

CADRE DE GESTION

Échantillon de référence régional québécois (ERRQ)

Responsable de la banque :

Simon Girard

Directeur Projet BALSAC

Université du Québec à Chicoutimi
555, boulevard de l'Université
Chicoutimi (Québec) G7H 2B1

Département des sciences fondamentales
Bureau P4-2130
418-545-5011 poste 2595
1-800-463-9880 poste 2595 (sans frais)
simon2_girard@uqac.ca

TABLE DES MATIÈRES

1	DESCRIPTION DE LA BANQUE	3
1.1	CONTEXTE	3
1.2	NATURE DES DONNÉES	3
1.2.1	<i>Critères d’admissibilité</i>	3
1.2.2	<i>Recrutement et collecte des données</i>	4
1.2.3	<i>Consentement à participer à la banque</i>	4
1.3	VOCATION	4
2	GESTION ADMINISTRATIVE	4
2.1	PERSONNES RESPONSABLES	4
2.2	INSTALLATIONS ET ÉQUIPEMENTS	5
2.3	CONSERVATION ET PÉRENNITÉ DES DONNÉES	5
3	GESTION DES DONNÉES	5
3.1	ACCÈS À LA BANQUE	5
3.2	GESTION DES DEMANDES	5
3.3	MESURES DE CONFIDENTIALITÉ DES DONNÉES	6
3.4	TRAITEMENT DES DONNÉES	6
3.5	MODALITÉS DE RETRAIT D’UN PARTICIPANT	6
3.6	DÉCOUVERTES FORTUITES	7
4	DIFFUSION DES RÉSULTATS ET PUBLICATIONS	7
5	SUIVI ÉTHIQUE	7
6	RÉVISION	7
	ANNEXE 1 : FORMULAIRE DE DEMANDE D’ACCÈS AUX DONNÉES DE L’ÉCHANTILLON DE RÉFÉRENCE RÉGIONAL QUÉBÉCOIS (BANQUE ERRQ)	8
	ANNEXE 2 : FORMULAIRE D’ENGAGEMENT CONCERNANT L’ACCÈS AUX DONNÉES DE L’ÉCHANTILLON DE RÉFÉRENCE RÉGIONAL QUÉBÉCOIS (BANQUE ERRQ)	12
	ANNEXE 3 : COMPTE DES DONNÉES DE LA BANQUE ERRQ	13

1 Description de la banque

1.1 Contexte

L'Échantillon de référence régional québécois (ERRQ) est une banque de données constituée dans le cadre du projet de recherche « Groupe de référence en épidémiologie génétique et en génétique des populations humaines au Québec », initié par le professeur Damian Labuda, du Centre de recherche du Centre hospitalier universitaire Sainte-Justine (CHU Ste-Justine), en collaboration avec la professeure Hélène Vézina, de l'Université du Québec à Chicoutimi (UQAC). L'objectif visé par ce projet était la création d'une banque d'échantillons génomiques couplés à des données généalogiques à l'intention de la communauté scientifique qui œuvre dans la réalisation de diverses études sur la structure et l'histoire génétique de la population québécoise. Les données de la banque ERRQ peuvent aussi servir de cohorte de référence pour des études sur la susceptibilité génétique individuelle à diverses maladies ou sur la réceptivité des individus à certains traitements médicamenteux.

Cette banque était initialement conservée dans le laboratoire du professeur Damian Labuda, au CHU Ste-Justine. En 2022, le Projet BALSAC a pris le relai de la gestion des données. Ce cadre de gestion a pour objectif d'informer les utilisateurs de la banque et le personnel du Projet BALSAC des principes et des règles balisant la gestion et l'accès aux données génétiques et généalogiques qui sont associées à la banque ERRQ.

Enfin, pour une meilleure compréhension du présent document, il est à noter que le Projet BALSAC est une infrastructure de recherche, hébergée à l'UQAC, visant, notamment, à construire et à exploiter le fichier BALSAC. Ce dernier est constitué principalement des données provenant des actes de l'état civil couvrant la population du Québec depuis le début du 17^e siècle jusqu'à la période actuelle ce qui permet la reconstitution de généalogies.

1.2 Nature des données

À la suite du recrutement des participants de la cohorte, deux types de données ont été générées et composent la nature des données accessibles à travers la banque ERRQ, soit des données génétiques de géotypage ainsi que des données généalogiques. Aucune donnée phénotypique n'a été recueillie et cette cohorte est exempte de biais médical.

Le nombre de données inclus dans chacun des volets de la cohorte est décrit en [Annexe 3 : Compte des données de la cohorte ERRQ](#).

1.2.1 Critères d'admissibilité

Le recrutement des participants s'est fait à travers plusieurs régions du Québec. Les critères d'inclusion des individus recrutés étaient d'être âgé de 18 ans et plus et d'avoir des parents ou des grands-parents nés ou mariés au Québec. Selon la région de recrutement, certains de ces critères ont été resserrés afin de prendre en compte l'histoire du peuplement d'origine européenne au Québec et d'assurer une représentation juste de la population régionale.

1.2.2 Recrutement et collecte des données

Le recrutement et la collecte ont débuté en 2003 et se sont terminés en 2014. Les données de la banque ont fini d'être collectées et aucun participant ne sera recontacté.

1.2.3 Consentement à participer à la banque

Les participants ont signé un formulaire de consentement permettant l'utilisation secondaire des données. Le présent Cadre de gestion respecte tous les termes inscrits dans les formulaires.

1.3 Vocation

L'utilisation et l'exploitation de la banque visent la production et la valorisation de la connaissance scientifique. L'accès aux données de la banque s'inscrit dans cette vocation en étant réservé aux individus œuvrant dans un contexte de recherche scientifique, réalisée dans un cadre éthique et désintéressé, sans aucune conséquence d'ingérence dans la vie privée d'autrui et sans aucun intérêt de mise à profit commerciale, personnelle ou abusive de l'information. En aucun cas, la banque ERRQ ne peut être utilisée avec une visée de mise à profit commerciale de l'information et des données qu'elle contient. Parallèlement, aucun participant ne peut tirer profit de l'intégration de ses informations génétiques ou généalogiques au sein de la banque ERRQ.

Si un changement vocationnel est envisagé, un amendement au présent Cadre de gestion sera produit et déposé pour approbation au Comité d'éthique de la recherche de l'UQAC (CER-UQAC). La demande devra expliquer de façon détaillée les changements projetés et les raisons qui les motivent.

2 Gestion administrative

2.1 Personnes responsables

La banque ERRQ est sous la responsabilité du Projet BALSAC, par l'entremise de sa direction. La direction du Projet BALSAC a comme mandat de voir à la conservation, au développement, à la valorisation et à l'exploitation de l'ensemble des données qui sont sous sa responsabilité, y compris de la banque ERRQ. Le Projet BALSAC est sous la supervision de l'UQAC, telle que mandatée par l'*Entente interuniversitaire relative au Projet BALSAC*¹. Les données génétiques et généalogiques de la banque ERRQ constituent une ressource institutionnelle. À ce titre, les chercheurs ne peuvent revendiquer d'autres droits que ceux consentis par les participants de la banque ERRQ et les règlements de l'UQAC.

La direction du Projet BALSAC est supportée dans son mandat par le Comité directeur du Projet BALSAC et du Comité d'accès dont la composition relève du *Cadre de gestion des données du fichier de population BALSAC*².

¹ *Clauses 5.1.3 et 5.1.4, I-C-242 – Entente cadre universités partenaires BALSAC 2011*

² *Point 3.2 Structure de gestion, Cadre de gestion – Fichier BALSAC – version approuvée le : 25 mars 2024. Numéro de référence : 2024-1644*

2.2 Installations et équipements

Les données de la banque ERRQ sont conservées sur les serveurs internes sécurisés du Projet BALSAC, à l'UQAC. La sécurité des données est garantie par plusieurs mesures physiques, informatiques et structurelles, définie par l'ensemble des mesures mises en place par l'UQAC.

2.3 Conservation et pérennité des données

Les données génétiques et généalogiques de la banque ERRQ sont conservées au Projet BALSAC. Advenant le cas où le Projet BALSAC ne peut plus assumer la responsabilité des données, des démarches seront effectuées afin qu'une autre instance prenne le relai. Si elles ne peuvent être transférées, les données génétiques ainsi que les liens qui les relient aux données généalogiques du fichier BALSAC seront détruits de manière sécuritaire selon les modalités en vigueur à l'UQAC.

3 Gestion des données

3.1 Accès à la banque

Les demandes d'accès aux données de la banque ERRQ sont réservées aux chercheurs, œuvrant dans un contexte scientifique, affiliés à une institution de recherche reconnue. Une demande d'accès prend toujours la forme d'un prêt de données, transmis sous la forme d'un jeu de données entièrement anonymisé au chercheur principal de la demande d'accès.

Les chercheurs qui font une demande d'accès aux données ERRQ doivent passer par le Service aux chercheurs du Projet BALSAC en présentant le formulaire présenté à l'[Annexe 1 : Formulaire de demande d'accès aux données de l'Échantillon de référence régional québécois](#). Cette demande d'accès peut être faite conjointement à une demande d'accès au fichier BALSAC et s'inscrire dans le formulaire de demande d'accès prévu à cet effet dans le *Cadre de gestion des données du fichier de population BALSAC*³.

3.2 Gestion des demandes

Le Service aux chercheurs du Projet BALSAC est responsable de la gestion et du traitement des demandes d'accès aux données de la banque ERRQ. Les demandes reçues seront soumises au Comité d'accès aux données du Projet BALSAC qui évaluera la recevabilité de la demande selon les critères présentés sur le [site internet du Projet BALSAC](#). En tout temps, le Comité d'accès se réserve le droit de questionner ou refuser les demandes d'accès qui lui sont présentées.

Le Comité d'accès suit en tout temps les mêmes procédures d'évaluation et de gestion des demandes d'accès pour la banque ERRQ, que celles dictées au point 5 du *Cadre de gestion des données du fichier de population BALSAC*⁴. Les chercheurs sont tenus d'utiliser les données de la banque ERRQ exclusivement pour le projet qui a été approuvé par le Comité d'accès du Projet

³ *Annexe 2 : Formulaire de demande d'accès au Fichier BALSAC, Cadre de gestion – Fichier BALSAC – version approuvée le : 25 mars 2024. Numéro de référence : 2024-1644*

⁴ *Point 5 Procédures d'accès à des fins de recherche, Cadre de gestion – Fichier BALSAC – version approuvée le : 25 mars 2024. Numéro de référence : 2024-1644*

BALSAC. Ils doivent également signer un formulaire d'engagement (voir [Annexe 2: Formulaire d'engagement concernant les données extraites de l'Échantillon de référence régional québécois \(banque ERRQ\)](#)) et s'engager à détruire les données faisant l'objet de cette entente à la fin du projet pour lequel ils se sont vus autoriser un accès ou à soumettre une demande de prolongation d'accès au Service aux chercheurs du Projet BALSAC.

Le Service aux chercheurs du Projet BALSAC tient un registre des demandes et des projets en cours qui utilisent les données de la banque ERRQ.

3.3 Mesures de confidentialité des données

Les données génétiques et généalogiques associées à la banque ERRQ sont considérées comme des données confidentielles. Conformément, tous les membres du personnel du Projet BALSAC ont signé des formulaires d'engagements à la confidentialité qui assure le respect de la portée confidentielle des données auxquels ils ont accès dans le cadre de leur fonction.

Similairement, le Comité d'accès aux données évalue les demandes d'accès en s'assurant du respect de la confidentialité des données et de la non-traçabilité des individus concernés. Tous les jeux de données sont anonymisés, et lorsque possible, présentés dans un format agrégé.

Les données génétiques sont conservées sur les serveurs sécurisés internes du Projet BALSAC sous forme de fichier PLINK binaire avec le code de l'individu comme identifiant. Les données généalogiques sont déjà partie intégrante du fichier BALSAC et le code de l'individu est utilisé comme identifiant pour faire le lien avec les données génétiques. Les données nominatives contenues dans le fichier BALSAC ne sont pas associées aux données génétiques ou généalogiques de la banque ERRQ et ne sont pas partagées avec les chercheurs.

3.4 Traitement des données

Les données génétiques ont été génotypées avec deux puces de génotypage différentes, soit Illumina Omni Express (~740,000 SNPs) et Illumina Omni 2.5 (~2.5M SNPs). Les données généalogiques ont été intégrées au fichier BALSAC, avec lesquelles ont été jumelées afin de constituer des ascendances généalogiques complètes. Les chercheurs qui souhaitent en connaître davantage sur les données peuvent communiquer avec le Service aux chercheurs du Projet BALSAC.

3.5 Modalités de retrait d'un participant

Les participants peuvent retirer leurs données généalogiques et génétiques de la banque ERRQ à tout moment en écrivant un courriel au Projet BALSAC, tel qui avait été mentionné dans les formulaires de consentement des participants du projet de recherche initial intitulé « Groupe de référence en épidémiologie génétique et en génétique des populations humaines au Québec ». Cependant, tel que cela avait été précisé, en cas de demande de retrait, les données introduites

dans le fichier BALSAC ayant permis la reconstruction généalogique demeureront dans la base de données du fichier BALSAC, mais celles-ci ne seront plus associées à la banque ERRQ.

3.6 Découvertes fortuites

Les découvertes fortuites qui découlent des projets de recherche conduits sur les données de la banque ne seront pas communiquées puisque le Projet BALSAC ne possède pas les coordonnées des participants et aucun moyen de les contacter.

4 Diffusion des résultats et publications

Les équipes de recherche ayant eu accès aux données de la banque ERRQ doivent le mentionner dans toutes publications ou présentations relevant de leur utilisation. Une formulation est présentée sur le [site internet du Projet BALSAC](#). Les résultats de recherche découlant de l'utilisation des données de la banque ERRQ, à l'exclusion des données elles-mêmes, demeurent la propriété du chercheur.

Également, les chercheurs doivent communiquer au Projet BALSAC, la liste des publications et communications appuyées sur des données de la banque ERRQ. Lorsque cela est possible, les chercheurs sont encouragés à remettre au Projet BALSAC une copie informatisée de leur article publié.

5 Suivi éthique

Afin d'assurer un suivi de l'approbation du Cadre de gestion, le Projet BALSAC s'engage à informer le CER-UQAC par écrit et dans les meilleurs délais des situations suivantes si elles se présentent :

- 1) De toute modification au Cadre de gestion, et ce, préalablement à l'application de ce changement.
- 2) De tout événement imprévu et sérieux qui surviendrait dans le déroulement des activités liées au présent Cadre de gestion et qui affecterait le respect des engagements envers les participants, ou dans le cas d'un bris de confidentialité.
- 3) De tout changement de vocation de la banque, de sa pérennité ou de ses responsables.
- 4) De tout changement dans la liste des personnes qui ont des accès privilégiés aux données génétiques et généalogiques brutes.

Une demande de renouvellement annuel de la certification éthique pour la banque de données et son Cadre de gestion doit être présentée au CER-UQAC. Le Projet BALSAC conserve une liste des demandes des chercheurs incluant le titre du projet proposé et le numéro attribué par le comité d'éthique de l'établissement auquel est rattaché le chercheur. Une liste des publications découlant de la banque est aussi conservée.

6 Révision

Une révision approfondie du Cadre de gestion est entreprise tous les cinq (5) ans. À défaut d'une révision dans le délai imparti de cinq (5) ans, la dernière version approuvée par le CER-UQAC demeure en vigueur.

ANNEXE 1 : Formulaire de demande d'accès aux données de l'Échantillon de référence régional québécois (banque ERRQ)

1. RENSEIGNEMENTS GÉNÉRAUX				
Titre du projet				
Échéancier d'accès aux données				
Date de début				
Date de fin				
Chercheur.euse principal.e du projet				
Statut : <input type="checkbox"/> PROFESSEUR.E <input type="checkbox"/> STAGIAIRE POSTDOCTORAL <input type="checkbox"/> ÉTUDIANT.E AUX CYCLES SUPÉRIEURS				
Nom, Prénom				
Établissement				
Téléphone				
Courriel				
Chercheur.euse responsable (Pour demande provenant de stagiaire postdoctoral et étudiant.e.s)				
Nom, Prénom				
Établissement				
Téléphone				
Courriel				
Membres de l'équipe de recherche (listez toutes les personnes qui auront accès aux données, y compris les étudiants)				
Nom, Prénom	Rôle	Courriel	Établissement	Autorisation
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON
				<input type="checkbox"/> OUI <input type="checkbox"/> NON

2. DESCRIPTION DU PROJET DE RECHERCHE

Résumé du projet

Inclure les hypothèses et objectifs de recherche, les données et sujets humains visés par la recherche et la méthodologie employée.

3. FINANCEMENT DU PROJET DE RECHERCHE

Confirmé	
Demandé	

4. DONNÉES DEMANDÉES EN PROVENANCE DE LA BANQUE ERRQ

Fichier généalogique de base pour tous les individus de la banque ERRQ

Cela inclut les variables suivantes : *ID individu; Sexe de l'individu; ID du père de l'individu; ID de la mère de l'individu; Lieu de mariage des parents; Date de mariage des parents; Statut migratoire; Lieu d'origine*

Données génétiques pour tous les individus de la banque ERRQ

Les données sont dans un format PLINK.

Données spécifiques à certaine région seulement

Données génétiques :

- Abitibi
- Beauce
- Côte-Nord
- Gaspésie complet
 - Gaspésie – Origine acadienne
 - Gaspésie – Origine anglo-normande
 - Gaspésie – Origine canadienne-française
 - Gaspésie – Origine loyaliste
- Montréal et les environs
- Outaouais
- Québec et les environs

Données généalogiques :

- Abitibi
- Beauce
- Côte-Nord
- Gaspésie complet
 - Gaspésie – Origine acadienne
 - Gaspésie – Origine anglo-normande
 - Gaspésie – Origine canadienne-française
 - Gaspésie – Origine loyaliste
- Montréal et les environs
- Outaouais
- Québec et les environs

Ajout de variable(s) supplémentaire(s)

Inscrire une liste des variables demandées. Une brève justification doit accompagner la demande pour chaque variable supplémentaire inscrite dans la liste. À mettre en annexe de la demande.

(La liste des variables disponibles dans le fichier BALSAC se retrouve à l'adresse Internet suivante : <http://www.uqac.ca/balsac/>)

5. UTILISATION DE DONNÉES EXTERNES AUX DONNÉES DE L'ERRQ

Dans le cadre de votre projet de recherche, comptez-vous utiliser des données venant d'une cohorte ou d'une autre base de données que celles de la banque ERRQ ? OUI NON

Précisez la cohorte ou la base de données :

- CARTaGENE
- Fichier BALSAC
- Autre :

Précisez la nature des données auxquels vous avez accès :

Avez-vous déjà obtenu l'approbation de travailler avec cette cohorte ou ensemble de données ?

- OUI NON *(Si oui, inclure la confirmation en annexe de votre demande.)*

6. DOCUMENT À JOINDRE À LA DEMANDE

Certification éthique

Pour avoir accès à ces données, votre demande d'accès doit être accompagnée d'une approbation de votre CER institutionnel.

Justificatif des variables demandées

Une justification est demandée pour les variables supplémentaires qui ne sont pas incluses dans un fichier généalogique de base.

Confirmation d'accès à une cohorte ou à un ensemble de données autre que les données de l'ERRQ

Si votre projet de recherche utilise des données venant d'une cohorte ou d'une base de données autre que celles de la banque ERRQ, inclure votre autorisation d'accès aux données à cette demande.

Signature du/de la chercheur.euse principal.e

Date

Signature du/de la chercheur.euse responsable
(pour demande provenant de stagiaire postdoctoral et étudiant.e)

Date

ANNEXE 2 : Formulaire d'engagement concernant l'accès aux données de l'Échantillon de référence régional québécois (banque ERRQ)

Chercheur.euse principal.e : _____

Titre du projet : _____

À titre de chercheur.euse principal.e, je déclare avoir lu le cadre de gestion de l'Échantillon de référence régional québécois (banque ERRQ) et à m'y conformer. Je m'engage également à :

- Tenir une liste des personnes qui utilisent les données et leur fournir les informations requises afin qu'elles respectent le présent engagement;
- Ne pas réclamer de propriété intellectuelle sur des données extraites de l'Échantillon de référence régional québécois (banque ERRQ).
- Utiliser les données exclusivement pour la réalisation du projet que j'ai soumis et qui a été approuvé par le comité d'accès aux données du Projet BALSAC;
- Ne pas redistribuer les données qui m'ont été transmises dans le cadre de cette entente;
- Reconnaître l'utilisation des données de l'Échantillon de référence régional québécois (banque ERRQ) dans toute activité de diffusion de résultats résultants d'analyses faites à partir des données de ce fichier, selon le modèle de citation approprié présenté sur notre site internet à <https://balsac.uqac.ca/citer-balsac/>).
- Informer le Projet BALSAC de toute production issue des données reçues (articles scientifiques, chapitres de livres, rapports de recherche, affiches, communications, écrits de vulgarisation scientifique et de transfert de connaissances, mémoires ou thèses).
- Détruire les données faisant l'objet de cette entente au plus tard cinq ans après la signature du présent engagement ou soumettre une demande de prolongation.

En foi de quoi, j'ai signé à _____, ce _____ 20_____.

Signature du.de la chercheur.euse principal.e : _____

Signature du.de la chercheur.euse responsable : _____
(pour demande provenant de stagiaire postdoctoral et étudiant.e)

ANNEXE 3 : Compte des données de la banque ERRQ

RÉGIONS	DONNÉES GÉNÉTIQUES	DONNÉES GÉNÉALOGIQUES
Abitibi	24	15
Beauce	46	67
Côte-Nord	71	72
Gaspésie (totale)	395	393
Gaspésie origine acadienne	102	102
Gaspésie origine anglo-normande	93	91
Gaspésie origine canadienne-française	103	106
Gaspésie origine loyaliste	94	92
Montréal et les environs	217	158
Outaouais	17	14
Québec et les environs	53	54
TOTAL	815	818