

【報告様式A】

日本医療研究開発機構
再生医療実現拠点ネットワークプログラム事業
事後評価報告書

I 基本情報

研究開発課題名： (日本語) AI を用いた重症心筋症に対する再生医療の Precision medicine の実践
(英語) Development of Autologous Skeletal Myoblast Cell Sheet Transplantation for Dilated Cardiomyopathy Aiming Precision Medicine

研究開発実施期間：令和2年6月1日～令和5年3月31日(予定)

研究開発代表者 氏名：(日本語) 宮川 繁
(英語) Shigeru Miyagawa

研究開発代表者 所属機関・部署・役職：
(日本語) 国立大学法人大阪大学・大学院医学系研究科・教授
(英語) Professor, Department of Cardiovascular Surgery, Osaka University Graduate School of Medicine

II 研究開発の概要

研究開発の成果およびその意義等

本研究代表者らは末期重症心不全に対して、組織工学を応用した温度感応性培養皿を用いて作成した筋芽細胞シートを心表面に移植する治療法（以下、筋芽細胞シート移植）に関する一貫した独自の研究を続けており、重症虚血性心筋症患者に対し自己骨格筋筋芽細胞シート移植術は、2015年に「再生医療等製品（ハートシート®）」として条件及び期限付き承認を得るといった、治療効果を示してきた。一方、本治療を拡張型心筋症にも応用して行った第I相臨床試験において、心機能改善効果、運動耐容能効果などを検証したところ、虚血性心筋症で認められた治療効果が必ずしもすべての拡張型心筋症患者では認められず、心機能や運動耐容能の改善が得られる症例（レスポンド）と、得られない症例（ノンレスポンド）が存在することが明らかとなった（Miyagawa S, J Am Heart Assoc. 2017）。本研究課題では、本治療法の治療メカニズムを鑑みながら、治療の受け手である患者の病態を正確に判断し、本治療法に反応しうると予測される患者にターゲットを絞ることが重要であると考え、患者の術前の生体情報・分子生物学的情報・移植される細胞シートの特性、および術後のアウトカム情報を取得し、統計解析および人工知能(AI)による機械学習によって、本治療法に反応する患者(レスポンド)を術前に正確に予測・選別するリスク因子解析モデルを作成することで、拡張型心筋症に対する Precision medicine を確立することを目的とした(図1)。特に、侵襲の高い手術によって採取する自家組織を原料とする製品に関しては、術前に治療予測を行うことは極めて重要な課題であり、自己骨格筋筋芽細胞シートのみならず、他の自家組織由来製品においても治療反応性を見極める手法は非常に意義のあることであり、本研究開発がその一助となるものと考えている。

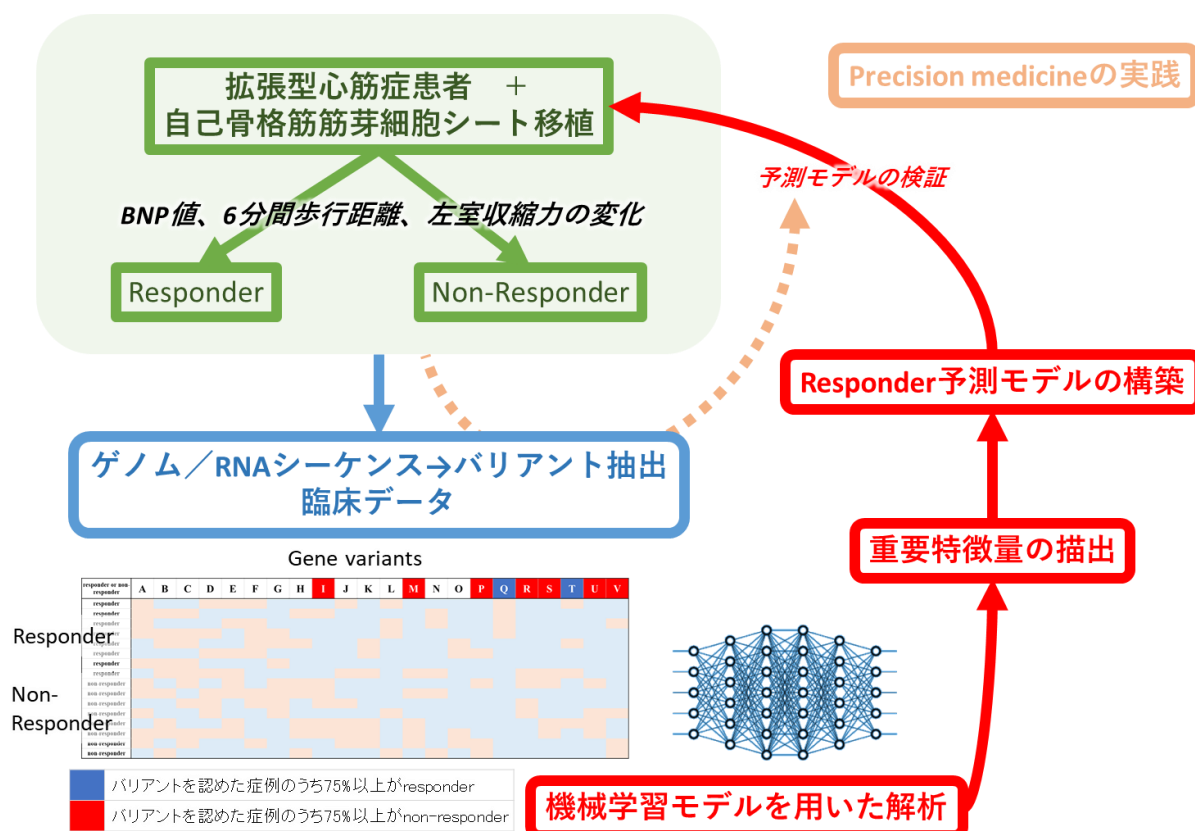


図1 Precision medicine の実践に向けた研究スキーム

【ゲノム情報及び臨床情報の収集、解析】

まず、自己骨格筋由来筋芽細胞シート移植を試行された対象者の絞り込みを行った。当初エントリー

候補患者 19 名のうち、参加拒否と外来通院困難を除く 16 名で解析を行った。その後外来通院困難であった 1 名が令和 4 年 9 月に来院されたため、最終的に検体は 17 名となった。症例計 17 症例に対して研究参加の同意を取得し、臨床情報、ゲノムおよび RNASeq 解析に必要な血液検体、および心筋組織検体を入手した。レスポンドーおよびノンレスポンドーについては BNP 値、6 分間歩行テスト、左室駆出率を指標として定義し 2 群に分けた。これらの症例について血液検体を用いてゲノム DNA を抽出した。ゲノムについては Illumina 社 NovaSeq/HiSeq、および Agilent 社 SureSelect エンリッチメントキットを用いて、まず全エクソーム配列解析を実施した。得られた FASTQ データは国際標準に準拠した GATK Bestpractice の解析パイプラインを用いて FASTQ ファイルの QualityCheck、リファレンスゲノム (hg19) へのマッピングおよび重複リードの除去等のステップを経た。調整後の BAM ファイルに対して GATK HaplotypeCaller を用いて VCF ファイルを作成した。得られたバリエーションリストは心筋症として報告のある遺伝子を各種データベースから幅広く抽出しバーチャルパネルを実施した。また病態から推定される一般集団内でのアリル頻度や ACMG (American College of Medical Genetics) によるバリエーション病原性評価のガイドラインなどにより絞り込みを行い、1) SNV 2) Short In/Del 3) Exon splicing の領域について層別化の上、症例ごとに評価を行った。上記の過程を経て一定のアルゴリズムにより絞り込まれたバリエーションはさらに専門のキュレーターにより既報の論文情報や変異の CADD (Combined Annotation Dependent Depletion) や PROVEAN (Protein Variation Effect Analyzer) などのバリエーション重篤度予測アルゴリズム等を用いてより詳細な評価を行った。結果としてレスポンドー、ノンレスポンドー問わず全症例から複数の病原性バリエーションを検出したものの、ノンレスポンドー群では心筋症として報告頻度が高い ACTN2 や BAG3 遺伝子上の病原性変異を有している症例が多く、その分布には傾向の違いが見られた。さらに遺伝学的に疾患の原因となりうるコピー数多型 (CNV) や構造多型 (SV) を評価、解析するため上記全症例に対して全ゲノム解析も実施した。全ゲノムデータはヘテロ接合型変異に対するリードの分布およびゲノム全体へのリード分布が全エクソーム解析より均一であり、これらの症例に対する上記 CNV、SV 評価を高精度に行うための解析パイプラインを開発し解析を行った。結果としてノンレスポンドー症例で病態に関連する可能性がある遺伝子を含む領域 (8 番染色体短腕約 2.2Mb) の CNV Gain の領域を新たに抽出することが出来た。

続いてトランスクリプトーム解析ではノンレスポンドー (6 症例) およびレスポンドー (1 症例) の心筋サンプルを用いて RNA-seq を実施した。得られたデータについて、1) ノンレスポンドーと GTE x コントロールデータを用いた発現量比較解析、2) ノンレスポンドー群におけるパスウェイ解析を行った。1) の結果、複数の病態関連遺伝子候補が抽出され、全エクソーム/全ゲノム解析にてノンレスポンドーのみが保有しているバリエーション情報と統合することにより数個の候補遺伝子をインフォマティクス解析から抽出することが出来た。また 2) の結果からはノンレスポンドーで活性が有意に変化するパスウェイ (酸化リン酸化に関連するパスウェイ) が同定された。そこで病態層別化因子の探索手法として 2 群のミトコンドリアゲノム領域についても検討を行ったところ、ノンレスポンドー群において心筋症に関する既知バリエーションが同定され、2 か所の高病原性未知バリエーションの検出率がレスポンドー群で有意に高いことが判明した。この因子については更なる詳細な検討が必要ではあるもののレスポンドーにおける心筋保護の因子として機能している可能性も示唆された。

【再生医療等臨床研究の実施】

並行して上記レスポンドー予測基盤技術のバリデーションを行うため、臨床研究に向けた研究計画 (第二種再生医療等提供計画) の策定を行った。ゲノム解析研究に同意を得て検体採取を行った拡張型心筋症患者を対象とした臨床研究計画を策定、再生医療等製品である「ハートシート®」の適用外使用として、認定再生医療等委員会 (大阪大学第一特定認定再生医療等委員会) から実施の了承をえられた。

折しも新型コロナウイルス感染症の蔓延による受診控えにより、予定していた例数を達成することが出来なかった。2023年3月に1例の拡張型心筋症患者について実施計画書に基づき自己骨格筋芽細胞シート移植を実施し観察・検査スケジュールにしたがい臨床情報を取得した。引き続きこれまでの既存症例における臨床データやゲノムデータを用いて作成した機械学習モデルを用いてレスポンドかノンレスポンドかの予測を行い、本症例の6ヶ月の経過をみてモデルの精度を検証する。総症例数が少ないため、このバリデーションデータも含めて新たに機械学習モデル、全ゲノム解析データベースを構築し、今後の症例も蓄積してモデルの精度を向上させ、precision medicineの実現を目指す。

The principal investigators have developed a regenerative therapy for end-stage severe heart failure, in which myoblast cell sheets are transplanted to the cardiac surface (myoblast transplantation). Myoblast transplantation for patients with severe ischemic cardiomyopathy has shown therapeutic efficacy. However, when this therapy was also applied to dilated cardiomyopathy (DCM) in a Phase I clinical trial, it was found that the therapeutic effect was not observed in all patients. In this research project, we considered it important to accurately determine the pathological condition of the patients who were to receive the treatment and to target patients who were predicted to respond to this treatment. Therefore, by obtaining the patient's preoperative biological information, molecular biological information, characteristics of the transplanted cell sheet, and postoperative outcome information, and by creating a risk factor analysis model to accurately predict and select patients who will respond to this therapy (responder) before surgery through statistical analysis and machine learning using artificial intelligence (AI), it would be able to develop a risk factor analysis model to accurately predict and select patients who will respond to this therapy before surgery.

First, we narrowed down the subjects to whom myoblast transplantation was attempted. Of the 19 patients who were initially candidates for entry, the final number of specimens was 17. Consent for study participation was obtained for a total of 17 cases, and clinical information, blood samples required for genomic and RNASeq analysis, and myocardial tissue samples were obtained. Responders and non-responders were divided into two groups based on the BNP value, 6-minute walk test, and left ventricular ejection fraction. Genomic DNA was extracted from peripheral blood of these cases. For the genomic analysis, whole-exome sequencing was first performed using NovaSeq/HiSeq from Illumina and the SureSelect enrichment kit from Agilent. The obtained FASTQ data underwent steps such as QualityCheck of FASTQ files, mapping to the reference genome (hg19), and removal of duplicate reads using the GATK Bestpractice analysis pipeline conforming to international standards. A VCF file was created using GATK HaplotypeCaller for the adjusted BAM file. For the obtained variant list, genes reported to cause DCM were extensively extracted from various databases and a virtual panel was performed. In addition, we narrowed down the allele frequency in the general population estimated from the pathology and the guidelines for variant pathogenicity evaluation by ACMG (American College of Medical Genetics), etc. 1) SNV 2) Short In/Del 3) Exon splicing area was stratified and evaluated on a case-by-case basis. Variants narrowed down by a certain algorithm through the above process are further provided by a specialist curator with information on published papers and variant severity prediction algorithms such as CADD (Combined Annotation Dependent Depletion) and PROVEAN (Protein Variation Effect Analyzer). was used for more detailed evaluation. As a result, multiple pathogenic variants were detected in all cases, regardless of whether they were responders or non-responders, but in the non-responder group there were cases with pathogenic mutations in the ACTN2 and BAG3 genes, which are frequently reported as DCM. In many cases, differences in trends were observed in their distribution. In addition, whole genome analysis was performed for all the above cases in order to evaluate and analyze copy number variations (CNV) and structural variants (SV) that could be genetically responsible for the disease. The distribution of reads for heterozygous mutations and the

distribution of reads over the entire genome of the whole-genome data are more uniform than the whole-exome analysis. I did the analysis. As a result, we were able to newly extract a region of CNV gain in the region (about 2.2 Mb on the short arm of chromosome 8 containing a gene that may be related to pathology in non-responder cases.

Subsequently, RNA-seq was performed using myocardial samples from non-responders (6 cases) and responders (1 case) for transcriptome analysis. The obtained data were subjected to 1) comparative analysis of expression levels using non-responders and GTEx control data, and 2) pathway analysis in the non-responders group. As a result of 1), multiple pathogenic gene candidates are extracted, and several candidate genes are extracted from informatics analysis by integrating with variant information possessed only by non-responders in whole exome/whole genome analysis. In addition, from the results of 2), a pathway (a pathway related to oxidative phosphorylation) was identified in which the activity significantly changed in non-responders. Therefore, we also examined the mitochondrial genome regions of the two groups as a method of searching for pathological stratification factors, and we found that known variants related to DCM were identified in the non-responder group, and we found that the detection rate of two highly pathogenic unknown variants was significantly higher in the responder group. It was suggested that this factor functions as a cardioprotective factor in responders, although further detailed investigation is required.

In parallel, to validate the responder prediction technology, one patient with DCM underwent myoblast transplantation in accordance with the implementation plan, and clinical information was obtained according to the observation and examination schedule. A machine learning model created using clinical and genomic data from existing cases was used to predict whether the patient was a responder or a non-responder, and the accuracy of the model was verified after 6 months of follow-up in this case. Since the total number of cases is small, we will construct a new machine learning model and whole genome analysis database including this validation data and accumulate future cases to improve the accuracy of the model and realize precision medicine.