DELIVERABLE 11.12

NEC – REQUIREMENT NO. 15

WORK PACKAGE NUMBER: WP11

WORK PACKAGE TITLE: ETHICS REQUIREMENTS

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Document Information

iReceptor Plus Project Information	
Project full title	Architecture and Tools for the Query of Antibody and T-cell Receptor Sequencing Data Repositories for Enabling Improved Personalized Medicine and Immunotherapy
Project acronym	iReceptor Plus
Grant agreement number	825821
Project coordinator	Prof. Gur Yaari
Project start date and duration	1st January, 2019, 48 months
Project website	http://www.ireceptor-plus.com

Deliverable Information	
Work package number	WP11
Work package title	Ethics requirements
Deliverable number	D11.12
Deliverable title	NEC – Requirement No. 15
Description	Personal data are transferred from a non-EU country to the EU (or another third state), confirmation that such transfers comply with the laws of the country in which the data was collected must be specified in the grant agreement and submitted as a deliverable.
Lead beneficiary	time.lex
Lead Author(s)	Liesa Boghaert; Jos Dumortier
Contributor(s)	
Revision number	
Revision Date	

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Status (Final (F), Draft (D), Revised Draft (RV))	D
Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to a group specified by the consortium (RE), Confidential for consortium members only (CO))	CO

Document History			
Revision	**Date**	**Modification**	**Author**
1	09.06.2019	Initial version	Jos Dumortier and Liesa Boghaert
1	24.06.2019	Review and edits	Gur Yaari
1	25.06.2019	Correct names	Milena Mirkis
1	26.06.2019	Comments	Artur Rocha
1	26.06.2019	Comments	Brian Corrie
2	28.06.2019	Final version	Jos Dumortier and Liesa Boghaert

Approvals				
	Name	**Organisation**	**Date**	**Signature (initials)**
Coordinator	Dr. Gur Yaari	Bar Ilan University	30.06.2019	GY
WP Leaders				

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Table of Contents

Executive Summary 5
1 Introduction 6
1.1 Purpose and scope of the deliverable 6
1.2 iReceptor Plus and the GDPR 6
2 Identification of non-EU beneficiaries involved in iReceptor Plus 8
3 Analysis of data processing activities in iReceptor Plus 9
3 Relevant national laws on transfer of personal data 11
3.1 Norway 11
3.2 Israel 11
3.3 Canada 12
3.4 United States of America 13
4 Assessment of compliance with non-EU country laws on transfer of personal data and further recommendations 14
5 Conclusion 15

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Executive Summary

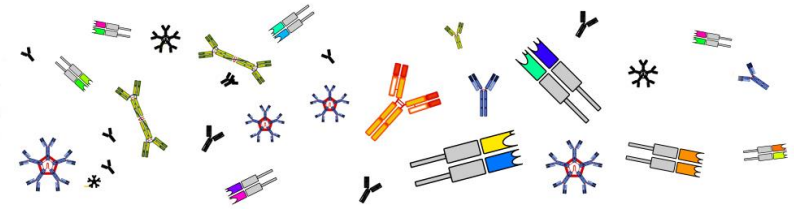
This deliverable is dedicated to confirming that any transfer of personal data from a non-EU country to the EU (or another third state) complies with the laws of the country in which the data was collected.

To this aim, the deliverable first provides a short introduction to the objectives of the iReceptor Plus project and their relationship to the subject of the deliverable. The document then proceeds by identifying the beneficiaries involved in iReceptor Plus that run a repository in a non-EU country. In the following sections of the deliverable, the data processing activities performed in the course of the project are analysed, so as to assess them in light of the relevant national data protection laws. From this assessment it appeared that the data processing activities carried out in the context of the first phase of this research project do not involve any processing of *identifiable information,* but rather anonymous data. This leads to the conclusion that the national data protection laws of the countries in which the non-EU repositories are located, do not apply to the iReceptor Plus research project as currently performed.

Nevertheless, to mitigate any risk that throughout the project identifiable information would nonetheless be processed or that anonymous information would be re-identified, the beneficiaries will (1) further investigate on national guidance with regard to the anonymisation of health information for research purposes, and (2) monitor and periodically re-assess the identifiability of the information that is processed in the initial phase and further phases of the iReceptor Plus project, with the aim of ensuring proper compliance to the highest standards of data protection.

Such re-assessment will, in our opinion, in any case be necessary as of month 24 of the project.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



1 Introduction

1.1 Purpose and scope of the deliverable

As a part of Work Package 11 (Ethics requirements), Deliverable 11.12 aims at (1) analysing the processing activities performed in the iReceptor Plus project and (2) providing confirmation that any transfer of personal data from a non-EU country to the EU (or another third state) throughout the project complies with the laws of the non-EU country in which the data was collected.

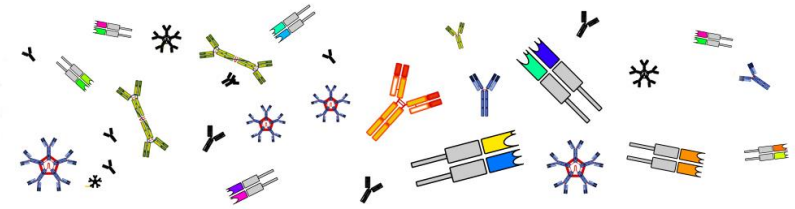
In terms of scope, it is important to emphasize that this deliverable will only deal with data processing activities that are performed in the course of the project and will not consider any data processing that will be carried out once the iReceptor Plus research project is finalised and fully operational. This limitation of scope corresponds to the description of deliverables in the Grant Agreement.

1.2 iReceptor Plus and the GDPR

The objective of iReceptor Plus as a research project is to build a common scalable platform to integrate distributed repositories of Adaptive Immune Receptor Repertoire sequencing data (AIRR-seq data) for enabling improved personalized medicine and immunotherapy for diseases with an immune component. iReceptor Plus will be designed as a network of federated repositories that facilitates data queries and advances analyses through a centralized web portal (the iReceptor Plus Scientific Gateway). In essence, this means that iReceptor Plus will be a freely and openly available software platform, which enables (1) the querying of multiple repositories at once by common metadata searches, (2) the analysis of federated data aggregated from multiple repositories and (3) the integration of these data with other types of large scale human health, biological and genomic data.

To develop this software, beneficiaries based both within and outside the European Union (EU) will bundle their expertise. Within the EU, beneficiaries from France, Germany, Spain, Portugal and Belgium are involved in the project. The non-EU beneficiaries in the project are based in Norway, Israel, Canada and the United States of America. Given that the iReceptor Plus project will result in a software platform that enables access to and analysis of both AIRR-seq data and non-AIRR-seq data (such as clinical and biological data) stored in distributed repositories, it is important to analyse what the exact processing activities during the development phase of the project are, and if these processing activities comply with the laws of the non-EU countries involved in relation to the transfer of personal data from these countries to the EU or another third state.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



The following sections of this deliverable will therefore:

- **IDENTIFY** the non-EU beneficiaries involved in iReceptor Plus (**section 2**);
- **ANALYSE** the data that will be processed by beneficiaries during the project (**section 3**);
- **SET OUT** the laws of relevant non-EU countries with regard to the processing of personal data, as well as their scope of application (**section 4**);
- **ASSESS** the applicability of and compliance with the laws of non-EU countries with regard to the processing of personal data and **RECOMMEND** on any further steps to take (**section 5**);

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

2 Identification of non-EU beneficiaries involved in iReceptor Plus

One of the key defining elements of the iReceptor Plus research project is that it is designed as a network of federated repositories that facilitates data queries and advances analysis through a centralized web portal. This distributed, federated approach is central to the iReceptor Plus concept as it allows each institution to maintain control over their data and stay compliant with local legislation. This approach however also entails that the AIRR-seq data that are processed in the development phase of the project, will be stored in the individual repositories that underlie the iReceptor Plus network.

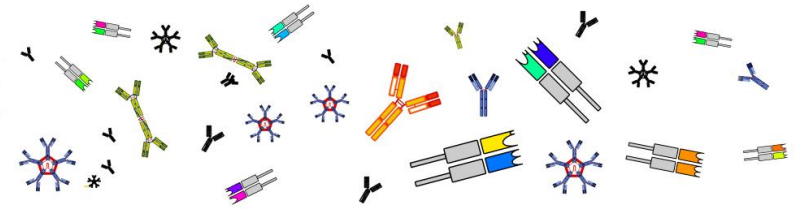
Knowing that iReceptor Plus involves beneficiaries based both within and outside the European Union, it is important to identify the beneficiaries of whom the repository will be part of the iReceptor Plus network of repositories and in which non-EU country these beneficiaries are based.

Below are listed the non-EU beneficiaries that are running a repository that will be part of the iReceptor Plus research project, as well as the location of those repositories.

Beneficiary		**Storage location**
Bar Ilan University	BIU	Israel
Simon Fraser University	SFU	Canada
University of Toronto	UTORONTO	Canada
University of Haifa	UH	Israel
University of Oslo	UiO	Norway
Medgenome Inc.	MedGenome	United States of America
10X Genomics	10X	United States of America
Clalit health services	Rabin MC	Israel
University of Texas System	UTSW	United States
Oslo University Hospital	OUS	Norway

Knowing that iReceptor Plus is designed as a network of repositories, it is likely that in the course of the iReceptor Plus, data will be transferred from the above non-EU repositories to the EU or from a certain non-EU repository to another.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



3 Analysis of data processing activities in iReceptor Plus

It is useful to recall that iReceptor Plus will be developed as a freely and openly available software platform, which operates on three levels. First of all (1), the iReceptor Plus platform will give access to publicly available AIRR-seq data in the AIRR Data Commons[1] ('public AIRR-seq data'[2]). On a second level (2), the iReceptor Plus platform will provide an intermediate level of sharing of non-publicly available AIRR-seq data that can only be accessed if common consent structures or reciprocal data transfer agreements (DTA) are put in place ('controlled AIRR-seq data'[3]). Finally (3), iReceptor Plus will provide the ability to integrate public AIRR-seq data with non-public non-AIRR seq data (such as information extracted from individual electronic health records), without exposing the non-public non-AIRR seq data.

This way, the iReceptor Plus platform will not only provide access to large amounts of public data but will also facilitate the comparison and integration of non-public data with public data.

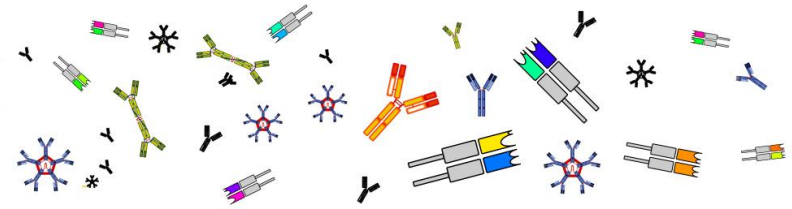
- The **public AIRR-seq data** (including metadata) are data that were originally used to perform scientific research, and which were afterwards deposited in a repository as a pre-requisite to the publication of the research paper the data support.

  In the context of iReceptor Plus, the individual repositories include both AIRR-seq data that come directly from research studies performed by the scientific entity curating the repository, as well as AIRR-seq data that were taken from other large repositories such as SRA[4] or ENA[5].

  Every deposit of AIRR-seq data into a repository is always performed under the

[1] The AIRR Data Commons is the network of all distributed AIRR-seq data repositories.
[2] Public AIRR-seq data are AIRR-sequences, including metadata that are made available to all through a public repository.
[3] Controlled AIRR-seq data are AIRR-sequences, including metadata that are not made available to all through a public repository, but can only be accessed if access is authorized by the data owner through a reciprocal data transfer agreement or common consent structure concluded with the person or entity envisaging to access the data. In essence, the data owner 'controls' the access to the data.
[4] The Sequence Read Archive (SRA) is a bioinformatics database that provides a public repository for DNA sequencing data, especially the 'short reads' generated by high-throughput sequencing, which are typically less than 1000 base pairs in length. he archive is part of the International Nucleotide Sequence Database Collaboration (INSDC), and run as a collaboration between the National Centre for Biotechnology Information (NCBI), the European Bioinformatics Institute (EBI), and the DNA Data Bank of Japan (DDBJ).
[5] The European Nucleotide Archive (ENA) is a repository providing free and unrestricted access to annotated DNA and RNA sequences.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



supervision of and with the authorisation of the ethics committee responsible for the research data concerned. Each ethics committee is responsible for following ethical guidelines and complying with the GDPR and national data protection laws. Given that most national laws foresee an obligation of anonymisation before research data can be published, the AIRR-seq data deposited in the public repositories should be considered anonymous.

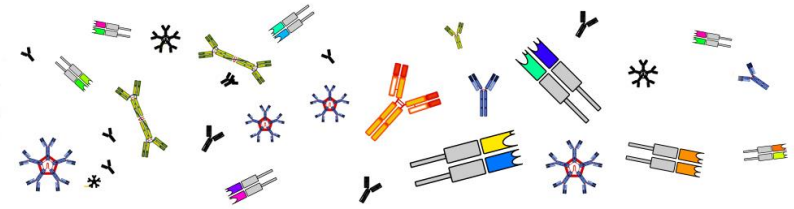
- The **controlled AIRR-seq data** (including metadata) are data that do not originate from publicly available repositories but come from research entities that have themselves performed the research concerned. Since those research entities were in contact with the people from whom the samples underlying the data were taken, they might be able to retrace the people to whom the data relate, unless anonymisation techniques were already applied to the data. This leads to the conclusion that some of the controlled AIRR-seq data may be non-anonymous data.

- The **non-public non-AIRR seq data** that can be integrated with public AIRR-seq data via the iReceptor Plus platform might be anonymous or non-anonymous data depending on the question if the entity curating these data has applied anonymisation techniques to the data. However, when integrating these non-public non-AIRR seq data with public AIRR-seq data, the non-public non-AIRR seq data will always remain with entity requesting the integration with the public AIRR-seq data and will never be shared with other parties. Consequently, this integration capability of the iReceptor Plus platform will never lead to a transfer of personal data from an EEA-country to a non-EEA country and does not raise concerns in the context of this deliverable.

During the **initial development phase** of the iReceptor Plus project, **only public AIRR-seq data curated in the distributed repositories will be processed.** Given that access to controlled AIRR-seq data requires common consent structures or reciprocal transfer agreements, these data will not be part of the initial development phase (M1-M24) of iReceptor Plus.

After the initial development phase, when the security and access control measures required to enable researchers to manage data according to their data transfer agreements and consent structures, **controlled AIRR-seq data will also be processed**. Knowing that these data may in some cases be non-anonymous data and thus personal data, as this later stage of the project, the requirements for transfers of personal data (as set out above) will have to be complied with. This will be confirmed in future deliverables that deal with re-assessing the need for GDPR-compliance and data management compliance (such as D3.4, M24).

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



3 Relevant national laws on transfer of personal data

3.1 Norway

Although Norway is not part of the European Union, Norway is a member of the European Economic Area (EEA)[6], making it part of the EU's single market. That is why, by decision of 6 July 2018, the EEA Joint Committee announced the incorporation of the General Data Protection Regulation (GDPR) into the EEA Agreement, making the GDPR directly applicable to the three EAA states: Norway, Iceland and Liechtenstein.

Consequently, for the purpose of this deliverable, Norway should not be seen as a non-EU country from which data can be transferred to an EU-country but should rather be seen as a country that is part of the EU single market and as such governed by the GDPR.

For these reasons, any transfer of personal data from the Norwegian beneficiaries (University of Oslo and Oslo University Hospital) should be assessed in light of the GDPR and should thus be discussed in the context of Deliverable 11.2, which deals with transfers of personal data from the EU to a non-EU country or international organisation.

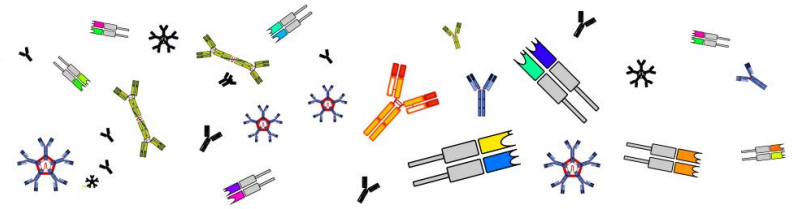
3.2 Israel

In Israel, data protection law is primarily laid down in the 'Protection of Privacy Law, 5741-1981' and regulations[7] promulgated pursuant thereto, in particular the 'Privacy Protection (Transfer of Data to Databases Abroad) Regulations, 5761-2001', which deals with the transfer of data to databases abroad.

The Israeli 'Protection of Privacy Law, 5741-1981' protects privacy in databases. For the purpose of this law, the 'information' that is protected are 'data on the personality, personal status, intimate affairs, state of health, economic position, vocational qualifications, opinions and beliefs of a person'. Moreover 'sensitive information' means (1) data on the personality, intimate affairs, state of health, economic position, opinions and beliefs of a person and (2) information that the Minister of Justice determined by order, with the approval of the Constitution, Law, Justice Committee of the Knesset is sensitive information.

[6] The EEA was established in 1992 via the EEA agreement, an international agreement enabling the extension of the EU's single market to certain non-EU member parties, namely Norway, Iceland and Liechtenstein.
[7] Note that the term 'regulations' in this sentence does not have the same meaning as it does in EU law.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



Although Israeli data protection law does not explicitly address anonymisation, de-identification or pseudonymisation, the definition of 'information' suggests that data stripped from all identifying information are outside the scope of the Israeli law. In this context, the Israeli data protection regulator (the 'Privacy Protection Authority') considers data to be identifiable where the relevant entity or its business collaborator has the ability to derive data subjects' identities using other data sets. Moreover, a public committee has been charged with establishing standards for anonymising medical information.

3.3 Canada

In Canada, there are several federal, provincial and territorial privacy statutes that govern the protection of personal information in the private, public and health sector.

Given that the Canadian repositories in iReceptor Plus are ran by the beneficiaries Simon Fraser University (SFU) and University of Toronto (UTORONTO), this deliverable will mainly refer to the Canadian data protection law applicable to the public sector.

Under Canadian law, universities are in principle covered by provincial laws. This means that, in the context of this deliverable, the essential laws are those of British Columbia and Ontario, the provinces in which the repositories of respectively SFU and UTORONTO are located.

In British Columbia, the 'Freedom of Information and Protection of Privacy Act' is the general public sector privacy law. Besides that, there is a specific privacy law relating to health records, namely the E-Health (or Personal Health Information Access and Protection of Privacy Act). Moreover, the Information and Privacy Commissioner for British Columbia has issued a guidance document for access to data for health research.[8] In Ontario, there also is a 'Freedom of Information and Protection of Privacy Act' for the public sector and a 'Personal Health Information Protection Act'.

Although not identical, the above acts all have similar definitions of 'personal information' and 'personal health information'. From these definitions, it appears that an essential element for information to be considered personal, is that the information is identifiable. Consequently, it is clear that the provincial data protection laws of British Columbia and Ontario do not apply to anonymous information.

[8] Available at: https://www.oipc.bc.ca/guidance-documents/2115.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)

Furthermore, Canadian data protection laws do not seem to distinguish between domestic and international transfers[9] of data.

3.4 United States of America

The US has several sector-specific and medium-specific privacy and data protection laws both at the national, state and territorial level.

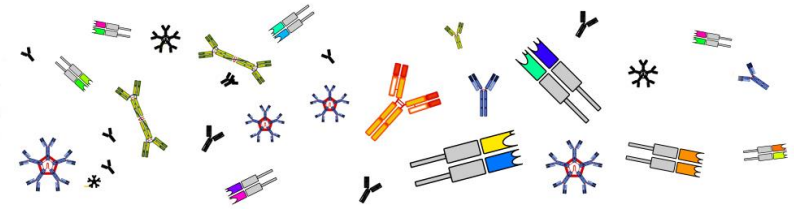
The beneficiaries of iReceptor Plus based in the United States of America (US) are Medgenome Inc. (MEDGENOME), 10X Genomics Inc (10X) and University of Texas System (UTSW). The repositories of the beneficiaries MEDGENOME and 10X are located in the state of California, the repository of UTSW is located in the state of Texas.

The most prominent data protection law in California is the 'California Consumer Privacy Act' (CCPA). In Texas, two bills have been filed to enact similar data protection laws in Texas, namely the 'Texas Consumer Privacy Act' (House Bill 4518) and the 'Texas Privacy Protection Act' (House Bill 4390).

Again, these (proposed) laws contain differing, but highly similar definitions of personal information, which each refer to the identifiability of the information concerned. As such, it is clear that also in the United States of America, data protection laws only apply to information about an identifiable person.

[9] The term 'transfer' does not have the same meaning under Canadian law as under EU law.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



4 Assessment of compliance with non-EU country laws on transfer of personal data and further recommendations

From the above sections, it appeared that in the initial phase of the iReceptor Plus project, only anonymous data will be processed. Furthermore, it appeared that the laws that govern the non-EU repositories involved in iReceptor Plus only apply to processing activities and transfer of data that concern *identifiable information.*

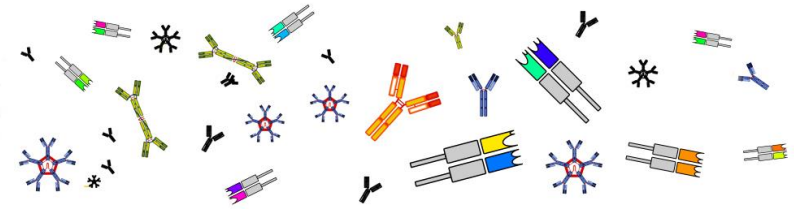
Consequently, it can be concluded that the processing activities carried out in the first phase of the iReceptor Plus project are not subject to the national data protection laws of the countries in which the non-EU repositories are located and that, to our current understanding, there are currently no requirements that should be complied with for the AIRR-seq data stored in the non-EU repositories to be transferred to the EU or another third country.

Nevertheless, the consortium is aware that the assessment of the identifiability of the genetic data concerned is a dynamic exercise which cannot be decided upon once and for all. The beneficiaries therefore commit to periodically re-assess the risk of re-identification at every stage of processing and to further investigate any national guidance on anonymisation of health information for research purposes.

Should it at a further stage (e.g. in M24) become clear that the processing activities carried out in the research project do involve identifiable information in the sense of the relevant national data protection laws, then the beneficiaries ascertain that the requirements of those laws will be analysed on a more thorough level, so as to guarantee strict compliance.

That is why the technically skilled beneficiaries of the project will monitor the data processing activities performed in the project and will report on any changes thereof to the legal and ethics partners involved in iReceptor Plus. This will allow iReceptor Plus to re-assess its current compliance whenever necessary, so as to remain compliant throughout the entire project. Such re-assessment will, in our opinion, in any case prove essential as of month 24 of the project.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)



5 Conclusion

The analysis performed in this deliverable has shown that in the initial phase (M1-M24) of the iReceptor Plus project only anonymised AIRR-seq data will be processed. Knowing that anonymous data are not governed by the national data protection laws of the non-EU repositories involved in iReceptor Plus, those national laws are not applicable to the iReceptor Plus processing activities as currently performed. Consequently, compliance with these national data protection laws are not an issue in this first phase of the project. Nevertheless, to mitigate any risk that throughout the project personal information would nonetheless be processed or that anonymous information would be re-identified, the beneficiaries will (1) further investigate on national guidance with regard to the anonymisation of health information for research purposes, and (2) monitor and periodically re-assess the identifiability of the information that is processed during the iReceptor Plus project, with the aim of ensuring proper compliance to the highest standards of data protection.

This project is funded by the European Union's H2020 Research and Innovation Programme under Grant Agreement No. 825821 and Canadian Institutes of Health Research (CIHR)