DATA PROCESSING AGREEMENT
FOR THE SUBMISSION AND DISTRIBUTION OF PERSONAL DATA BY THE EUROPEAN GENOME-PHENOME ARCHIVE

Version 1.4, June 2024

Upon completion of the Data Processing Agreement, the Data Controller, on its own behalf or on behalf of or upon the mandate of the Data Producer submits Personal Data (genetic and phenotypic data) of research participants to the European Genome-phenome Archive (EGA). Personal Data is made accessible and distributed upon granted access by the Data Access Committee.

This Data Processing Agreement (hereinafter DPA) addresses the requirements of Data Protection Laws for processing Personal Data by EMBL and CRG, as the Data Processor on behalf of the Data Controller, and the relationship between and accountability of the Data Controller and the Data Processor in processing Personal Data, as well as their respective rights and obligations.

1. INTRODUCTION

1. The European Genome-phenome Archive (EGA) is a service for permanent archiving and sharing of all types of personally identifiable genetic and phenotypic data (Personal Data) resulting from biomedical research projects, jointly managed by the European Molecular Biology Laboratory (EMBL) and the Centre for Genomic Regulation (CRG). The submitted Personal Data are accessible and distributed under controlled access policy, whereby access decisions reside with the Data Access Committee (DAC), created or defined by the Data Producer for each dataset, and covered by a Data Access Agreement (DAA), defining the terms and conditions of the use of a specified dataset.
2. The European Molecular Biology Laboratory is an intergovernmental institution headquartered in Heidelberg, Germany, established by an agreement of 1973 (link) and enjoying privileges and immunities within the framework of its founding act of 1973, general principles of public international law and conventions signed with its host countries. Accordingly, it has the power to self-regulate data protection reflecting the essential principles of European Data Protection Law and focusing on scientific research as necessary for the fulfilment of its objectives and for the exercise of its functions by way of its Internal Policy No 68 on General Data Protection (IP No 68).
3. The European Bioinformatics Institute (EMBL-EBI) is the UK site of EMBL, established in Hinxton, UK, upon the Agreement of 1995 between the Government of the United Kingdom of Great Britain and Northern Ireland and EMBL concerning the European Bioinformatics Institute, enjoying the same position, privileges and immunities as EMBL and exercising the activities of EGA for EMBL.
4. The Centre for Genomic Regulation (Fundació Centre de Regulació Genòmica – CRG) is a public non-profit Spanish foundation, mainly governed by the Generalitat of Catalonia through the

Departments of Research and Health, and additionally by the Pompeu Fabra University, the Spanish Ministry of Science & Innovation and "la Caixa" Banking Foundation. The mission of the CRG as an international biomedical research institute of excellence is to discover and advance knowledge for the benefit of society, public health and economic prosperity. CRG has the status of research centre in Catalonia, being identified as CERCA centre. CRG is subject to GDPR according to Article 3(1) of the GDPR.

5. Personal Data submitted to and stored by EGA were collected by the Data Producer and/or the Data Controller from individual Data Subjects (research participants) whose consent agreements or informed consent forms authorise data release only for specific research use.

2. PURPOSE OF THIS DPA

The purpose of this DPA is to identify processing activities of the Data Processor on behalf of the Data Controller, whilst providing its service (the EGA Service/s) to the Data Controller and the global scientific community, to store submitted dataset(s) of genetic and phenotypic Personal Data with EGA and to distribute these dataset(s) upon the Data Controller's and/or Data Producer's DAC approval to various Recipients, requesting access to these data. Signature of this DPA entails the commitment to comply with the rights and obligations defined herein and with the requirements established by the applicable Data Protection Laws regarding processing of the Personal Data defined in Section 3.

3. PROCESSING OF PERSONAL DATA

For the purposes outlined in Section 2 and further detailed in Appendix 1, EMBL and CRG act as Joint Data Processors (hereinafter referred to jointly as the Data Processor) of Personal Data of research participants ("Data Subjects") embedded in the submitted dataset(s), on behalf of the Data Controller that submits the Personal Data to the EGA.

The Data Processor provides contact details of the Data Protection Officer or any other person mandated to give information on data protection matters of the Data Processor on its Internet page https://ega-archive.org/data-protection/privacy-notice/ega-dac.

4. TRANSFER OF DATA AND APPLICABLE DATA PROTECTION LAWS

Any Data Controller subject to GDPR, may transfer Personal Data to the Data Processor, partly being managed by an international organisation, upon relying on a derogation of the transfer being necessary for important reasons of public interest under Article 49(1)(d) and Article 49(4) of the GDPR.

Any Data Controller not subject to GDPR and/or subject to any other national law on data protection shall duly rely on the appropriate legal basis for transfer of Personal Data to the Data Processor, under the Data Controller's applicable law.

Either party shall comply with Data Protection Laws applicable on its operations.

5. DATA PROCESSOR'S RESPONSIBILITIES:

1. The Data Processor shall only process Personal Data as instructed by the Data Controller, as documented in this DPA, including with regard to transfers of Personal Data to a third country or an international organisation. The Data Processor may also process the Personal Data where it is

required to do so by any applicable law; in such case, the Data Processor will inform the Data Controller of that legal requirement before Processing the relevant Personal Data, unless that law prohibits such information on important grounds of public interest.

2. The Data Processor will use reasonable efforts to follow any other Data Controllers instructions as long as they are required by the Data Protection Laws, applicable to Data Processor, technically feasible and do not require changes to the EGA Service.
3. The Data Processor shall, as soon as reasonably possible, inform the Data Controller if the Data Processor believes that any instruction of the Data Controller is in breach of Data Protection Laws, or is otherwise unable to implement it.
4. The Data Processor shall process Personal Data only for the purpose of providing it to the scientific community through the Data Processor.
5. The Data Processor shall implement and maintain appropriate technical and organisational measures as set out in Appendix 2. The Data Controller understands and agrees that these measures are subject to technical progress and development, and the Data Processor is therefore expressly allowed to implement alternative measures provided they maintain or exceed the general security level described in Appendix 2.
6. The Data Processor shall ensure that confidentiality applies to Personal Data and that access is strictly limited to the personnel who have committed themselves to confidentiality and have received appropriate training of their responsibilities.
7. The Data Processor shall not link or combine Personal Data to other information or archived data available in a way that could re-identify research participants.
8. Taking into account the nature of the Processing, Data Processor shall assist Data Controller by appropriate technical and organisational measures, insofar as this is reasonably possible, for the fulfilment of the Data Controller’s obligation to respond to requests for exercising the Data Subject's rights laid down in Chapter III of the GDPR and/or any Data Protection Laws applicable to the Data Controller in case it is not subject to GDPR.
9. The Data Processor shall keep all of the Personal Data secure from any unauthorised or accidental use, access, disclosure, damage, loss or destruction.
10. The Data Processor shall without undue delay after becoming aware of and in accordance with Data Protection Laws, notify the Data Controller of any confirmed incident concerning unauthorised destruction, loss, alteration, disclosure of, or access to Personal Data (“Security Incident”) via Data Controller’s contact point, shared with the Data Processor at the time of the submission of Personal Data. The Data Processor shall together with the notification provide the contact details of the Data Protection Officer or other contact point where more information about the incident can be obtained.
11. Upon request of the Data Controller, the Data Processor shall delete or return all Personal Data to the Data Controller.

12. The Data Processor shall make available to the Data Controller all information, where necessary and technically feasible, for the Data Controller to demonstrate compliance with its obligations.
13. Data Processor shall assist the Data Controller in ensuring compliance with its obligations pursuant to Articles 32 to 36 of GDPR and/or any similar obligations under Data Protection Laws applicable to the Data Controller in case it is not subject to GDPR, taking into account the nature of the processing and the information available to the Data Processor.
14. Data Processor shall redirect all third party requests regarding Personal Data or information about the Processing activities to the Data Controller, whether the request is made by a Data Subject, a Supervisory Authority or any other third party, unless such requests cannot legally be redirected to the Data Controller.
15. Data Processor shall distribute Personal Data submitted by Data Controller only upon decision to grant access to specific dataset(s) made by the relevant DAC.

6. USE OF SUB-PROCESSORS

The Data Processor may use one or more Sub-Processors to process Personal Data during the course of submission, storage, and distribution of Personal Data. The Data Processor must enter into a written agreement with each Sub-Processor, requiring the Sub-Processor to comply with terms no less protective than these Terms.

Sub-Processor(s) may only Process Personal Data for the purposes of this DPA. The Data Processor remains responsible for all Sub-Processors it uses to carry out the Processing.

The Data Processor shall keep record of a list of Sub-Processors that will be provided to the Data Controller upon request. Any changes thereof will be notified on Data Processor's website. Unless the Data Processor receives an objection to a change within 30 days of the notification, that change will be deemed approved. In case of Data Controller's objection, both Data Controller and Data Processor shall use all reasonable endeavours in good faith to resolve the objection.

7. DATA CONTROLLER'S RESPONSIBILITIES

Data Controller is responsible for ensuring that Personal Data embedded in datasets and transmitted to the Data Processor are pseudonymised and encrypted. For the avoidance of doubt, it is clarified that the Data Processor shall be entitled to temporarily decrypt the files for the purposes outlined in Section 2 hereof and further detailed in Appendix 1 hereto.

Data Controller is responsible at each time:

- to demonstrate the existence of Data Subjects' informed consent for participation in the research project or any other suitable legal basis for processing Personal Data,
- to obtain appropriate ethical or other approvals for collecting Personal Data embedded in the submitted datasets,
- to comply with and to be able to demonstrate compliance with applicable Data Protection Laws,
- to control distribution of Personal Data, and

- to notify the Data Processor of any change in the identity and/or the contact details of the responsible Principal Investigation or the DAC.

In case of change in the identity of the Data Controller, for example due to transfer of the Personal Data (and the Data Controller's rights and obligations) to another institution, the Data Controller party hereto shall notify the Data Processor in advance, to allow for the timely conclusion of a new DPA between the Data Processor and the new Data Controller.

8. AUDIT

The Data Processor will contribute to audits conducted by the Data Controller by providing appropriate documentation. In exceptional cases and upon prior agreement with the Data Processor, the Data Controller may conduct on-site inspections. Any direct or indirect cost incurred by the Data Processor through an audit request will need to be reimbursed by the Data Controller, and any on-site inspection will need to be (i) pre-agreed in terms of objective, scope, timing and process, and (ii) without prejudice to the privileges and immunities granted to EMBL.

9. TERM AND TERMINATION

This DPA takes effect on the Effective Date, on which the submission of the Personal Data to the Data Processor will be processed and will remain in effect whilst any part of the Data Controller's data is stored by the Data Processor. It can be terminated by either party with 60 days' notice, provided that any Personal Data provided by the Data Controller will be returned or deleted upon expiration of the notice period.

10. LIABILITY AND INDEMNIFICATION

Data Controller shall indemnify and hold harmless the Data Processor (EMBL and CRG) from and against any losses, damages and expenses (including without limitation legal fees) awarded against the Data Processor (EMBL and CRG) and arising from a claim brought as a result of the Data Controller's breach of its obligations under this DPA or applicable Data Protection Laws ("Claim"); provided, however, that the indemnity shall not extend to any Claim arising from (i) a negligent act or omission of the Data Processor (EMBL and CRG) and/or their personnel, (ii) any misconduct by the Data Processor (EMBL, CRG)and/or their personnel and/or (iii) any breach of this DPA or applicable Data Protection Laws by the Data Processor (EMBL and CRG).

Each Party shall be deemed liable for the damage caused by any of its processing activities that do not comply with its obligations under this DPA and applicable Data Protection Laws.

The Data Processor (EMBL and CRG) shall be exempt from any liability under the previous paragraph if they prove that they are not in any way responsible for the event giving rise to the damage.

11. DISPUTE RESOLUTION

The Parties shall endeavour to settle their disputes amicably.

Any controversy or claim arising out of, or relating to, this DPA (including the enforceability or breach thereof, any question regarding its existence, validity or termination) or relating to the EGA Service shall be resolved using the internal dispute resolution mechanisms of EGA as follows:

Initially, the EGA Operational Phase will take place with meeting/s between the Data Controller and the relevant EGA staff to resolve the dispute amicably. On a second stage, if the controversy is not solved, the Legal Management Phase will be implemented, through the meeting/s between the Legal teams of EMBL, CRG and the Data Controller with the aim to come to an understanding of the dispute and agree on a resolution. In case the dispute is not solved, the third phase will consist of the Direction Management Phase, with the negotiation between the legal representatives of the EMBL, the CRG and the Data Controller.

Once these Phases have been exhausted and, at the end, the Parties have not been able to resolve the dispute, the Parties shall refer the dispute to arbitration under the WIPO Expedited Arbitration Rules ("Rules"), which Rules are deemed to be incorporated by reference into this section (Arbitration Phase). Notwithstanding the foregoing, the arbitrator shall not be authorised to award punitive damages with respect to any such claim or controversy, nor shall any party seek punitive damages relating to any matter under, arising out of or relating to this DPA or the EGA Service in any other forum. If any arbitration is commenced by either Party, the substantially prevailing Party in that arbitration or action is entitled to recover from the other Party its attorneys' fees and costs (including arbitration fees and costs and expert witness fees) incurred in connection therewith. The entire arbitration shall be conducted and concluded in no later than ninety (90) days after service of the arbitration demand, unless the arbitrator has reasonable grounds to extend this deadline, and it may do so without the approval of either party. The extension and the reasoning thereof shall be notified to both parties. A written demand for arbitration must be delivered within one (1) year from the date on which the EGA Services to which the claim relates were provided. Failure to comply with this provision shall be a complete bar to any claim. The place of arbitration will be the legal seat of the defendant party and, in case of a collective claim against EMBL-EBI and CRG, the place of arbitration shall be Barcelona. The language to be used in the arbitral proceedings shall be English, unless otherwise agreed upon, and the governing substantive law to be used shall be of England and Wales.

Nothing herein shall be deemed or interpreted as a waiver, express or implied, of any privileges or immunities accorded to EMBL, being one managing party of the Data Processor, by its constituent documents or international law or as the acceptance by EMBL of the jurisdiction of (i) the courts of any country, including in case of injunctive relief sought, or (ii) any national regulatory and/or supervisory authority.

12. DEFINITIONS

Unless otherwise agreed or defined in these Terms, all capitalised terms used will have the meanings given to them below:

"Data Controller", "Data Processor", "Personal Data", "Data Subject", "Processing", "Genetic data", "Pseudonymisation" shall have the same meaning as described in the GDPR, without prejudice to EMBL's privileges and immunities.

For the sake of clarity, "Data Controller" shall mean an organisation that submits the Personal Data to the Data Processor on its own behalf or on behalf of the Data Producer.

“Data Producer” shall mean an organisation that collected the samples and generated any associated analyses of the Personal Data. Where the Data Producer submitted those data to the Data Processor, it shall also be the Data Controller.

“DAC” means Data Access Committee, a body of one or more individuals, named by the Data Producer and/or Data Controller, granting data release to external requestors, these being Recipients.

“Data Protection Laws” shall mean the GDPR and any national implementing laws, regulations and secondary legislation and any other laws and regulations relating to the processing of Personal Data and privacy which apply to a Party; and, if applicable, the guidance and codes of practice issued by any competent data protection supervisory authority; and for the Data Processor, partly managed by an intergovernmental institution, EMBL’s IP No 68 and any other rules regulating data protection, as in force at each time, further provided that any reference to GDPR or any other national implementing law in relation to EMBL is simply for convenience and does not imply a waiver of any privileges and immunities applicable to EMBL.

“EGA service” shall mean the data centre facilities, servers, networking equipment, and host software systems (e.g., virtual firewalls) that are under Data Processor’s control and used to provide its Services.

“EGA Security Standards” shall mean the security standards attached to these Terms as Appendix 2.

“Security Incident” shall mean a breach of Data Processors’ security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, Data Controller’s data on systems managed or otherwise controlled by Joint Data Processors. Security incidents do not include unsuccessful attempts or activities that do not compromise the security of Data Controller’s data (including unsuccessful log-in attempts, pings, port scans, denial of service attacks and other network attacks on system firewalls or networked systems).

“Sub-Processor” shall mean any processor which the Data Processor engages to carry out specific processing activities on behalf of the Data Processor.

“Effective date” shall mean the date on which the Data Controller accepted this DPA.

“Recipient” shall mean any natural or legal person, to whom the Personal Data are disclosed upon DAC granted access to the data.

“Supervisory Authority”; as long as GDPR applies, shall mean a supervisory authority of a Member State concerned pursuant to Article 55(2) of the GDPR, and for EMBL, shall mean the Data Protection committee pursuant to Article 20 of its Internal policy No 68.

13. SIGNATURE

This Agreement shall be executed by electronic means, having the same legal effect and enforceability as being executed by original handwritten signatures. This Agreement shall enter into force upon its signature by the Data Controller (the Effective Date, as defined in Article 12(9) hereof).

SIGNED for and on behalf of:	SIGNED for and on behalf of:
Name:	Name:
Position:	Position:
Signature:	Signature:
Date:	Date:

SIGNED for and on behalf of:

Name:

Position:

Signature:

Date:

Appendix 1

Subject Matter and Details of the Data Processing

Subject Matter

Data Controller's submission of Personal Data to EGA service, EGA processing of such data, and distribution to recipients upon DAC granted access.

Duration of the Processing:

Processing takes place from the submission of data from the Data Controller to the Data Processor and until the end of Data Processor's services, including, if applicable, any period during which provision of the EGA Services may be suspended and any post-termination period during which the Data Processor may continue providing the Services requested by the Data Controller for transitional purposes, unless deletion or return of the data is requested by the Data Controller.

Nature and Purpose of the Processing

Provision of the EGA Services for the purpose of sharing deposited data with the international scientific community and the general public.

EGA services consist of all activities required for the achievement of the aforementioned purpose, including the following:

1. facilitation of data discovery;
2. ensuring datasets are updated to current standards;
3. application of data compression; and
4. generation and provision of quality control metrics, such as, for example, validation of sample species, sex, ancestry, tumour-normal sample, as well as the identification of sample duplication and the curation of study type and reference genome.

Categories of Personal Data processed (special data)

Genetic and phenotypic data of research participants.

The Categories of Data Subjects to whom the Project Personal Data relate

Research participant - individuals/Data Subjects whose explicit consent for processing personal data is provided to the Data Producer and/or the Data Controller in the relevant consent forms or agreements that authorise data release for specific research purposes.

Appendix 2

European Genome-phenome Archive: Security Overview

Version 1.0, March 2019

Authors: Thomas Keane (EMBL-EBI), Dylan Spalding (EMBL-EBI), Jordi Rambla (CRG, Barcelona), Arcadi Navarro (CRG, Barcelona), Paul Flicek (EMBL-EBI), Helen Parkinson (EMBL-EBI)

The European Genome-phenome Archive (EGA) is a controlled access archive for consented human data. The EGA does not grant or deny access to data, this is done by the Data Access Committee (DAC) of the relevant Data Controller and EGA applies these permissions to the access to data on behalf of the Data Controller (Figure 1). This document provides an overview of EGA's practices in ensuring the security of data stored at EGA. As security is a prime concern of the EGA, the EGA is a member of the Global Alliance for Genomics and Health (GA4GH - https://www.ga4gh.org/) Data Security work stream. The EGA contributes and helps develop the recommendations outlined the GA4GH Security Technology Infrastructure document[1], which defines guidelines, best practices, and standards for building and operating an infrastructure that promotes responsible data sharing in accordance with the GA4GH Privacy and Security Policy[2].

[1] https://github.com/ga4gh/data-security/blob/master/DSIP/DSIP_v4.0.md
[2] https://www.ga4gh.org/product/data-privacy-and-security-policy/

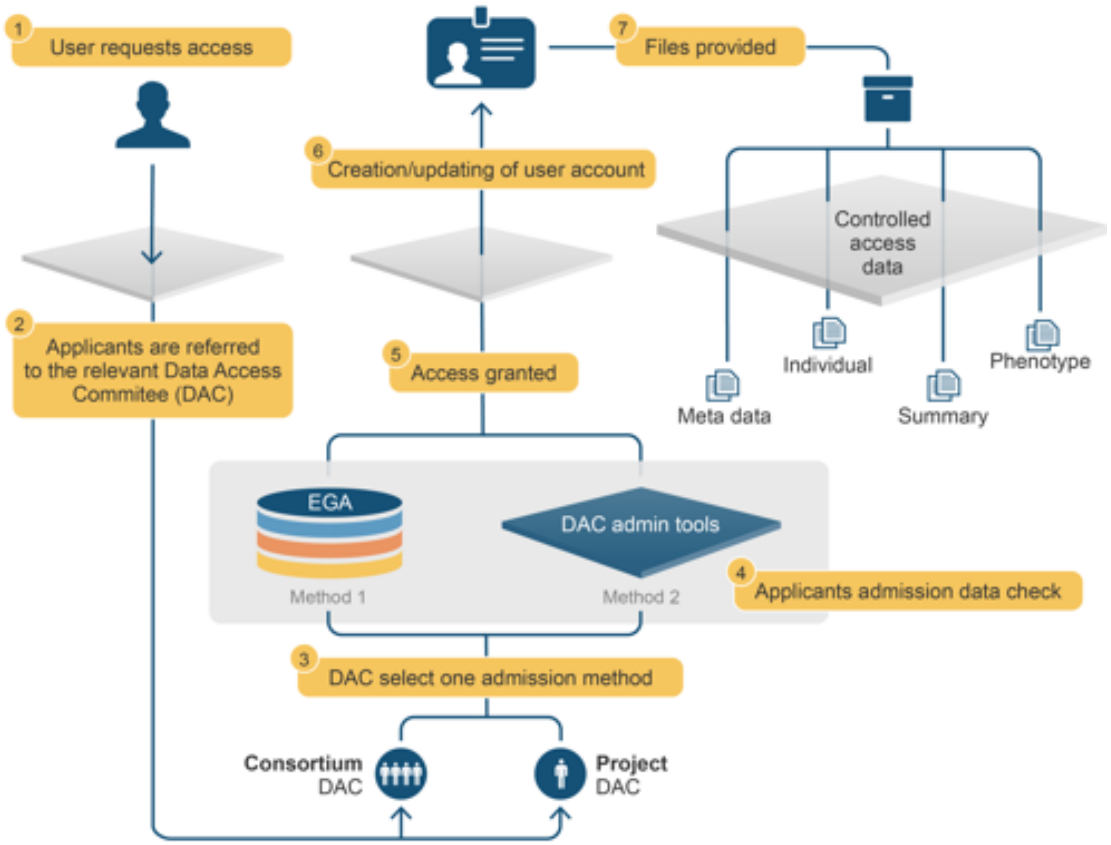

Figure 1: Process of applying for access to data held in the EGA. The user makes a request to access data controlled by a DAC. The DAC informs the EGA of the decision, and if access is granted, the EGA creates an account for the user (if the user does not already have an account) and grants permission to the data for that user.

The key points of EGA security strategy are:

1 Regular Risk Assessment

- The EGA regularly identifies and assesses risk related to the following:
    - Breach of confidentiality,
    - Breach of privacy or autonomy,
    - Malicious or accidental corruption or destruction of data archived at EGA,
    - Disruption of services provided by the EGA.

2 Risk mitigation

- The EGA implements and maintains safeguards to minimise the risks identified above in accordance with the 5 control objectives identified below and outlined in the GA4GH Security and Infrastructure document[3].
- If a data breach is discovered, the EGA applies a defined protocol to minimise damage.

3 Identity and authorisation management

- The EGA authenticates the identity of individuals or software accessing controlled access data held at the EGA.
- The EGA ensures an appropriate level of assurance (LoA) is applied to the identity consistent with the risk associated with that individual, such as multi-factor authentication for DACs.
- The EGA provides the minimum access rights and privileges consistent with the user's identity, allowing access consistent with the GA4GH Privacy and Security Policy, as determined by the appropriate DAC.

4 Audit Logs

- The EGA maintains a set of logs recording:
    - Changes to user access rights,
    - Data access requests,
    - Resource usage.

5 Cryptography, communication security, and data integrity

- The EGA ensures data transmission integrity using a hash function.
- All data transmitted to or from the EGA is end-to-end encrypted.
- All data at EGA is stored using strong encryption.
- Encryption keys are not stored in the same system as the encrypted data.
- All data archived at EGA must be accompanied by a signed submission statement ensuring appropriate consent or ethical approval has been obtained, and is in accordance with all applicable laws and regulations.

The EGA has a defined protocol defining the response in the event of a security breach, and is continuing to work with the GA4GH Data Security Work Stream to help define best practice and associated standards for breach responses.

[3] https://github.com/ga4gh/data-security/blob/master/DSIP/DSIP_v4.0.md

6 Control Objectives

The following control objectives are defined with the aim to implement technology safeguards to prevent the incidences identified below:

- Control Objective 1: Unauthorised access, use, or disclosure of confidential and private data.
- Control Objective 2: The discovery, access,and use of individuals' genomic and health-related data, and individual identities, other than as authorised by applicable jurisdictional law, institutional policy, and individual consents.
- Control Objective 3: Accidental or malicious corruption or destruction of data.
- Control Objective 4: Disruption, degradation, and interruption of services enabling access to data.
- Control Objective 5: Potential security attacks and misuse of authorised accesses and privileges.