

CTGA Database Information Submission Form

Help Topics

Everyday the Centre for Arab Genomic Studies receives a variety of publications from scientists in the Arab World or abroad to be considered for the CTGA Database on genetic disorders in Arab populations. Curators of the CTGA Database are working continuously to update each and every record of the database following rigorous steps that include paper review, editing, and curation. At present, CTGA Database curators create about 30 entries and update an equivalent number each month with an average of 5-10 publications reviewed per day. Geneticists working on genetic disorders in Arab individuals are hailed to contribute to the growth of the CTGA Database. For this reason, curators of the CTGA Database prepared the CTGA Database Information Submission Form to speed up the process of data submission, review, and publication in the CTGA Database. When submitting the form, it is always preferable to send also supportive documents, such as the publication(s) cited (in hard copy or electronic formats).

Proper filling of the CTGA Database Information Submission Form is a key factor in speeding up the process of data curation in the CTGA Database. Accordingly, this help document provides tips and extensive examples on how to use the CTGA Database Information Submission Form, which can be used for reporting clinical or molecular data related to a single disease or gene locus entity.

I. Name of suggested Disorder or Gene

This box should be filled with the primary title of the disorder. Whenever available, alternative names of the disorders or commonly used abbreviated names may also be reported. If more than one name is indicated, please separate with commas. If the form is used to submit molecular (mutation) data, then full-name of the gene locus is to be reported along with alternative names and symbols. Maximum capacity of this box is 100 characters.

Example 1 (genetic disorder):

[Sickle Cell Disease, SCD, Sickle Cell Anemia, SCA, Hemoglobin SS Disease](#)

Example 2 (genetic disorder):

[Absence of Abdominal Muscles with Urinary Tract Abnormality and Cryptorchidism, Prune Belly Syndrome, Eagle-Barrett syndrome, Triad Syndrome, Abdominal Muscular Deficiency Syndrome](#)

Example 3 (gene locus):

[1-@Acylglycerol-3-Phosphate O-Acyltransferase 2, AGPAT2, Lysophosphatidic Acid Acyltransferase-Beta, LPAAT-Beta](#)

Example 4 (gene locus):

[5,10-@Methylenetetrahydrofolate Reductase, MTHFR](#)

II. Classification of the Disease

The CTGA Database follows the WHO International Classification of Disease (Ver. 10) as a standard to classify disease records. If the submission form is used to report data on a genetic disease, the appropriate choice should be selected. Available options include:

1. Congenital malformations, deformations, and chromosomal abnormalities
2. Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
3. Diseases of the circulatory system
4. Diseases of the digestive system
5. Diseases of the ear and mastoid process
6. Diseases of the eye and adnexa
7. Diseases of the genitourinary system
8. Diseases of the musculoskeletal system and connective tissue
9. Diseases of the nervous system
10. Diseases of the respiratory system
11. Diseases of the skin and subcutaneous tissue
12. Endocrine, nutritional and metabolic diseases
13. Certain infectious and parasitic diseases
14. Mental and behavioural disorders
15. Neoplasms
16. Pregnancy, childbirth and the puerperium
17. Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified

Example 1:

[Absence Defect of Limbs, Scalp, and Skull](#); select: *Congenital malformations, deformations and chromosomal abnormalities*

Example 2:

[Adenomatous Polyposis of the Colon](#); select: *Diseases of the digestive system*

If the CTGA Database Information Submission Form is used to report data for a gene locus, then leave this section blank as classification is not required.

III. Mode of Inheritance

This section is to indicate the mode of inheritance of the disorder or gene locus reported in the form. Two selectors are available to give combinations of many choices. Available options include:

Selector 1

1. Autosomal
2. X-linked
3. Y-linked
4. Mitochondrial
5. Multifactorial

Selector 2

1. Recessive
2. Dominant

A box is also provided to describe any additional notes, if necessary. Maximum capacity of this box is 100 characters.

Example 1 (genetic disorder):

[Adrenoleukodystrophy](#); selector 1: *X-linked*

Example 2 (genetic disorder):

[Al-Gazali Syndrome](#); selector 1: *Autosomal*; Selector 2: *recessive*

Example 3 (genetic disorder):

[Congenital Failure of Autonomic Control](#); selector 1: *Autosomal*; Selector 2: *recessive*; Additional Notes: *dominant with reduced penetrance (or paternal gonadal mosaicism)*

Example 4 (genetic disorder):

[Hemoglobin--Beta Locus](#); selector 1: *Autosomal*; Additional Notes: *Autosomal dominant for some such as methemoglobinemia, polycythemia, and Heinz body hemolytic anemia Autosomal recessive for others such as sickle cell disease and thalassemia major.*

IV. OMIM Number

This is an optional field. However, providing the international number of the disease or gene locus as adapted by the [Online Mendelian Inheritance in Man \(OMIM\) database](#) will make it easier for the curators of the CTGA database to properly integrate suggested data into the corresponding record.

Two boxes are provided in this section:

1. Disease OMIM Number: If the submission form is used to propose data on a genetic disorder, please provide the corresponding OMIM number. There is no need to include the symbols adapted by the OMIM Database, only the six digit number is necessary. The maximum capacity of the box is 6 characters only.
2. Gene(s) OMIM Number: If the submission form is used to suggest data on a genetic disorder, and if the disorder is mapped to a locus (loci), please provide the OMIM numbers of the corresponding locus (loci). If the disease reported is not yet linked to a specific gene, please leave this box blank. The maximum capacity of this box is 46 characters; to allow multiple entry reporting. In case multiple entries are reported, please separate with commas.

Example 1 (genetic disorder):

[Hereditary Nonpolyposis Colorectal Cancer Type 1](#); Disease OMIM Number: 120435; Gene(s) OMIM Number: *leave blank*

Example 2 (gene locus):

[Collagen, Type XVII, Alpha-1](#); Disease OMIM Number: *leave blank*; Gene(s) OMIM Number: 113811

Example 3 (genetic disorder with a defined gene locus):

[Deafness, Autosomal Recessive 9](#); Disease OMIM Number: 601071; Gene(s) OMIM Number: 603681

V. Description

A general description of the genetic disorder reported in the submission form. Basic requirements in this section may include the world prevalence of the disease, its clinical symptoms, diagnosis, and treatment modalities. No references are required in this section. If the submission form is used to report gene mutations, you may then skip this section and proceed to section VI. Molecular Genetics. Maximum number of characters allowed in this section is 1000 characters.

Example 1:

[Fibrochondrogenesis](#); Description: *Fibrochondrogenesis is a rare disorder with one case recorded among 1,158,067 live-births registered by the Spanish Collaborative Study of Congenital Malformations (ECEMC). Until March 2003, only a limited number of cases with fibrochondrogenesis were recorded worldwide. Fibrochondrogenesis is a neonatally lethal rhizomelic chondrodysplasia distinguished from other forms of lethal dwarfism by broad long-bone metaphyses, pear-shaped vertebral bodies, and*

characteristic microscopic changes of cartilage: unique interwoven fibrous septa and fibroblastic dysplasia of chondrocytes. Fibrochondrogenesis is distinguished radiologically by the widening of the metaphyses of the long bones, and, on lateral X-ray of the spine, by a median fissure of the body of the vertebra without any loss of vertebral height. Omphalocele is the only internal anomaly known to be associated with this disease.

Example 2:

[Frontonasal Dysplasia](#); Description: *Frontonasal dysplasia is a rare midline anomaly characterized by malformations of the central portion of the face, especially of the forehead, the nose, and the philtrum. True ocular hypertelorism; broadening of the nasal root; a median facial cleft affecting the nose, upper lip, and palate; uni- or bilateral clefting of the alae nasi; lack of formation of the nasal tip; anterior cranium bifidum occultum; and a V-shaped prolongation of the hair onto the forehead are the main features. Frontonasal dysplasia may be associated with congenital heart abnormalities, in particular, tetralogy of Fallot, vertebral anomalies, agenesis of the corpus callosum, Dandy-Walker malformation, a short neck, short limbs, polydactyly of the hands and feet, and cryptorchidism.*

Example 3:

[Fryns Syndrome](#); Description: *Fryns syndrome is a rare autosomal recessive disorder, multiple congenital anomaly syndrome with an incidence of 0.7-1 in 10,000 births. The syndrome is characterized by congenital diaphragmatic hernia, unusual facies and distal limb hypoplasia. The spectrum of distal limb hypoplasia includes short and broad hands, short digits, short or absent terminal phalanges, hypoplastic or absent nails, and clinodactyly. Neurologic and cardiac malformations have been reported in up to 72% and 88% of Fryns syndrome cases, respectively. Fryns syndrome is usually associated with stillbirth and death soon after birth. Patients who survive the neonatal period represent 14% of reported cases. Characteristics of survivors include less frequent congenital diaphragmatic hernia and milder lung hypoplasia, absence of complex cardiac malformations, frequent early myoclonus, and most often, severe neurologic impairment. A significant inter and intra-familial phenotypic variability as well as discordant phenotype in monozygotic twins has been reported. Detection of fetal hydrops, cystic hygroma, and multiple pterygia have allowed prenatal ultrasonographic diagnosis as early as in the 11th week of gestation.*

VI. Molecular Genetics

A general description of the gene locus reported in the submission form. Basic requirements in this section may include a general description of the protein product, its properties, suggested function, sites of expression, size of the gene, number of exons and introns, and common mutations. No references are required in this section. Maximum number of characters allowed in this section is 1000 characters.

Example 1:

[G Protein-Coupled Receptor 56](#); Molecular Genetics: *The mammalian cerebral cortex is characterized by complex patterns of anatomical and functional areas that differ markedly between species, but the molecular basis for this functional subdivision is largely unknown. Mutations in GPR56, which encodes an*

orphan G protein-coupled receptor (GPCR) with a large extracellular domain, cause a human brain cortical malformation called bilateral frontoparietal polymicrogyria (BFPP). BFPP is characterized by disorganized cortical lamination that is most severe in frontal cortex. GPR56 consists of 14 exons covering 15 kb of genomic sequence and has a 3-kb open reading frame. The pattern of the mouse *Gpr56* expression and the anatomy of BFPP imply that *Gpr56* most likely regulates cortical patterning. The most severely affected cortical regions in BFPP are strikingly thin, and many forms of PMG show cortical alterations with reduction of the normal six cortical layers to four, suggesting possible roles in cell fate control that would be consistent with the expression of GPR56 in progenitor cells.

Example 2:

[Galactosylceramidase](#); Molecular Genetics: The galactosylceramidase gene (*GALC*) encodes a lysosomal enzyme which catabolises degradation of several galactolipids such as galactosylceramide, galactosylsphingosine and galactosyldi-glyceride. The galactosylceramidase gene (*GALC*) is about 60 kb in length, consists of 17 exons, contains ten GC-box-like sequences within the promoter region and one potential YY1 element, and one potential SP1 binding site. Nearly 70 mutations, including polymorphisms in every one of the 17 exons have been identified in individuals with Krabbe disease. The 30-kb deletion, which always occurs with the C>T 502 (R>C 168) polymorphism, makes up approximately 45% of mutant alleles in the population with European ancestry. Missense mutations causing the infantile form of Krabbe disease are found in both subunits, although more seem to be found in the coding region for the 30-kd subunit. One mutation, G>A at position 809, resulting in G>D substitution at amino acid position 270, always results in the later-onset form of Krabbe disease. A number of small deletions and insertions have been identified, and these result in a frame shift and premature termination. Even missense mutations very near the 3' end of the coding region that result in amino acid changes near the carboxyl end of the 30-kd subunit result in clinical disease.

Example 3:

[Glyceronephosphate O-Acyltransferase](#); Molecular Genetics: Dihydroxyacetonephosphate acyltransferase (DHAPAT, or DAPAT; EC 2.3.1.42), a key enzyme in the biosynthesis of ether phospholipids, is localized exclusively within peroxisomes. DHAPAT and alkyl-DHAP synthase are responsible for the first two steps of plasmalogen biosynthesis. As the etherphospholipid plasmalogens are major constituents of myelin phospholipids, defective development of myelin is expected in individuals with deficiencies of these enzymes. Patients with rhizomelic chondrodysplasia punctata (RCDP) can be subdivided into three subgroups based on biochemical analyses and complementation studies. The largest subgroup contains patients with mutations in the *PEX7* gene encoding the PTS2 receptor. This results in multiple peroxisomal abnormalities which includes a deficiency of acyl-CoA:dihydroxyacetonephosphate acyltransferase (DHAPAT), alkyl-dihydroxyacetonephosphate synthase (alkyl-DHAP synthase), peroxisomal 3-ketoacyl-CoA thiolase and phytanoyl-CoA hydroxylase, although there are differences in the extent of the deficiencies observed. Patients in the two other subgroups have been reported to be either deficient in the activity of DHAPAT (RCDP type 2) or alkyl-DHAP synthase (RCDP type 3) while no other abnormalities could be observed. By fluorescence in situ hybridization, DHAPAT was mapped to human chromosome 1q42.

VII. Epidemiology in the Arab World

This is the most important part of the CTGA Database Information Submission Form. It is used to report the occurrence of a genetic disease or a genetic mutation in an Arab individual. The most commonly accepted source of information would be a scientific paper cited in a peer-reviewed local or international journal. Personal communications may also be accepted on condition to receive supportive documentation. This section is not intended to report abstracts of published papers since they are readily available in bibliographic databases such as a <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed> and SCI-Expanded.

In case the CTGA form is used to report the occurrence of a genetic disorder in Arab individuals, an extended summary of the published paper is required. Key features should include: a description of the patient(s) reported (age, sex, etc.), the family history, clinical description, key symptoms and diagnostics that played major role in defining the disease, and treatment modalities or medical interventions utilized.

In case the CTGA form is used to report the occurrence of gene mutations in Arab individuals, the extended summary should include: description of the patient(s) reported (age, sex, disease features), methods utilized in the molecular characterization, the gene mutation, and possible effects on the expression of the corresponding gene or locus.

This section contains three boxes to report data from a maximum of three separate publications. Please use each box to report data from a single clinical or molecular paper discussing one disease or a gene locus and do not exceed 1000 characters for each paper summary. To report results of a single study, please ignore the boxes labeled "Study 2" and "Study 3". To report more than three studies related to a disease or gene, please use additional forms.

Example 1 (genetic disorder, single study):

[Gray Platelet Syndrome](#); Epidemiology in the Arab World; **Study 1:** *Alkhairy (1995) reported the gray platelet syndrome in a Palestinian female teenager from Saudi Arabia with menorrhagia and thrombocytopenia from menarche at 11 years. This was the first report of gray platelet syndrome in an Arab patient. The proband was an 18 year old who presented for the first time with a history of dizziness and general fatigability of two weeks duration. She had a history of easy bruising and epistaxis since childhood. Three of her 5 sibs, all male, had a bleeding tendency. All 3 bled for 24 hours after hospital-based circumcision and required multiple ligations to achieve hemostasis. All 4 affected sibs displayed typical morphology of gray platelet syndrome. The parents were first cousins of Palestinian origin with normal hematologic parameters. The fact that four siblings of both sexes are simultaneously involved indicates that their disorder is inherited in autosomal recessive manner; though gonadal mosaicism for a dominant gene defect cannot be excluded.;* **Study 2:** leave blank; **Study 3:** leave blank.

Example 2 (genetic disorder, multiple studies):

[Jejunal Atresia](#); Epidemiology in the Arab World; **Study 1:** *Farag and Teebi (1989) described two Arab brothers with 'apple peel' jejunal atresia whose Palestinian parents were consanguineous. The two sibs*

died after surgical intervention.; **Study 2:** Farag et al. (1993) reported two other affected sibs in the same family: later-born female and male infants. The diagnosis was confirmed in the female patient by plain abdominal X-ray. Laparotomy showed jejunal atresia with "apple peel" variant. The baby died 40 days after the operation. The male infant was born after delivery of a phenotypically normal girl and one early spontaneous abortion. Abdominal X-ray confirmed the diagnosis. Laparotomy showed upper jejunal atresia. Resection anastomosis was done to remove a gangrene in the small gut. The baby died 67 days after the operation.; **Study 3:** leave blank.

Example 3 (gene locus, single study):

[Growth/Differentiation Factor 5](#); Epidemiology in the Arab World; **Study 1:** In 2003, Al-Yahyaee et al. studied the clinical, radiographic, and genetic characteristics of an Omani family with four affected children. Genetic studies showed transition A1137G and deletion delG1144 mutations in the gene encoding the cartilage-derived morphogenetic protein-1 (CDMP-1) in this family. The A1137G was a silent mutation coding for lysine, whereas the delG1144 predicted a frameshift mutation resulting in a presumable loss of the CDMP-1 biologically active carboxy-terminal domain. The affected siblings were homozygous for the delG1144 mutation while parents were heterozygous.; **Study 2:** leave blank; **Study 3:** leave blank.

Example 4 (gene locus, multiple studies):

[H Factor 1](#); Epidemiology in the Arab World; **Study 1:** Ohali et al. (1998) described the clinical course, complement components, and pathological findings of 10 infants with autosomal recessive hemolytic-uremic syndrome (HUS). All patients were the offspring of four Bedouin couples from an extended and highly inbred Bedouin family. Serum factor H levels were greatly decreased or absent in 4 patients tested and moderately decreased in 15 of 23 healthy unaffected siblings and patients.; **Study 2:** In year 1999, Ying et al. conducted further analyses on the family of Ohali et al. (1998). Ying et al. (1999) performed linkage analysis to investigate a possible linkage between the disorder and the markers near the complement factor H (CFH) gene. Mutation analysis of the entire 20-exon coding region of the HF1 gene revealed homozygosity for a single C-T missense mutation, resulting in a ser1191-to-leu amino acid change. Functional analyses demonstrate that the mutant CFH is properly expressed and synthesized but that it is not transported normally from the cell. The study of Ying et al. (1999) was the first to report that a recessive, atypical, early-onset and relapsing HUS is associated with the CFH protein and that a CFH mutation affects intracellular trafficking and secretion.; **Study 3:** leave blank.

Example 5 (gene locus, multiple studies):

[Hemoglobin - Alpha Locus 1](#); Epidemiology in the Arab World; **Study 1:** El-Kalla and Baysal (1998) studied alpha-thalassemia in the United Arab Emirates and examined the alpha globin genes of 418 cord blood samples from newborn UAE nationals using microcolumn chromatography, isoelectric focusing, alkali denaturation, spectrophotometry, PCR, hybridization and DNA sequencing. The frequency of alpha-thalassemia among these neonates was very high (49%). The most common mutation was the 3.7 Kb deletion (68.6%, frequency 0.2847). One newborn was found to be compound heretozygous for the 3.7 Kb and the 4.2 Kb deletions. Four different non-deletional alpha globin mutations (alpha-T) were also identified; which were responsible for about 6% of the total mutations. These were: alpha-PA-1, alpha-PA-2, HbCS, and alpha-5nt del. Hb Barts was found in all cases with alpha-T mutations.; **Study 2:** Baysal

(1998) also studied the genotype-phenotype correlation of 17 UAE nationals with HbH disease or Hb like conditions, and characterized five different alpha-thalassemia determinants (--MED-I, 3.7 Kb deletion, alpha-PA-1, alpha-CS, alpha-5nt del). Two brothers, both with alpha-CS alpha/ alpha-CS alpha presented with totally different conditions. One was symptomatic from early infancy, while the other was completely asymptomatic.; **Study 3:** In a study on the hemoglobinopathies in the United Arab Emirates, Baysal (2001) examined the alpha globin genes of 418 cord blood samples from newborn UAE nationals using PCR, hybridization and DNA sequencing. The frequency of alpha-thalassemia among these neonates was very high (49%). Additionally, Baysal (2001) identified four non-deletional mutations in 3% of the chromosomes (alpha-T). Baysal (2001) studied 28 alpha thalassemic UAE nationals, with HbH. The poly a-1 mutation [alpha-PA-1 (AATAAA-AATTAAG)] was the most common mutation (47.4%). In total, nine different alpha-thalassemia genotypes were identified.

VIII. Country of Origin/Residence of Individual(s) Analyzed

This is also a very important component of the CTGA Database Submission Form. In many papers, we have realized the complete absence of information on the ethnicity or place of origin of patients described with genetic disorders. Such information is vital to achieve the purpose of the CTGA Database to give a clear scope on the distribution of genetic disorders in Arab populations.

Boxes in this section use ready-made options that can be selected to fill information regarding the country of origin and country of residence of patients described in section VII. Use the boxes labeled Study 1, Study 2, and Study 3 to match with the reports described in section VII under Study 1, Study 2, and Study 3, respectively.

If in a single study multiple patients are described, please use the box labeled "Additional Notes" (maximum capacity: 100 characters) to list the countries of origin or residence of individuals studied.

Example 1:

[Joubert Syndrome 1](#); Epidemiology in the Arab World; **Study 1:** Sztriha et al. (1999) analyzed a 7-year-old male from the United Arab Emirates with Joubert syndrome born to consanguineous parents of Palestinian origin...; **Study 1 - Selector Origin:** Palestine; **Study 1 - Selector Residence:** United Arab Emirates.

Example 2:

[Krabbe Disease](#); Epidemiology in the Arab World; **Study 1:** Al-Essa et al. (2000) studied a 2-year, 6-month-old Saudi male with infantile Krabbe's disease with fluorine-18-labeled-2-fluoro-2-deoxyglucose positron emission tomography (FDG PET) scan...; **Study 1 - Selector Origin:** Saudi Arabia; **Study 1 - Selector Residence:** Saudi Arabia.

IX. Reference

In this section, please indicate the detailed coordinates of the references used in Section VII. The CTGA database uses the style followed in [PubMed](#) for citation reporting. Three boxes are provided to cite multiple references used in Section VII under Study 1, Study 2, and Study 3. Each box is set with a maximum capacity of 500 characters.

Example:

[L-2-Hydroxyglutaricacidemia](#); Epidemiology in the Arab World; **Study 1:** Duran et al. (1980) reported a 5-year-old boy of Moroccan (Berber) origin who had nonspecific mental and motor delay and growth deficiency...; **Study 2:** Larnaout et al. (1994) described three brothers from Tunisia, born to a consanguineous family, with a progressive neurological disorder associated with L-2-hydroxyglutaric aciduria...; **Study 3:** blank; References; **Study 1:** Duran M, Kamerling JP, Bakker HD, van Gennip AH, Wadman SK. L-2-Hydroxyglutaric aciduria: an inborn error of metabolism? *J Inherit Metab Dis.* 1980; 3(4):109-12.; **Study 2:** Larnaout A, Hentati F, Belal S, Ben Hamida C, Kaabachi N, Ben Hamida M. Clinical and pathological study of three Tunisian siblings with L-2-hydroxyglutaric aciduria. *Acta Neuropathol (Berl).* 1994; 88(4):367-70. **Study 3:** leave blank.

X. Date of Contribution

This box is to indicate the date of submission of the form following the style: day.month.4-digit-year. This date will appear besides the contributors' name in the corresponding section of a CTGA record.

Example:

29.5.2006