EGEA Collection: A Biobank Devoted to Asthma and Asthma-related Phenotypes

Rachel Nadif¹, Emmanuelle Bouzigon², Nicole Le Moual¹ and Valérie Siroux³

¹ INSERM, U1168, VIMA: Aging and chronic diseases. Epidemiological and public health approaches, Villejuif, F-94807, France; Univ Versailles St-Quentin-en-Yvelines, UMR-S 1168, F-78180, Montigny le Bretonneux, FR
² INSERM, U946, Genetic Variation and Human Diseases Unit, Univ Paris Diderot, Univ Sorbonne Paris Cité, Paris, FR
³ Institute for Advanced Biosciences, Centre de recherche UGA-Inserm U1209-CNRS UMR 5309, Site Santé, Allée des Alpes, 38700 La Tronche, FR

Corresponding author: Rachel Nadif (rachel.nadif@inserm.fr)

The Epidemiological study on the Genetics and Environment of Asthma (EGEA) is a 20-year cohort including cases with asthma (n = 388), their first degree relatives (n = 1244) and population-based controls (n = 415) recruited in the early 90’s from five French cities. Participants were extensively characterized regarding environment and respiratory health, and a specific effort was made to trace and bank biological samples (ISO 9001 certification since 2006). Access to data and samples is opened to researchers wishing to develop new scientific collaborative programs. The survey has already led to more than a hundred papers with almost one third in collaboration with national and international teams.

Keywords: Asthma; Asthma-related phenotypes; Environment; Genetics; Biology


(1) Bioresource Overview

Project description
A better understanding of the aetiology of asthma is important for developing preventive measures and new asthma management. The Epidemiological study on the Genetics and Environment of Asthma (EGEA) is aimed to provide public health information for improving the respiratory health of the population. EGEA is a unique 20-year cohort set up to investigate the genetic and environmental factors and their interactions in asthma and asthma-related phenotypes (bronchial hyperresponsiveness, atopy), and to clarify the phenotypic heterogeneity [1, 2]. The first EGEA survey (EGEA1) included cases with asthma, recruited in five chest clinics, their first-degree relatives and population-based controls, recruited in the early 90’s in 5 French cities (total: n = 2047). A follow-up of the participants has been completed in 2003–2007 (EGEA2), including 1601 subjects with complete examination, almost exclusively adults. A third survey (EGEA3) consisting in a self-completed questionnaire has been initiated in 2011. The questions relate to the respiratory health, the allergic status, and environmental factors. Data collection ends in 2013 with an overall response rate of 79.2%. A biological collection is associated to EGEA1 and to EGEA2.

The survey has already led to more than a hundred papers on the phenotypic, environmental, biological and genetic aspects with almost one third in collaboration with French, European or International research teams [3–10]. The scientific projects of the study are organized around 6 working groups (genetics, biology, phenotypes, occupational exposures, air pollution and nutrition). The EGEA study is part of international networks, thanks to its inclusion in European programs, both on environmental (FP7-ESCAPE on Air Pollution) and genetic (FP6-GABRIEL and its extension to an international program, TAGC project) aspects. The EGEA group was leading main papers in these European projects.

Classification (1)

Human.

Species
N/A.
**Classification (2)**
Epidemiological biobank in the field of asthma: biological samples and associated data e.g. personal characteristics, health, environmental, genetic and biological data.

**Context**

**Spatial coverage**
At EGEA1, asthma cases were recruited from respiratory or allergy clinics and controls were recruited from electoral rolls, a check-up centre or surgery clinics. Participants were examined in seven centers from five French cities (followed by latitude and longitude coordinates): Grenoble (45.182359, 5.718384), Lyon (45.754269, 4.855957), Marseille (43.295532, 5.380554), Montpellier (43.602604, 3.878174) and Paris (48.883979, 2.373047). Adults were recruited in all five cities. Children were recruited in Paris, Grenoble and Marseille.

**Temporal coverage**
The start and end dates of the biological collection are 1991 to 1995 (EGEA1), and 2003 to 2007 (EGEA2, 11.4 years apart on average after EGEA1, range 5.3–14.9 years). There is no biological collection at EGEA3 (see Figure 1 for a global overview of the collection).

**Temporal coverage for accessibility**
N/A.

---

### (2) Methods

**Steps**
EGEA collection maintained a high quality of samples and data through a professional quality management system (QM-system). EGEA collection is certified ISO 9001:2008 since 2006. A specific effort was made to trace and bank biological samples.

All work processes during handling and storage of biological samples are based on standard operating procedures (SOPs). A first biological collection has been established at EGEA1 including DNA, serum and plasma. The biological collection has been extended at EGEA2, including lymphoblastoid cell lines, DNA, RNA, serum, plasma, erythrocytes and exhaled breath condensate (EBC) collected with an RTube™ (Respiratory Research Inc., Charlottesville, VA, USA) under standardized procedures. Exhaled fraction of nitric oxide (FeNO) was also available for almost half of the adult participants. Detailed information on design and questionnaires is available at [https://egeanet.vjf.inserm.fr/index.php/en/subheading “presentation”](https://egeanet.vjf.inserm.fr/index.php/en/subheading “presentation”).

**Stabilization/preservation**
DNA, RNA, serum, plasma and erythrocytes were obtained from blood collection tubes according to the workflow described in Figure 2, and preserved as follows:

- **DNA**: buffy-coat obtained from EDTA (ethylene diamine tetra acetic acid) tubes, stored in cryotubes.
Serum: obtained from rapid serum tube with clot activator, stored in straws and in cryotubes.
Plasma: obtained from EDTA and lithium-heparin tubes, stored in straws and in cryotubes.
Erythrocyte: obtained from EDTA and lithium-heparin tubes, stored in straws after addition of glycerol.
RNA: obtained from whole blood preserved in Paxgen tubes for RNA stabilization for no longer than 6 months at –80°C, then extracted and stored in cryotubes.
Lymphoblastoïds cell lines: obtained from peripheral lymphocytes from lithium-heparin tubes, and stored in vials.

Exhaled breath condensate: from RTube™ (Respiratory Research Inc., Charlottesville, VA, USA) according to a standardised method. Briefly, the RTube was rinsed with deionised water and dried thoroughly. Participants breathed orally at tidal volumes into a mouthpiece attached to a cold condenser (–20°C). They were seated comfortably with a headrest. All headrests and back seats were tilted slightly to avoid any saliva contamination during breathing manoeuvres. EBC was stored in cryotubes.

**Type of long-term preservation**
Up to 50 aliquots per participant in two banking facilities has been implemented. The following long-term preservation types are used: cryotubes and straws.

**Storage temperature**
Straws of serum, EDTA or heparin plasma, erythrocytes, and vials of lymphoblastoïds cell lines are conserved at –196°C in the vapor phase of liquid nitrogen. Cryotubes of buffy-coat, serum, EDTA or heparin plasma, and EBC are conserved at –80°C in two distinct biobanks.

**Shipping temperature from patient/source to preservation or research use**
In each collection center, the biological samples were aliquoted within 1:15 hour after blood withdrawal and centrifugation at 4°C if necessary. The aliquots were immediately stored at –20°C and transferred to –80°C within two days. They were shipped to the biobanks on dry ice (–80°C) within the month.

**Shipping temperature from storage to research use**
Biological samples are shipped on dry ice (–80°C).

**Quality assurance measures**
The EGEA collection (environmental, phenotypic, genetic and biological data) is referenced in the Biobanking and BioMolecular Resources Research Infrastructure (BBMRI-ERIC®) with the ID:FR_BB-0033-00043, and certified ISO 9001:2008 and renewed since 2006 (https://certificats-attestations.afnor.org/certification=262711141114). The complete workflow of sample and data handling is embedded in a quality management system. Process control is based on standard operation procedures (SOPs) that can be found in the central workflow system.

**Source of associated data**
A major strength of the EGEA study lies in the detailed phenotypic characterization of the participants. They have been extensively phenotyped through: 1) Detailed questionnaires regarding upper and lower airway symptoms, allergic symptoms, childhood events including infections, medical history, 2) Lung function tests: spirometry with methacholine bronchial challenge, 3) Skin prick tests to 12 allergens, total Immunoglobulin E (IgE) levels, Phadiatop® test, and white Blood Cell count.

Detailed data to environmental factors such as active and passive smoking, presence of domestic animals considering different windows of exposure, indoor domestic occupational exposures and air pollution have been recorded. More detailed data have been recorded at the second survey (EGEA2), more specifically regarding medication, diet and quality of life. A third survey (EGEA3) consisting in a self-completed questionnaire has been initiated in 2011. The questions relate to the respiratory health, the allergic...
status, and environmental factors. Data collection ends in 2013 with an overall response rate of 79.2%. All questionnaires based on the British Medical Research Council/CECA, the American Thoracic Society and the ECRHS questionnaires are accessible on the web site.

(3) Bioresource description

Object name
The EGEA collection

Bioresource name
EGEA collection

Bioresource location
The EGEA biological collection is banked in two locations: 1) Biothèque Internationale Jean Mérieux, EFS Rhône-Alpes, Site d’Annemasse 1, Route de Taninges, (Annemasse F-74100) and 2) Centre d’Investigation Clinique 1403, Hôpital Cardiologique, Bd du Professeur LECLERCQ, (Lille F-59037), FRANCE.

Straws are banked at Annemasse, and cryotubes are duplicated at Annemasse and Lille.

Bioresource contact
E-mail: Egea.cohorte@inserm.fr

Bioresource URL
Web: https://egeanet.vjf.inserm.fr/

Identifier used
The identifier used for clinical data is an anonymized 9-digit number including a 6-digit family number and a 3-digit number for each participant. For any external collaboration requesting access to EGEA data, a new and unique identifier number is generated when transferring the extracted dataset. Only the EGEA coordinator centre keeps the correspondence between the EGEA identification number and the identifier specifically generated for the collaborative project. The identifier used for samples has 3 more digits indicating the type of fluid and the sample number. This identifier is sent to the laboratories in a data file associated with the shipment without any explanation or clinical data.

Bioresource type
Asthma and asthma-related phenotypes (rhinitis, atopic dermatitis, allergic sensitization, bronchial hyperresponsiveness).

Type of sampling
EGEA is a cohort that combines both a family study including asthma cases and a population-based control group.

Anatomical site
N/A.

Disease status of patients/source
Asthma.

Clinical characteristics of patients/source
EGEA combines both a family study and a case-control study. Probands (asthmatics and controls) were ascertained in five French cities (Paris, Lyon, Marseille, Montpellier and Grenoble). All probands and their two parents were born in France. Cases were selected from respiratory or allergy clinics and controls were selected from electoral rolls, a check-up centre or surgery clinics in seven centres of the five cities. Adults were selected in all five cities. Children were recruited in Paris and Grenoble.

Cases and controls were recruited through a standardized protocol. Inclusion criteria included place of birth, area of residence, age (7 to 70 years old), family structure and for cases with asthma a positive answer to four standardized questions on asthma based on international questionnaires (see Clinical diagnosis section below for more information, and references 1 and 2). More than 10,000 self-administered questionnaires were distributed in a systematic way. The questionnaires were distributed to all patients in the respiratory or allergy clinics. Similarly, for the controls, the questionnaires were sent (electoral rolls) or distributed (consultations of surgery clinics, check-up center). The exclusion criteria were: not born in France or parents not born in France, area of residence not adequate, not in the age group chosen, no inclusion criteria on asthma, or family structure not adequate.

Size of the bioresource
EGEA1 included 388 cases with asthma, their first-degree relatives (n = 1244) and 415 population-based controls. At EGEA2, among the alive cohort (n = 2002), 92% (n = 1845) completed a short self-completed questionnaire and 77% (n = 1543) participated in the second phase of the study. 58 new family members were also included. The EGEA2 population consisted of 1601 subjects with complete examination (1570 adults). At EGEA3, 1558 participants had filled in the self-completed questionnaire (response rate = 79.2%).

Vital state of patients/source
Vital status of asthma cases, their first degree relatives and controls has been recorded at each follow-up.

Clinical diagnosis of patients/source
Inclusion criteria chosen to define asthma in cases were as standardized as possible. The phenotype considered was based on self-reported positive responses to four questions, all from international, validated and standardized questionnaires, with the idea of having undisputable asthma according to questionnaire data (favoring specificity over sensitivity). Questions were from the British Medical Research Council (BMRC)/European Coal and Steel Community (ECSC)/American Thoracic Society (ATS), and European Community Respiratory Health Survey (ECRHS) questionnaires: Q1. “Have you ever had attacks
of breathlessness at rest with wheezing?” (BMRC/ESCS/ATS questionnaires), Q2. “Have you ever had asthma attacks?” (BMRC/ESCS questionnaires), if yes Q2a. “Was this diagnosis confirmed by a physician?” (ATS questionnaire), and Q2b. “Have you had an asthma attack in the last 12 months?” (ECRHS questionnaire). Inclusion criteria were also based on 3 (possibly 2) positive answers and a clinical evaluation of medical record by consensus from the five centers. Asthma in relatives of cases was defined as a positive answer to at least one of the first two questions [1, 2].

Pathology diagnosis
N/A.

Control samples
Among the whole EGEA samples, around 50% of the participants do not have asthma. Other diseases are self-reported (cancer, cardiovascular diseases, other respiratory diseases…).

Biospecimen type
DNA, RNA, serum, plasma, erythrocytes, exhaled breath condensate and lymphoblastoids cell lines. Serum, EDTA or heparin plasma and erythrocytes, and EBC are conserved in almost 50,000 cryotubes of 250 µL, and in almost 16,000 straws of 500 µL.

Size of the bioresource
Total number participants: 2,121
% participants with DNA: 97.6%
% participants with RNA: 87.3%
% participants with serum/plasma/erythrocytes: 89.4%
% participants with EBC: 85.1%
Genome-wide SNP data (Illumina 610K chip) and imputed data from Hapmap2 (2.5 Million SNPs) and 1000 Genome samples was registered under number IE-2011-591.

Release date
Data and biological samples are currently available according to the access criteria below.

Access criteria
The EGEA study is a “research platform”, with open access to the data and biobank to researchers (academic or industrial) wishing to develop new scientific program in collaborations with EGEA. The EGEA study is coordinated by V. Siroux (epidemiology, Principal Investigator), F. Demenais (genetics), I. Pin (clinical aspects) and R. Nadif (biology). Submitted scientific project (form located on https://egeanet.vjf.inserm.fr/index.php/en/contacts-en) are systematically reviewed by a steering committee for approval (in particular to avoid any overlap with ongoing projects). Applicants may be asked to pay a processing charge to data extraction and to shipment from the biobanks. Researchers are required to sign a materials transfer agreement, confirm that all samples have been used up at the end of their permission, and provide raw experimental data back to the EGEA.

(4) Reuse potential
Data from questionnaires and clinical examination may be reused according to access criteria above. A new approval from the EGEA coordination must be obtained to reuse biological data obtained from samples involved in research projects previously accepted.

EGEA cooperative group
Coordination: Siroux, V (epidemiology, PI since 2013); Demenais, F (genetics); Pin, I (clinical aspects); Nadif, R (biology); Kauffmann, F (PI 1992–2012).

Respiratory epidemiology: Inserm ex-U 700, Paris: Korobaeff, M (Egea1), Neukirch, F (Egea1); Inserm ex-U 707, Paris: Annesi-Maesano, I (Egea1-2); Inserm ex-U 1018, Villejuif: Kauffmann, F, Oryszczyn, M P (Egea1-2); Inserm, U 1168, Villejuif: Le Moual, N; Nadif, R, Varraso, R; Inserm U 1209 Grenoble: Siroux, V.

Genetics: Inserm ex-U 393, Paris: Feingold, J; Inserm U 946, Paris: Bouzigon, E; Demenais, F, Dizier, M H; CNG, Evry: Gut, I (now CNAG, Barcelona, Spain), Lathrop, M (now Univ McGill, Montreal, Canada).

Clinical centers: Grenoble: Pin, I, Pison, C; Lyon: Ecochard, D (Egea1), Gormand, F, Pacheco, Y; Marseille: Charpin, D (Egea1), Vervloet, D (Egea1-2); Montpellier: Bousquet, J; Paris Cochin: Lockhart, A (Egea1), Matran, R (now in Lille); Paris Necker: Paty, E (Egea1-2), Scheinmann, P (Egea1-2); Paris: Trousseau: Grimfeld, A (Egea1-2); Just, J.

Data and quality management: Inserm ex-U 1555 (Egea1): Hochez, J; Inserm U 1168, Villejuif: Le Moual, N; Inserm ex-U 780: Ravault, C (Egea1-2); Inserm ex-U 794: Chateigner, N (Egea1-2); Grenoble: Quentin, J (Egea1-2).

Constraints
None, the authorization to import/export the biological samples was registered under number IE-2011-591.

Ethics Statement
INSERM, RBM “Recherche BioMédicale” RBM 91-005 and RBM 01-11. CNIL “Commission Nationale de l’Informatique et des Libertés” n°109427 (04/1990), n°900198 (10/2000) and n°1769319 (2014). Ethical approval from the relevant institutional review board committee(s)(n°01-07-07,04-05-03, 04-11-13 and 04-11-18). DGS “Direction Générale de la Santé” n°2002/0106 and n°190048. Written informed consent was signed by all participants.

In the informed consent, it was specify that biological samples will only be used for scientific purposes and stored without any time limit for research in the context of this program. Research may be conducted in collaboration with public or private partners to enable faster or more efficient research. In this context, anonymized biological samples may be exchanged between laboratories.

Competing Interests
The authors have no competing interests to declare.

Author Roles
Rachel Nadif, EGEA Biobank coordinator.
Emmanuelle Bouzigon, EGEA Genetic database coordinator.
Nicole Le Moual, EGEA Epidemiological database coordinator.
Valérie Siroux, EGEA coordinator.

References