Plan Overview

A Data Management Plan created using DMP Tool

DMP ID: https://doi.org/10.48321/D1602K

Title: Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens

Creator: Giovana Torrezan - **ORCID:** 0000-0002-8659-5329

Affiliation: AC Camargo Hospital (accamargo.org.br)

Funder: São Paulo Research Foundation (fapesp.br)

Funding opportunity number: 2018/06269-5

Grant: https://bv.fapesp.br/pt/auxilios/102253/utilizacao-de-abordagens-genomicas-combinadas-para-investigar-as-causas-geneticas-de-tumores-raros-e/

Template: Digital Curation Centre (português)

Project abstract:

Os avanços recentes na genômica permitiram o reconhecimento de mais de uma centena de novos genes de predisposição ao câncer. Embora o estudo molecular destes genes esteja atualmente bem estabelecido para os tumores hereditários mais comuns, há uma série de tumores raros (incidência anual <6 a cada 100.000 indivíduos) que estão altamente associados a síndromes de câncer hereditário conhecidas, mas são frequentemente negligenciados devido à sua baixa frequência no espectro de tumores dessas síndromes. A identificação das causas genéticas dos tumores raros pediátricos e adultos é importante para a definição de melhores protocolos para rastreamento do paciente, acompanhamento clínico e aconselhamento genético familiar. Além disso, o conhecimento das bases moleculares da tumorigênese nesses tumores contribui para a compreensão de diferentes vias e processos celulares envolvidos na transformação celular e pode tornar-se relevante para a definição de novas terapias. Assim, o presente estudo tem como primeiro objetivo definir a frequência de variantes germinativas patogênicas em genes conhecidos de predisposição ao câncer em crianças, adolescentes e adultos jovens (<40 anos) com tumores raros. Para isso, avaliaremos as regiões codificantes de 113 genes conhecidos de penetrância alta a

moderada para tumores hereditários, utilizando sequenciamento de nova geração, em 200 pacientes. Ainda, selecionaremos um grupo de pacientes sem alterações germinativas provavelmente patogênicas no qual utilizaremos abordagens genômicas combinadas (sequenciamento de exoma germinativo e tumoral e análise de transcriptoma do tumor) para buscar novas causas genéticas e vias biológicas associadas ao desenvolvimento desses tumores. Com este trabalho, esperamos contribuir para a definição de estratégias de triagem genética mais eficazes e adequadas para a população brasileira com tumores raros e para a tradução de conhecimentos genéticos complexos em melhores cuidados clínicos.

Start date: 12-01-2018

End date: 12-01-2022

Last modified: 07-08-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Utilização de abordagens genômicas combinadas para investigar as causas genéticas de tumores raros em crianças, adolescentes e adultos jovens

Coleta de Dados

Que dados serão coletados ou criados?

Os dados gerados neste projeto são constituídos de:

1) dados clínicos dos participantes de pesquisa: histórico pessoal e familiar de neoplasias, dados histopatógicos dos tumores.

2) dados genéticos: dados brutos de sequenciamento NGS (arquivos FASTQ) e dados processados, como arquivos de mapeamento (BAM), de chamadas de variantes (VCFs), de variantes filtradas (planilhas excel).

Os resultados finais compilados de variantes de interesse clínico e dados clínicos serão depositados em um banco de dados do REDCap, com acesso restrito.

Como os dados serão coletados ou criados?

Os dados clínicos serão coletados a partir do prontuário eletrônico dos participantes em tabelas de excel nomeadas por data de atualização, com acesso controlado e restrito aos pesquisadores e alunos envolvidos no estudo. Os dados genéticos serão gerados por sequenciamento de nova geração de painéis genéticos, de exoma ou transcriptoma, e análises subsequentes em softwares próprios (VarSeq).

Documentação e Metadados

Que documentação e metadados acompanharão os dados?

O dados genômicos em formato BAM serão depositados no final do estudo no repositório Sequence Read Archive (SRA), com solicitação para liberação pública após a publicação dos manuscritos associados. Os metadados que acopanharão os dados genômicos serão os metadados padrão desse banco e incluem tipo de amostra, tipo de biblioteca, plataforma de sequenciamento, tipo de arquivo. Para os dados clínicos e variantes genéticas de impacto clínico depositados no banco do projeto no REDCap, será criado um dicionário de varáveis para cada campo do banco.

Ética e Conformidade Legal

Como você administrará qualquer questão ética?

O projeto foi aprovado pelo comitê de ética do A.C.Camargo Cancer Center (CAAE 84725318.4.0000.5432). Todos os participantes de pesquisa incluídos no estudo assinarão um Termo de Consentimento Livre e Esclarecido (termos do Biobanco ou do projeto), contemplando a resolução 466 do Conselho Nacional de Saúde e dando garantias aos direiros dos participantes. Toda

identificação do material biológico e resultados será feita por códigos para preservar a confidencialidade do paciente. Para os pacientes recrutados prospectivamente, os TCLEs são escaneados e esses arquivos são mantidos no servidor do A.C.Camargo Cancer Center, e as cópias físicas são armazenadas na sala do pesquisador responsável.

Como você vai gerenciar os direitos autorais e os direitos de propriedade intelectual (IP / IPR)?

O projeto conta com uma parceria internacional com o instituto Idibell (Espanha). Dados gerados em um dos objetivos do estudo serão compartilhados com pesquisadores dessa instituição, e para isso, foi firmado acordo de propriedade intelectual e de compartilhamento de dados entre o A.C.Camargo e o Idibell.

Armazenamento e Backup

Como os dados serão armazenados e terão backup durante a pesquisa?

Todos os dados gerados pela esquisa se encontram armazenados no servidor do A.C.Camargo Cancer Center, que apresenta política de acesso restrita por senha e *backup* semanal. Especificamente os dados brutos de sequenciamento (arquivos FASTQ) são armazenados mensalmente em sistema de fita e controlados pela facility de sequenciamento genômico e bioinformática e departamento de TI do ACCCCC.

Como você vai gerenciar o acesso e a segurança?

Todos os arquivos gerados serão armazenados no servidor do A.C.Camargo Cancer Center, que apresenta política de acesso restrita por senha e estrutura de pastas individuais para alunos e pesquisadores.

Seleção e Preservação

Quais dados são de valor a longo prazo e devem ser mantidos, compartilhados e / ou preservados?

Os dados genômicos e clínicos devem ser mantidos por um período de no mínimo 10 anos.

Qual é o plano de preservação a longo prazo do conjunto de dados?

O banco de dados REDCap, e os dados genômicos processados serão mantidos no servidor do ACCCC por no mínimo 10 anos. Os dados brutos do sequenciamento serão armazenados em fita por no mínimo 10 anos, sob responsabilidade da equipe de TI do ACCCC. Os dados alinhados (arquivos BAM) serão disponibilizados no banco SRA por tempo indeterminado e sem custo.

Compartilhamento de Dados

Como você vai compartilhar os dados?

O projeto conta com uma parceria internacional com o instituto Idibell (Espanha). Dados gerados em um dos objetivos do estudo serão compartilhados com pesquisadores dessa instituição, e para isso, foi firmado acordo de propriedade intelectual e de compartilhamento de dados entre o A.C.Camargo e o Idibell.

Dados brutos de sequenciamento anonimizados também serão compartilhados no repositório SRA. O identificador da submissão dos dados será disponibilizado nos artigos científicos relacionados ao estudo.

Existem restrições ao compartilhamento de dados requeridos?

Para o banco de dados completos, incluindo dados clínicos, o compartilhamento necessitará de acordo de compartilhamento de dados entre as instituições. Para os dados brutos de sequenciamento anonimizados, o acesso no repositório SRA será realizado sem restrições.

Responsabilidades e Recursos

Quem será responsável pelo gerenciamento de dados?

Pesquisadora principal: Giovana Torrezan - captura de dados, produção de metadados, qualidade de dados.

Data manager A.C.Camargo Cancer Center: Rafael Vanhoz - qualidade de dados e armazenamento.

Equipe das facilities de Sequenciamento Genômico e Bioinformática do A.C.Camargo Cancer Center: armazenamento e backup, arquivamento de dados e compartilhamento de dados.

Equipe de TI do A.C.Camargo Cancer Center: armazenamento e backup.

Quais recursos você precisará para entregar seu plano?

Os recursos humanos e físicos necessários para a geração, armazenamento e compartilhamento de dados estão disponíveis no centro de pesquisas do A.C.Camargo Cancer Center.