

PEDIGENE®

A Proven Comprehensive Approach to Information Management

Enhancing the Genomic Research Laboratory



Duke University Medical Center, Durham, North Carolina USA

PEDIGENE® Researchers: Pioneering Genetic Research in the 21st Century

In 1996 Duke University Medical Center established the Center for Human Genetics (CHG), a medical-center-wide program designed to attract physicians and scientists interested in pursuing genetics research. The CHG is comprised of genetic epidemiologists, biostatisticians, genome biologists, genetic counselors, physicians, physician's assistants and molecular biologists. Over the past 10 years, researchers at the CHG have identified genes for a wide range of common diseases, from the initial identification of the APOE gene for Alzheimer's Disease to identification of genes for heart disease, age-related macular degeneration, Parkinson Disease and muscular dystrophy. Ongoing collaborations span virtually all medical specialties including psychiatry, pulmonary medicine, surgery, cardiology, radiology, obstetrics & gynecology, ophthalmology, nephrology, and pediatrics among others.

The study of these diseases, requiring detailed coordination between the clinicians, epidemiologists, molecular biologists and statisticians, resulted in a comprehensive database system as the core of the research efforts. The PEDIGENE® information system, initially developed by Drs. Margaret Pericak-Vance, Vance, Haines, Speer and other CHG faculty together with lead programmer Carol Haynes and copyrighted by Duke University in 1997, manages the massive amounts of data produced during the search for a disease gene.

PEDIGENE® has continued to grow and develop as the scope and scale of genetic and genomic studies has increased, currently providing database services for over 180 studies. The PEDIGENE® team is always at work to keep PEDIGENE® a state-of-the-art data information system for human genetics research, including investigations of interactions between genes and between genes and the environment. The PEDIGENE® database designers work to ensure the appropriate evolution of this innovative, flexible system by implementing leading-edge algorithms and multimedia databasing. With a unique symbiosis of outstanding researchers and database designers, the PEDIGENE® team is prepared to pioneer genetic research in the 21st Century.

PEDIGENE® An Innovative Database Solution

PEDIGENE® is a registered trademark of the Duke Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, United States of America



PEDIGENE®

The PEDIGENE® information management system is a large-scale database system designed for the genetic laboratory. PEDIGENE® manages data from all phases of genetic, clinical and molecular analysis. The user-friendly PEDIGENE® system manages data through five modules:

- Clinical Module
- Patient Information Module
- DNA Bank and Tissue Repository Module
- Molecular Genetics Module
- Statistical Analysis Module

PEDIGENE® is a proven research tool that addresses the needs of a highly productive genomics laboratory. PEDIGENE® emphasizes data integrity as well as management and automation of the repetitive aspects of genetic studies. The system's robust security rigorously guards access to confidential genetic research data. A team of dedicated researchers and database designers constantly expands the system's state-of-the-art capabilities.

Clinical Module

PEDIGENE® archives complex clinical information

Research participant collection goals include obtaining DNA and tissue samples, family histories, and clinical examination results. The Clinical Module provides the techniques and tools to collect and record all forms of clinical research data. Sample database entries include information such as pertinent medical history and epidemiologic risk factor data. Database storage of this information allows assessment of parameters critical to underlying disease risk. PEDIGENE® specialists develop disease-specific forms to maintain accurate clinical participant profiles. Clinical information storage is customized for each disease. The clinical information is stored in different modules such as medical history, disease-specific test results, clinical exams and autopsies. Researchers can quickly update modules and independently add additional disease-specific modules, as needed.

PEDIGENE® stores multimedia resources

While PEDIGENE® routinely stores numerical and textual data, it also captures information sources such as x-ray and photographic images, audio recordings, and video recordings. Databasing these data allows on-line access to critical clinical materials needed to confirm diagnoses in study families.

PEDIGENE® maximizes participant utilization

Research participants can be followed as members of multiple groups in PEDIGENE®. Participants can be tracked as affected subjects in one group and simultaneously tracked as unaffected subjects in another disorder group. Through this PEDIGENE® multigroup option, ascertained participant information can be utilized for maximum efficiency and cost-effectiveness.

PEDIGENE® provides ready access to clinical data

In PEDIGENE®, clinical data are maintained by use of a relational database management system. This technology enables researchers to access data through a variety of industry-standard software packages. Researchers also have access to customized data entry programs, analysis reports, and query tools that have been developed by the PEDIGENE® database design team. All data undergo rigorous integrity checks throughout each processing stage.

PEDIGENE® rigorously maintains data security

PEDIGENE® clinical information is available only to the principal investigator, designated research staff, and system administrators. Data access by non-critical personnel or investigators is strictly prohibited. PEDIGENE® has built-in multi-layered security levels that permit appropriate access to each member of the research team.



Stores multimedia research data

PEDIGENE® databases the complex medical diagnostic images that form an integral source of clinical research data. These digitized files are readily accessible for clinical study and comparison through PEDIGENE®.

Captures disease-specific clinical information

PEDIGENE® specialists develop disease-specific forms such as the AMD Health and Activities Survey to create accurate and detailed clinical participant profiles.



Jane Doe
Alzheimer disease
Affected



Jane Doe
Heart disease
Unaffected

Maximizes research participation data

With PEDIGENE®'s multigroup capability, affected participants from one study can be simultaneously tracked as unaffected in other studies.

Patient Information Module

PEDIGENE[®] efficiently stores complex patient information

The Patient Information module rapidly and efficiently stores all pertinent family and individual information. Database entries include information on:

- Individuals
- Family relationships
- Diagnostic group relationships
- Diagnostic information
- Sample collection
- Phenotypic variables
- Genotypes
- Demographic information
- Family history

PEDIGENE[®] provides the tools for storing and retrieving complex family pedigrees, and groups of individuals with common characteristics such as disease control groups. Family ascertainment is tracked through Progeny (Progeny Software, LLC) and key variables are passed to PEDIGENE[®] through an extensive integration. Progeny's customizable screens allow for the collection of disease-specific variables during the ascertainment process as well. The comprehensive integration between Progeny and PEDIGENE[®] provides a powerful environment that embeds a state-of-the-art pedigree tool into a world-class genetic information management system.

PEDIGENE[®] tools speed family ascertainment and collection

PEDIGENE's use of Progeny makes family ascertainment easy and accurate. After locating a potential research participant, the research team must ascertain the individual and additional family members when appropriate. Progeny provides techniques and tools to speed ascertainment and collection of research participants and their families. For example, Progeny provides a quick and efficient graphical pedigree drawing tool to step the interviewer through the creation of a pedigree, where additional family members can be added with a single mouse click. Additionally, Progeny provides a wide array of user-definable pedigree presentation options, to fashion data for different audiences. Use of these techniques and tools speeds family ascertainment and provides excellent and customizable displays of pedigrees.

PEDIGENE[®] provides flexible report capability

PEDIGENE[®] has powerful result display capabilities. The user can select a desired sample information screen and view detailed information about a particular individual or group of individuals. PEDIGENE[®] reports provide every aspect of an individual's vital statistics: diagnosis, subset of clinical characteristics, ethnicity, and/or affection status. PEDIGENE[®]'s data report flexibility provides better data management and permits data analysis using differing analysis parameters.

Stores complex family information.
 PEDIGENE® routinely documents complicated family relationships and histories. Each family is assigned a unique group name and number when it is added to the PEDIGENE® database. Individuals are then assigned unique numbers within that family group.

Accelerates ascertainment.
 Experience-tested tools and techniques speed family ascertainment and collection of research data.



Patient Summary Report for PD-TST (48 inds)

File Edit

Save Copy Print Sort Filter Retrieve Close

Average ADD: 54 Average Age of Onset

	Study	Center	Family	Individual	Gender	Race	Date_of_Birth	Sample_Date	Exam_Date	ADE	ADD	Clinical_Status	Pedigree_Status
1	PD	TST	1	1	M	Caucasian	27 Aug 1916	15 Oct 1996		80	58	Affected	Not_Entered
2	PD	TST	1	100	F	Caucasian	04 Dec 1912				0	Unclear	Not_Entered
3	PD	TST	1	101	M							Affected	Not_Entered
4	PD	TST	1	102	M	Caucasian	15 Jul 1915	15 Oct 1996		81	0	Unclear	Not_Entered
5	PD	TST	1	103	F	Caucasian	27 Oct 1918	15 Oct 1996		77	68	Affected	Not_Entered

Produces individual, family and group reports.
 Individual, family and group vital statistics are readily accessible as PEDIGENE® reports.

DNA Bank and Tissue Repository Module

PEDIGENE® provides reliable and efficient DNA Bank management

The DNA Bank and Tissue Repository Module provides accurate and real-time knowledge of sample amounts, lab supplies and service billing. The AutoPure LS® DNA Extractor is used to provide DNA from large numbers of blood, tissue and saliva samples. Sample movement is tracked through the DNA Bank with the Nautilus™ Laboratory Information Management System (the Nautilus™ LIMS is produced by Thermo Fisher Scientific, Inc.) The Nautilus™ LIMS offers a secure, user-friendly means of signing-in and tracking samples and collecting, moving, and storing sample and laboratory data. The Nautilus™ LIMS routinely manages DNA Bank information such as:

- Requisition processing
- Sample type (blood, buccal swabs, cell lines, external DNA, Oragene saliva, Paxgene RNA , Paxgene DNA, plasma, blood filter cards, serum, hair, or tissue).
- Storage location and sample availability
- Extraction and allocation amounts
- Client, billing, and laboratory supply information

PEDIGENE® ensures the integrity of sample processing

To increase DNA Bank efficiency, many variables are tracked and databased through Nautilus™ LIMS at sample acquisition. These variables include:

- Quantity and quality of samples obtained
- Elapsed time between blood drawing and processing
- Availability of alternative DNA sources (cell lines, tissue, frozen blood, or blood filter cards)
- Processes or tests to be performed.

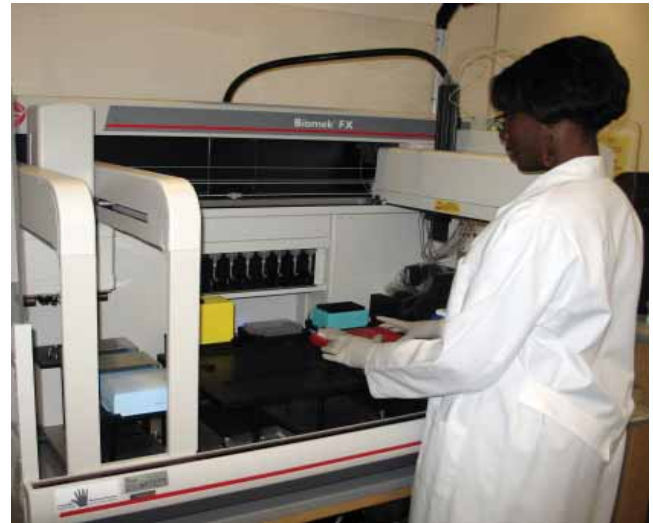
Sample information integrity is maintained through the use of bar codes, utilized from start to finish in the acquisition to genotyping phases. Sample data are automatically linked to the corresponding study, clinical, and auxiliary information.

PEDIGENE® simplifies billing and report creation

PEDIGENE®'s use of the Nautilus™ LIMS makes billing easy and accurate. Because each sample is tracked through log-in and processing, billing records can be created for any service performed on a sample. Such billing allows for consistency and flexibility in billing specific clients for different services. Billing invoices or reports are created in a few easy steps within Nautilus™ LIMS.

PEDIGENE® enhances competitive advantage

Nautilus™ LIMS provides power and flexibility of use through its comprehensive integration of data within the PEDIGENE® system. Competitive advantage in research depends upon the speed at which data are generated, accessed, analyzed and presented to the research community. The PEDIGENE® database design team is constantly developing new methods to organize and analyze the massive amounts of data required to isolate the genetic effects for common, complex genetic disease.



Assures quality control management

The DNA Bank and Tissue Repository Module simplifies data entry while protecting sensitive data. Nautilus™ LIMS tracks primers, probes, and other lab stock thereby providing the capacity to bulk order supplies and reduce laboratory costs.

Enhances competitive advantage

Nautilus™ LIMS facilitates integration of robots and genotyping equipment. Automated data collection and transfer reduces time and error. Customized and real-time reporting provides a competitive edge in human genome research.

Nautilus - plab
File Edit View Window Help

Exploring - Samples Logged In Today

Samples Logged In Today

DNA Bank

- Studies
- Samples
- Samples Logged In Today
 - 200801406
 - 200801407
 - 200801408
 - 200801409
 - 200801410
 - 200801411
 - 200801412
 - 200801413
 - 200801414
 - 200801415
 - 200801416
 - 200801417
 - 200801418
 - 200801419

Report Preview - Billing for DNA Bank by PI Summaries

100%

DUKE center
HUMAN GENOMICS

DUKE UNIVERSITY MEDICAL CENTER, Box 29
DNA BANK, GENOME SCIENCES RESEARCH BLDG
595 LaSalle Street, Room 3027
Durham, NC 27710

Bill to: Allison Ashley-K
Box 3445 DUMC
684-1805

Molecular Genetics Module

PEDIGENE® users readily access molecular analysis results

The PEDIGENE® Molecular Genetics Module is designed to provide users with access to accurate and reliable genotyping and mutation results with minimal effort. Data preparation steps have been streamlined to facilitate efficient and rapid statistical analyses. Each dataset undergoes rigorous quality control testing to verify that performance and quality control requirements are met.

PEDIGENE® facilitates large-scale SNP genotyping

Genomics Research Laboratory personnel perform large-scale genotyping of single nucleotide polymorphisms (SNPs) using the Illumina Beadstation 500G high throughput genotyping system. This platform allows researchers to genotype hundreds of thousands of SNPs per day. Smaller-scale genotyping is done using the ABI 7900 high-throughput DNA analysis system with the TaqMan assay. The CHG mutation laboratory is fully equipped, utilizing three Transgenomic WAVE Denaturing High-Performance Liquid Chromatography (DHPLC) units in conjunction with ABI 3100 and 3730 capillary electrophoresis sequencing to identify and verify all DNA variations and mutations. Experimental results from each of these platforms is submitted to standardized quality control procedures.

PEDIGENE® rigorously maintains genotyping integrity

Quality control samples are included in each experiment to provide within and across plate verification. CEPH samples are also included in each experiment, with the results automatically compared against databased values. Mendelian checks (where appropriate) and Hardy-Weinberg Equilibrium are performed on all experiments to assess genotyping quality. All problems must be resolved by the Genomics Research Laboratory personnel before the data is released for further analyses. To assist with these checks, complete descriptions of all markers are stored in the database, including primer sequences, allele sizes, frequencies for various control populations, map locations, sample quality and marker performance.

PEDIGENE® provides state-of-the-art bioinformatics resources

The Center for Human Genetics (CHG) has state-of-the-art bioinformatics resources. These include microarray databases and a full range of modern microarray data analysis tools. In addition, the SNPSelector module of PEDIGENE® is a web application for selecting SNPs for linkage studies or association studies at the level of the gene or the genome. SNPs are selected according to their quality, as defined by criteria including function, minor allele frequency, and linkage disequilibrium.

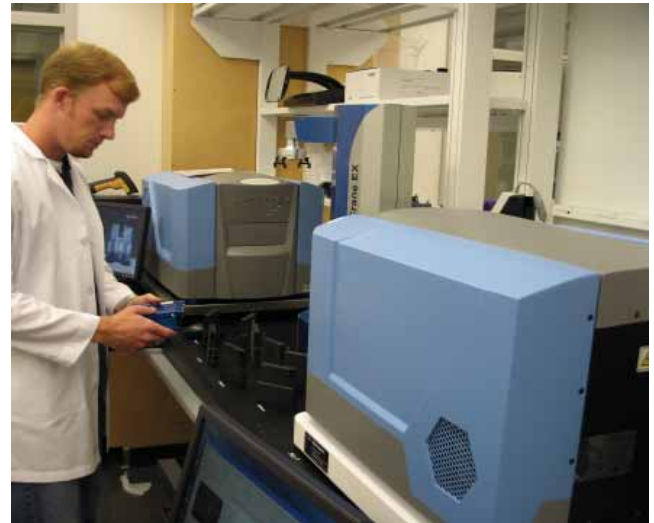
The CHG Variation Database is a proprietary repository for storing and analyzing human genome sequence variations detected in the CHG. It is used to link gene and genome variation to clinical data, and to facilitate research into genotype/phenotype correlations in Mendelian and complex diseases.

Bioinformatics resources are constantly being expanded and updated to remain at the forefront of this burgeoning field of research.



Process large blood and tissue sample shipments

Laboratory personnel can rapidly process large shipments of blood and tissue samples with the AutoPure LS® DNA Extractor and PEDIGENE®.



Identify and verify DNA mutations

The Illumina Beadstation whole-genome association studies allow comprehensive genomic coverage via HumanHap550 Genotyping Beadchips which contain more than 550,000 SNP loci with high-density tagSNP content.



Perform large-scale SNP genotyping

The ABI 7900 High Throughput DNA analysis system linked to the PEDIGENE® Molecular Genetics Module makes large-scale SNP genotyping quick and easy.



Handle large orders of samples efficiently and expeditiously

Using barcode tracking and custom robotics handling, lab personnel can easily manage multiple storage formats and platforms.

Statistical Analysis Module

LAPIS decreases genetic analysis time

The plethora of genetic analysis programs available today indicates a need for quick transitions between program input formats. We have developed the Linkage Analysis Pedigree Input System (LAPIS) which acts as an interface between the PEDIGENE® database and the genetic analysis programs. LAPIS provides options for automated data consistency checking and automated analysis based on input parameters. LAPIS supports the genetic analysis programs currently in use by statistical analysts. LAPIS provides a balance between flexibility of analyses by providing a wide range of standardized formats with data integrity by allowing direct access to the database without need for intermediate formats such as spreadsheets or text files.

LAPIS reduces error rates

LAPIS condenses the amount of time needed to prepare family and case/control data for genetic analysis. In addition, data error rates are dramatically reduced using this system.

LAPIS provides data flexibility

Allele frequencies can be calculated within LAPIS as the marker data is extracted from the database using a variety of options including:

- All individuals in the data set
- Unaffected spouses only
- A single affected individual from each family
- A single unaffected individual from each family
- A given set of control individuals

Families can be arranged in subsets for analysis, and counts of affecteds, unaffecteds, and sib pairs obtained.

Marker loci can be ordered within LAPIS using Mega2 mapfile data, so that input to the analysis programs appears in the desired map order. Quantitative data can also be utilized within LAPIS.

LAPIS-supported genetic analysis programs

Nonparametric linkage analysis:

- Genehunter
- SimIBD
- ASM
- ASPEX
- Siblink
- Ordered Subset Analysis

Parametric linkage analysis:

- Vitesse
- Fastlink
- Homog

Quantitative trait analysis:

- SAGE
- Solar

Association analysis:

- SNPEM
- TDT
- PDT
- TRANSMIT

Haplotyping analysis:

- Simwalk

Marker Characterization:

- GOLD

Genotype Error Checking:

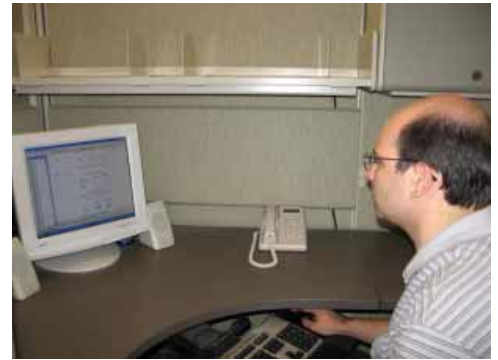
- PREST
- Relative
- Relpair

Power calculations:

- DESPAIR
- GASP
- Simlink
- SIMLA

Whole Genome Analysis::

- PLINK



Decrease analysis time

LAPIS provides quick data transition to currently used genetic analysis programs from the PEDIGENE[®] database.



Reduce error rates

Direct downloading of analyst selected data from the PEDIGENE[®] database drastically cuts error rates.



Gain data flexibility

With the LAPIS interface, analysts extract and order marker data from the PEDIGENE[®] database as needed for each study.

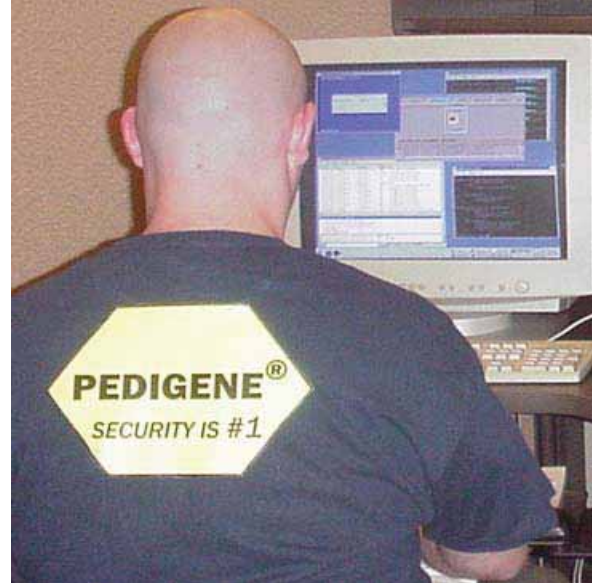
Security and Confidentiality

PEDIGENE[®] safeguards security and confidentiality

Confidentiality and security of data is of paramount importance in genetic studies, and PEDIGENE[®] rigorously safeguards data confidentiality.

In the PEDIGENE[®] system, security extends far beyond simple log-in authorization. Built-in security options protect unauthorized viewing or modification of data. Each PEDIGENE[®] database user is assigned a rolebased security level for each study he or she has access to. The security levels determine if the user can view individual names, disease affection status, genotypic results, and can limit the user to viewing only certain diagnoses and/or family groups. Access privileges must be granted by the supervising principal investigator for each new database user.

An extensive notification system automatically alerts researchers during each step of data acquisition, genotyping, and analysis, while numerous integrity checks ensure data validity.



A Proven Comprehensive Approach to Information Management

THE PEDIGENE® Team: Developing the Optimum System

The PEDIGENE® team is a multidisciplinary group of genetic researchers and database designers at the Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, USA. The primary mission of the Center of Human Genetics is to identify major susceptibility genes and epidemiologic risk factors that contribute significantly to the pathology of many common diseases. Our team has performed groundbreaking research in disorders such as Alzheimer disease, autism, amyotrophic lateral sclerosis (Lou Gehrig's disease), and Parkinson disease. The PEDIGENE® information management system provided the foundation for these research discoveries. It is the intent of the Center for Human Genetics to broaden the understanding of human genetics, to assist in the integration of Duke University Medical Center programs in collaborative research with academic partners, and to develop new therapies for the treatment of disease. PEDIGENE® will play an important role in transforming these goals into realities.

PEDIGENE® An Innovative Database Solution

PEDIGENE® is a registered trademark of the Duke Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, United States of America

PEDIGENE[®]

PEDIGENE[®]

A Proven Comprehensive Approach to Information Management

Duke Center for Human Genetics
Box 3445
Duke University Medical Center
Durham, NC 27710 USA

ph: 919-684-2063
fax: 919-684-0900

<http://www.chg.duke.edu>