79 consented patients have this phenotype. Any of these can be accessed for further inspection by clicking on the corresponding row.

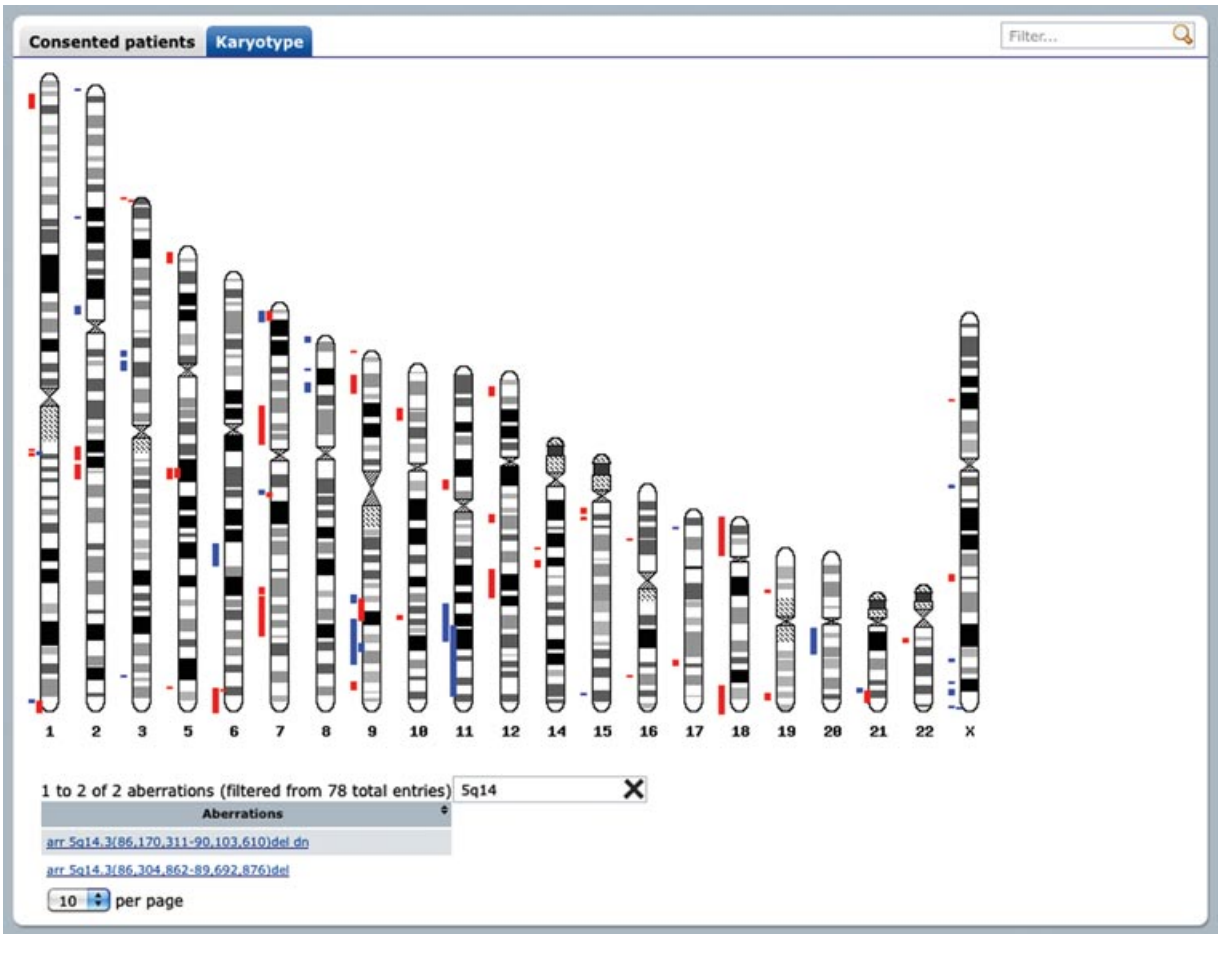
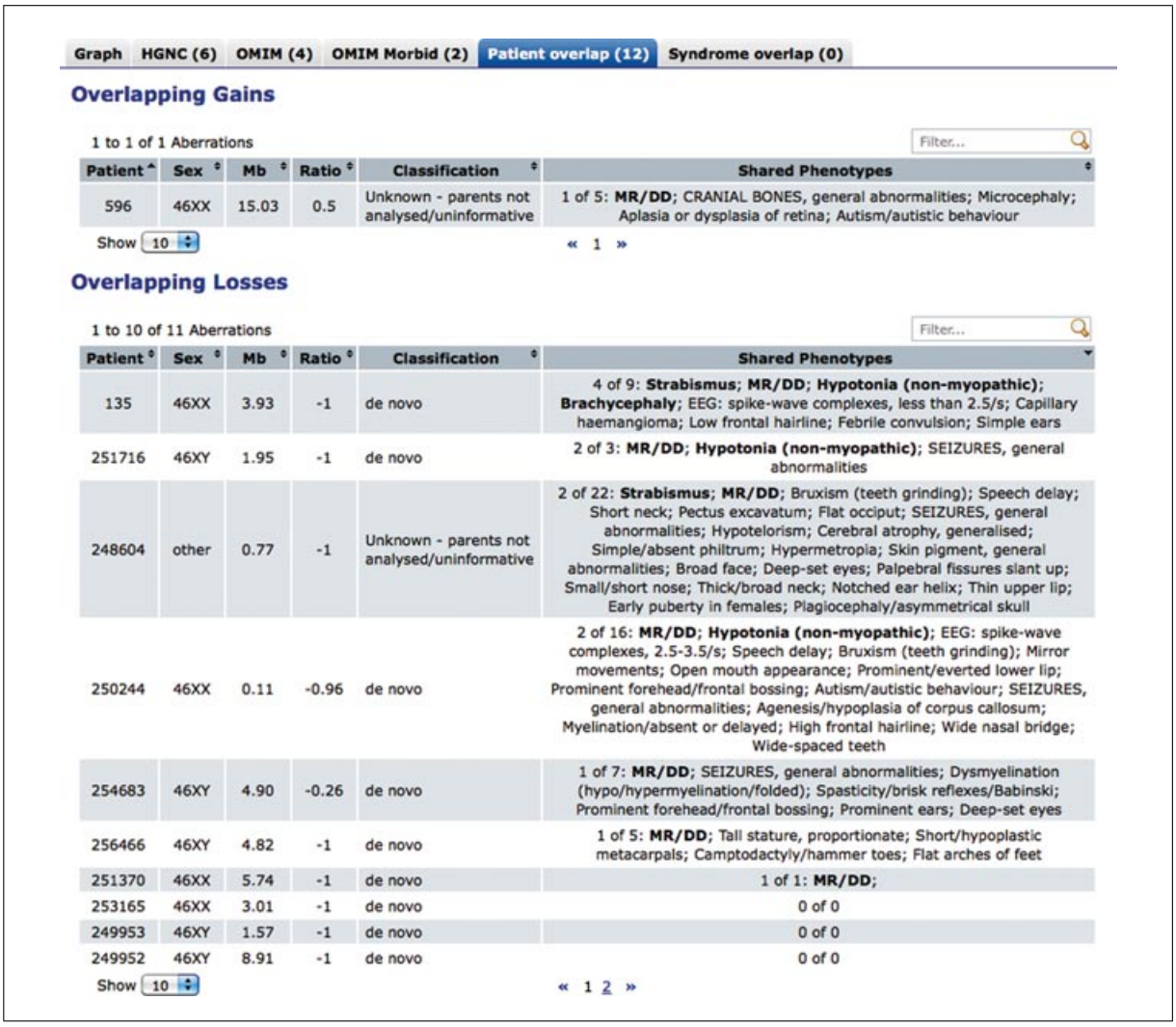


Figure 8.14.11 Karyotype showing all regions affected for patients with the brachycephaly phenotype. A total number of 78 aberrations are displayed in the graph. Below it, a table shows all matched aberrations listed. A filter box is provided allowing further refinement of results. 5q14 was typed in the filter box, reducing the list to two aberrations. Clicking on any of these listed aberrations the user is redirected to its patient report.



Graph | HGNC (6) | OMIM (4) | OMIM Morbid (2) | Patient overlap (12) | Syndrome overlap (0)

Overlapping Gains

1 to 1 of 1 Aberrations

Filter...

Patient	Sex	Mb	Ratio	Classification	Shared Phenotypes
596	46XX	15.03	0.5	Unknown - parents not analysed/uninformative	1 of 5: **MR/DD**; CRANIAL BONES, general abnormalities; Microcephaly; Aplasia or dysplasia of retina; Autism/autistic behaviour

Show 10 « 1 »

Overlapping Losses

1 to 10 of 11 Aberrations

Filter...

Patient	Sex	Mb	Ratio	Classification	Shared Phenotypes
135	46XX	3.93	-1	de novo	4 of 9: **Strabismus**; **MR/DD**; **Hypotonia (non-myopathic)**; **Brachycephaly**; EEG: spike-wave complexes, less than 2.5/s; Capillary haemangioma; Low frontal hairline; Febrile convulsion; Simple ears
251716	46XY	1.95	-1	de novo	2 of 3: **MR/DD**; **Hypotonia (non-myopathic)**; SEIZURES, general abnormalities
248604	other	0.77	-1	Unknown - parents not analysed/uninformative	2 of 22: **Strabismus**; **MR/DD**; Bruxism (teeth grinding); Speech delay; Short neck; Pectus excavatum; Flat occiput; SEIZURES, general abnormalities; Hypotelorism; Cerebral atrophy, generalised; Simple/absent philtrum; Hypermetropia; Skin pigment, general abnormalities; Broad face; Deep-set eyes; Palpebral fissures slant up; Small/short nose; Thick/broad neck; Notched ear helix; Thin upper lip; Early puberty in females; Plagiocephaly/asymmetrical skull
250244	46XX	0.11	-0.96	de novo	2 of 16: **MR/DD**; **Hypotonia (non-myopathic)**; EEG: spike-wave complexes, 2.5-3.5/s; Speech delay; Bruxism (teeth grinding); Mirror movements; Open mouth appearance; Prominent/everted lower lip; Prominent forehead/frontal bossing; Autism/autistic behaviour; SEIZURES, general abnormalities; Agenesis/hypoplasia of corpus callosum; Myelination/absent or delayed; High frontal hairline; Wide nasal bridge; Wide-spaced teeth
254683	46XY	4.90	-0.26	de novo	1 of 7: **MR/DD**; SEIZURES, general abnormalities; Dysmyelination (hypo/hypermyelination/folded); Spasticity/brisk reflexes/Babinski; Prominent forehead/frontal bossing; Prominent ears; Deep-set eyes
256466	46XY	4.82	-1	de novo	1 of 5: **MR/DD**; Tall stature, proportionate; Short/hypoplastic metacarpals; Camptodactyly/hammer toes; Flat arches of feet
251370	46XX	5.74	-1	de novo	1 of 1: **MR/DD**;
253165	46XX	3.01	-1	de novo	0 of 0
249953	46XY	1.57	-1	de novo	0 of 0
249952	46XY	8.91	-1	de novo	0 of 0

Show 10 « 1 2 »

Figure 8.14.12 Overlapping gains and losses for patient 249546. Losses are sorted in descending order by number of shared phenotypes. Note that shared phenotypes are highlighted in bold while the rest are in normal font.

13. In the Overlapping Losses table, click on the Shared Phenotypes heading to sort overlapping losses from greatest number of shared phenotypes to lowest (Fig. 8.14.12).

14. Because the shortest interval entry (patient 250244) is also classified as de novo, click on it to obtain further information.

    Patient 250244 has a deletion spanning only gene MEF2C and it has 16 phenotype terms, many of them similar to Nadia. At this point, decide whether to contact DECIPHER (decipher@sanger.ac.uk) to request further information about this patient. DECIPHER will then contact the submitting center who may, with the permission of the patient/family, contact you with further phenotypic details. In order to protect patient anonymity, a public user does not know which center in which country submitted the patient data until the point at which you receive contact from the submitting center.

BASIC PROTOCOL 3

ADDING A PATIENT TO DECIPHER AS A REGISTERED MEMBER

Membership in the DECIPHER Consortium provides a number of the following privileges to users above the free public access to the consented data. Registering with DECIPHER allows one the (1) ability to add patient data into the DECIPHER database to make use of DECIPHER's bioinformatics tools for interpretation of patient aberration data; (2) ability to share with the community linked-anonymized patient data (provided that the patient

has signed a consent form). This allows other clinicians/researchers to see this patient data so that new syndromes can be defined. (3) Access to unconsented patient data from the DECIPHER study of which they are a member is available. (4) Access to contact details for the responsible clinician in any deposited consented patient entry is allowed. This facilitates direct communication between consortium members and exchange of expert information. (5) Printing of DECIPHER patient reports is available. (6) Receipt of DECIPHER's monthly newsletter, informing users of the latest developments, and relevant annual events, such as the DECIPHER Symposium or the Genomic Disorders conference is provided.

Once logged into DECIPHER, a new tab called My Patients is seen.

As an example, Nathalie is a 5-year-old girl who was seen in the genetics clinic. She has osteopoikilosis and shows signs of developmental delay. She is in the 10^{th} percentile for height. She has an apparently normal female karyotype, 46,XX. A genomic array analysis has been carried out with the following results.

Nathalie was found to have a deletion on chromosome 12 starting at 64081319 and ending at 68947027. This deletion was verified with semi-quantitative FISH and testing of both parents, which revealed that it is de novo. Recently, an account in DECIPHER has been set up with write access and this information has been entered for further analysis and contrast with existing entries in the database.

Access DECIPHER

1. Click on My Patients.

   Clicking on this tab lists all studies and patients to which the user has access. Figure 8.14.13 shows an overview of all patients in the TGD study (dummy test data).

   Any of the fields in the tables can be sorted according to the table heading clicked. For example, clicking on the Lab ID heading sorts the patient with Lab ID 1 to the top row in the table. This can be a useful way to filter patients.

2. To add your patient data, click on Add Patient. A new page appears with the Overview tab selected and empty fields to be filled (Fig. 8.14.14).

3. Enter patient details with a unique Patient Lab ID (often the pedigree number of patient reference number) and Chromosomal Sex.

   When entering the Patient Lab ID, make sure that it is completely unique to this study; duplicate lab IDs for different patients in the same study will not be accepted.

4. Check the box indicating that you have consent.

   DECIPHER consent implies that a consent form from the Documents tab (Fig. 8.14.15) has been signed and understood by the patient or parent/guardian. Alternatively, for incompetent adults over 18 years, assent forms are provided for the parent or guardian to be signed on behalf of the patient. Consent in DECIPHER allows your patient data to be accessible via the search tool, aberration graphs, or lists of overlapping patients. The anonymized aberration and phenotypes will be immediately available in Ensembl and are regularly updated in the UCSC Genome Browser.

5. Click Create.

6. If consent is obtained, add a photograph by clicking Upload New Picture.

Add aberrations

7. Click on Aberrations Tab.

8. Click on Add Aberration to enter coordinate data.

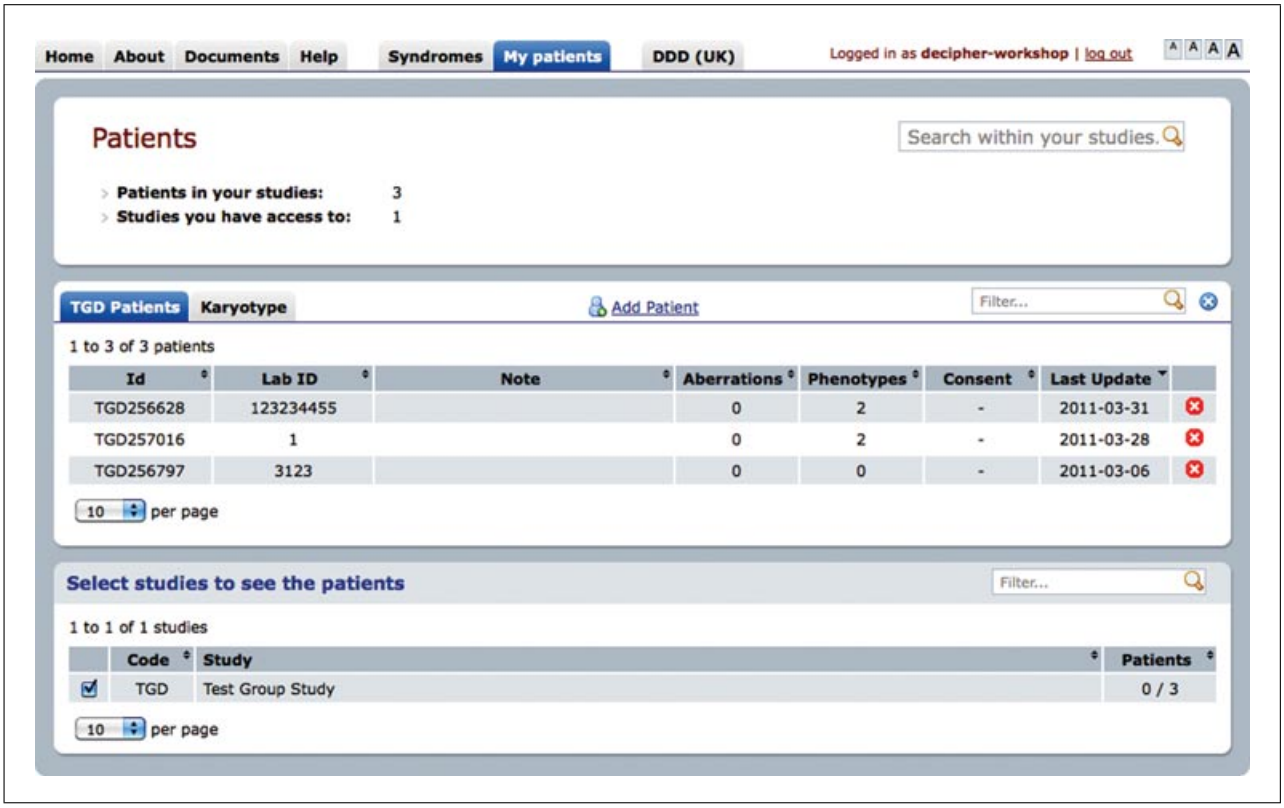


Figure 8.14.13 My Patients tab after logging into DECIPHER shows all studies and patients available to the user.

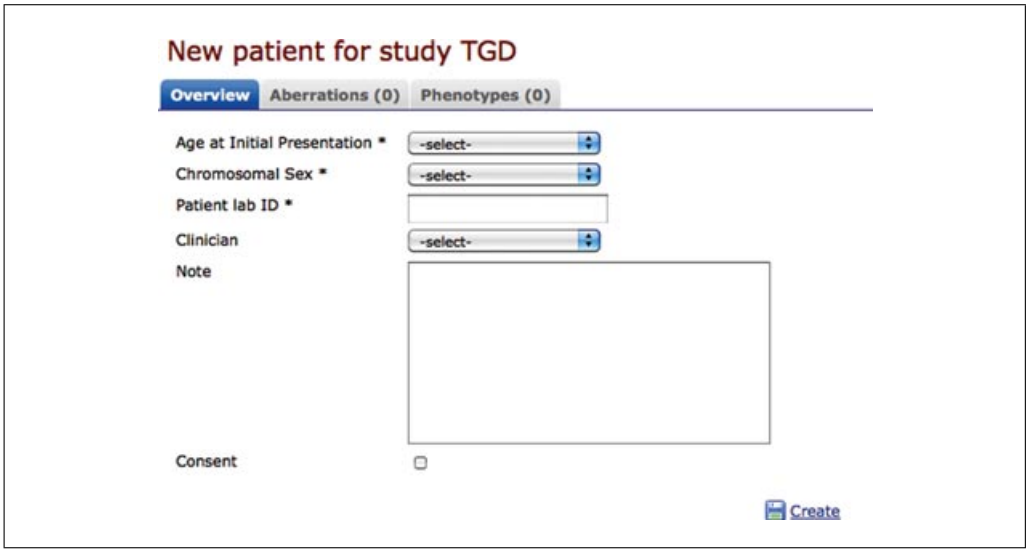


Figure 8.14.14 Adding a patient requires first some overview information about the patient. Asterisks denote required fields.

Figure 8.14.16 shows fields with values filled in for this patient aberration. Note that a mean ratio value for this aberration was not available, so the alternative Copy Number field was filled in. Currently, DECIPHER only accepts coordinate values in GRCh37/hg19 reference assembly. A reference manual on how to remap from NCBI36/hg18 to GRCh37/hg19 is provided via DECIPHER link (http://decipher.sanger.ac.uk/pdfs/assembly.pdf).

9. Fill in the Aberration form. In Classification Type, select De Novo and in Confirm By, select FISH.
10. When complete, hit Save.

Figure 8.14.15 Consent forms are required to be signed for sharing anonymized DECIPHER patient data. Forms are accessible via the Documents tab in the main menu and are available in several languages.

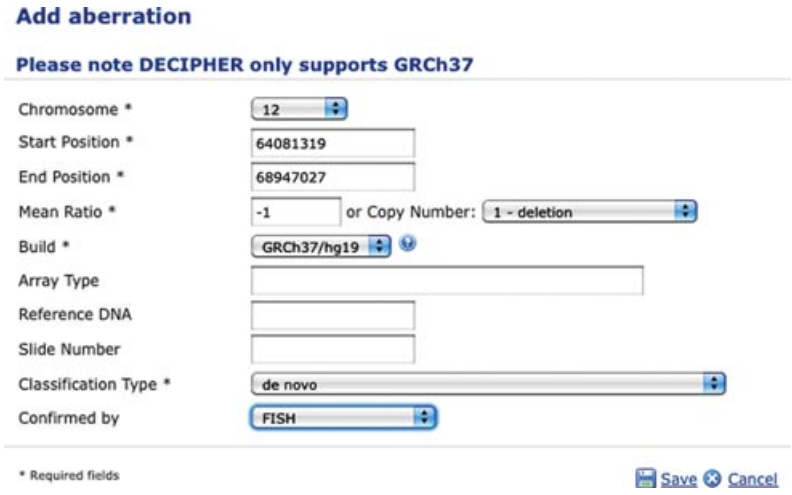


Figure 8.14.16 Pop-up for adding an aberration. Asterisks denote required fields and only aberrations in GRCh37/hg19 are allowed.

11. Investigate the Aberration tabs.

    Figure 8.14.17 shows the overlapping genes, OMIM genes, and OMIM morbid genes for this aberration. Since this aberration is a loss, it is shown in red in the Patient Track region of the graph.

12. Click the e! button to see the patient's details in Ensembl.

13. Click on the Syndrome Overlap tab.

    Note that this patient overlaps with the 12q14 Microdeletion Syndrome.

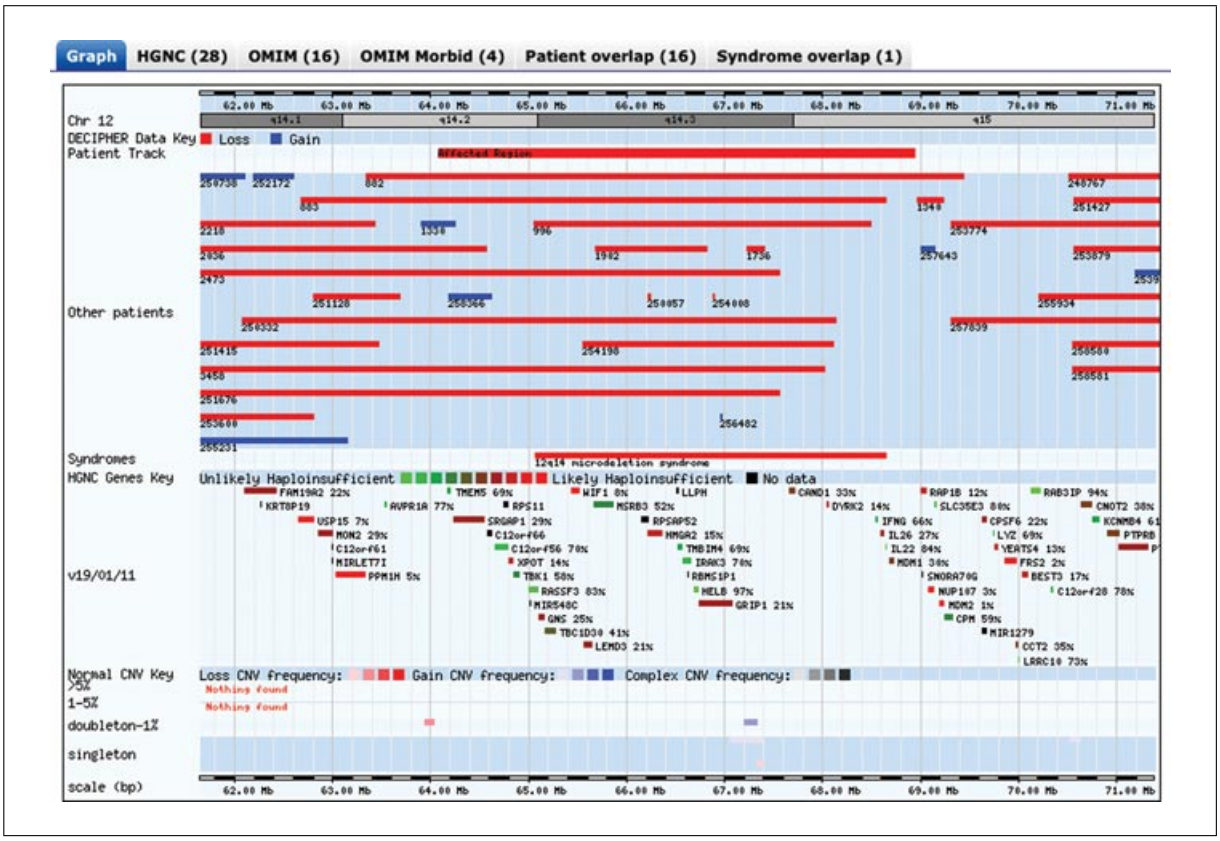


Figure 8.14.17 Aberration tabs for a deletion in chromosome 12 starting at 64081319 and ending at 68947027. For this region, 28 genes are overlapped of which 16 are OMIM. Of these OMIM genes, 4 are denoted as Morbid. In total, 9 consented patients and syndromes overlap this aberration.

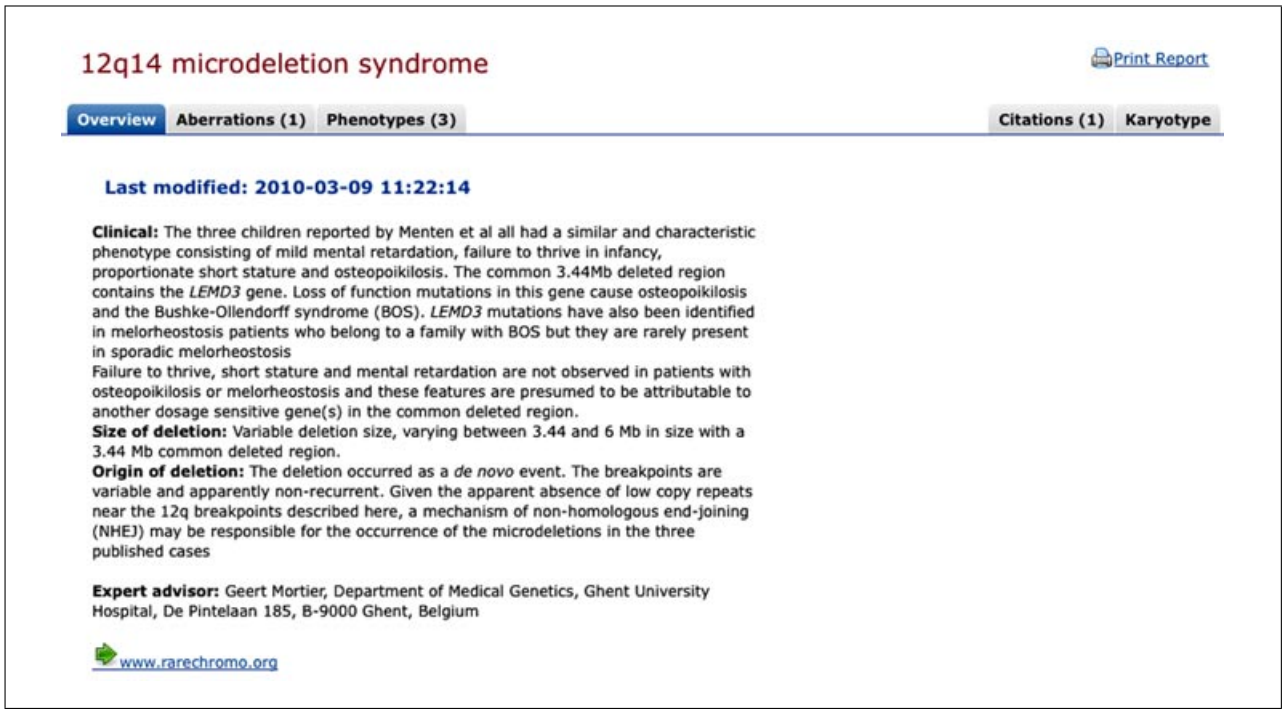


Figure 8.14.18 Overview information for 12q14 Microdeletion Syndrome. A synopsis of the clinical features, size of deletion, and expert advisor is provided, together with a link to Unique, a support group agency for families.

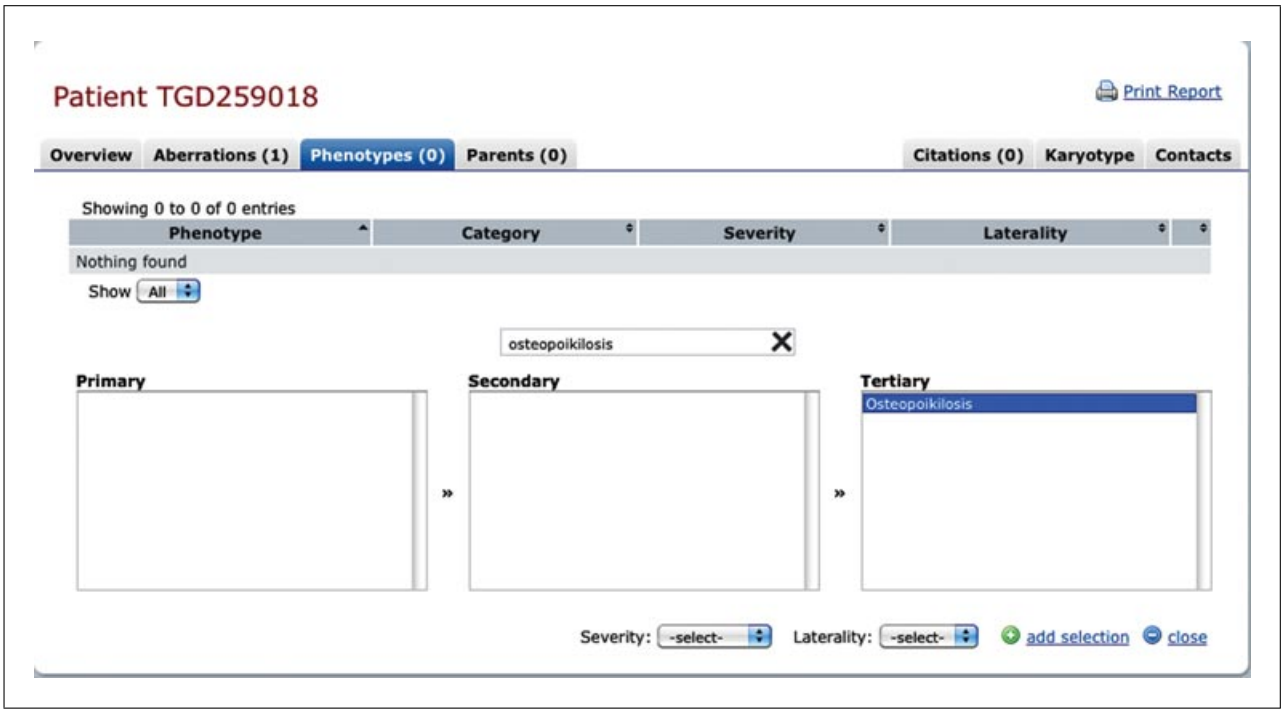


Figure 8.14.19 Interface for adding a LDDB phenotype term.

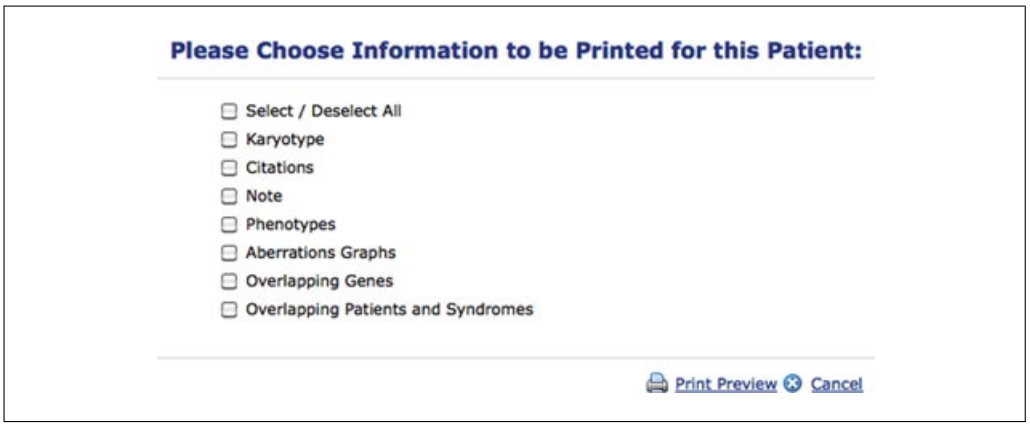


Figure 8.14.20 Pop-up window showing selectable items to be printed.

14. Click on the overlapping syndrome and explore the information this entry provides.
15. Click on the syndrome Overview (Fig. 8.14.18).

    The text included under the Overview tab reports that the LEMD3 gene has been implicated in osteopoikilosis and that short stature and mental retardation are not observed in patients with osteopoikilosis but are presumed to be attributable to another dosage sensitive gene(s) in the common deleted region. Nathalie has the LEMD3 gene affected as well as GNS, IFNG, and IRAK3 all catalogued in OMIM Morbid.

Add phenotypes

16. To add the clinical phenotype, select the Phenotype tab. Add the phenotype by either of the following two ways:

    a. by searching in the top input box *or*

    b. bynavigating the hierarchy.

17. Type “osteopoikilosis” in the search box and hit Search.
18. Select the term from the Tertiary set of terms and click Add Selection (Fig. 8.14.19).

19. If a case report or other publication about this patient has been written, add the relevant citation by adding the corresponding PMID.

Print report

20. Back in the Patient Report, click on Print Report.

21. From the resulting pop-up window (Fig. 8.14.20), click the Select All check box and Print Preview. By default, the overview information is shown.

22. Click on Print Document to print the patient report.

COMMENTARY

Background Information

Implementation of high-resolution whole-genome technologies such as genomic array analysis and exome/whole-genome sequencing is revealing an enormous wealth of, thus far, unseen genomic variation. Understanding the phenotypic consequences of this variation is one of the key challenges in genetic medicine. DECIPHER facilitates the association of rare genomic variants with their corresponding phenotypes by providing a secure collaborative environment for clinicians and researchers to share linked-anonymized patient data. To facilitate analysis, a suite of bioinformatics, search, and visualization tools are provided whose basic functionality is described in this protocol. DECIPHER aims to provide the most up-to-date integrated view of genomic and phenotypic data using the best bioinformatics resources, thereby enabling the diagnosis of patients with rare genetic disorders due to pathogenic copy number variation. For further support, desired functionality, or any queries, contact DECIPHER at *decipher@sanger.ac.uk*.

Acknowledgments

The authors would like to acknowledge Nigel Carter for helpful feedback on this manuscript, Tomas Fitzgerald for compiling the consensus CNV track data, and Ni Huang and Matthew Hurles for providing the haploinsufficiency scores. The authors are grateful to the Wellcome Trust for their financial support (grant no. WT077008). The authors thank the patients and their families for permission to include their information in the DECIPHER database. Finally, the authors would like to thank all of the DECIPHER Consortium Centers, a list of which can be found on the DECIPHER website.

Literature Cited

Baraitser, M., Tomiwa, K., Wilson, J., and Winter, R. 1989. A database of genetically determined neurological conditions for clinicians. *J. Neurol. Neurosurg. Psychiatry* 52:106.

Bruford, E.A., Lush, M.J., Wright, M.W., Sneddon, T.P., Povey, S., and Birney, E. 2008. The HGNC Database in 2008: A resource for the human genome. *Nucleic Acids Res.* 36:D445-D448.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., and Campbell, P. 2009. Origins and functional impact of copy number variation in the human genome. *Nature* 464:704-712.

Engels, H., Wohlleber, E., Zink, A., Hoyer, J., Ludwig, K.U., Brockschmidt, F.F., Wieczorek, D., Moog, U., Hellmann-Mersch, B., and Weber, R.G. 2009. A novel microdeletion syndrome involving 5q14. 3-q15: Clinical and molecular cytogenetic characterization of three patients. *Eur. J. Hum. Genet.* 17:1592-1599.

Firth, H.V. and Wright, C.F. 2011. The deciphering developmental disorders (DDD) study. *Dev. Med. Child Neurol.* 53:702-703.

Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. 2009. DECIPHER: Database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am. J. Hum. Genet.* 84:524-533.

Flicek, P., Amode, M., Barrell, D., Beal, K., Brent, S., Chen, Y., Clapham, P., Coates, G., Fairley, S., and Fitzgerald, S. 2010. Ensembl 2011. *Nucleic Acids Res.* 39:D800-D806.

Huang, N., Lee, I., Marcotte, E.M., Hurles, M.E., and Schierup, M.H.2010. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* 6:e1001154.

Jaillard, S., Drunat, S., Bendavid, C., Aboura, A., Etcheverry, A., Journel, H., Delahaye, A., Pasquier, L., Bonneau, D., and Toutain, A. 2010. Identification of gene copy number variations in patients with mental retardation using array-CGH: Novel syndromes in a large French series. *Eur. J. Med. Genet.* 53:66-75.

Le Meur, N., Holder-Espinasse, M., Jaillard, S., Goldenberg, A., Joriot, S., Amati-Bonneau, P., Guichet, A., Barth, M., Charollais, A., and Journel, H. 2010. MEF2C haploinsufficiency caused by either microdeletion of the 5q14. 3 region or mutation is responsible for severe mental retardation with stereotypic movements, epilepsy and/or cerebral malformations. *J. Med. Genet.* 47:22.

Mills, R.E., Walter, K., Stewart, C., Handsaker, R.E., Chen, K., Alkan, C., Abyzov, A., Yoon, S.C., Ye, K., and Cheetham, R.K. 2011. Mapping copy number variation by population-scale genome sequencing. *Nature* 470:59-65.

Robinson, P.N., Köhler, S., Bauer, S., Seelow, D., Horn, D., and Mundlos, S. 2008. The human phenotype ontology: A tool for annotating and analyzing human hereditary disease. *Am. J. Hum. Genet.* 83:610-615.

Sanborn, J., Benz, S., Craft, B., Szeto, C., Kober, K., Meyer, L., Vaske, C., Goldman, M., Smith, K., and Kuhn, R. 2010. The UCSC cancer genomics browser: Update 2011. *Nucleic Acids Research.* 39:D951-D959.

Stankiewicz, P. and Beaudet, A. 2007. Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic mental retardation. *Curr. Opin. Genet. Dev.* 17:182-192.

Winter, R.M. and Baraitser, M. 1987. The London Dysmorphology Database. *J. Med. Genet.* 24:509.

Zweier, M., Gregor, A., Zweier, C., Engels, H., Sticht, H., Wohlleber, E., Bijlsma, E.K., Holder, S.E., Zenker, M., and Rossier, E. 2010. Mutations in MEF2C from the 5q14. 3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum. Mutat.* 31:722-733.