

厚生労働科学研究費
がん対策推進総合研究事業（業務項目）研究報告書

喫煙量依存的な肺がんリスクを規定する遺伝素因の同定

研究分担者 河野隆志、白石航也、島津太一、アドリアン・シャルヴァ、片野田耕太（国立がん研究センター）、久保充明（理化学研究所）、松田文彦（京都大学）、松尾恵太郎、伊藤秀美（愛知県がんセンター）、醍醐弥太郎（東京大学）

研究要旨：

各施設で既に取得されているゲノム情報を用いて全ゲノムインピュテーションを行いつつ、年齢、性別、喫煙情報等の診療情報を収集した。喫煙との相互作用の検討やメタ解析を通してゲノム網羅的な関連解析を実施するための基盤情報の構築を行った。

A．研究目的

早期診断・外科治療のための高危険度群の捕捉が、肺がん死減少のための最も有効な手段である。喫煙は肺がんリスクを規定する主要因であり、受動喫煙によっても本邦肺がんリスクの1.3倍の上昇が示されている。したがって、喫煙による肺がんの罹患を効率よく減少させるため、本邦の政策の基盤となる「個別化肺がん予防効果の明確なエビデンス」が求められている。本研究では、肺がんリスクにおいて喫煙と相互作用する遺伝要因を用いて能動/受動喫煙者の肺がん絶対リスクを評価し、超高危険度群（相対危険度5以上）を把握することを目的とする。各施設（理研/東大、国がんセ/BBJ、京大/愛知県がんセ）で保有している既存のSNPデータを用いてゲノム網羅的な関連解析を行い、新規感受性遺伝子座の同定を行う。また多施設で収集された症例を用いて検証研究を実施するため、国立がん研究センターと京大と愛知県がんセンターが中心となり、症例を収集し、来年度以降実施予定の検証研究を実施するために必要な症例数の確保を行う。また能動・受動喫煙の情報が得られている症例の収集も合わせて行う。

B．研究方法

各施設で既に取得された肺がん症例のゲノム情報（BBJ・東大：Miki et al., Nat Genet 2010, BBJ・NCCH：Shiraishi et al., Nat Genet 2012, 愛知県がん・京大）と健常群として三つコホートサンプルから得られたゲノム情報を統合した。さらに喫煙と遺伝要因との相互作用を検討するため、喫煙情報が得られた症例のみを抽出した。まずは、インピュテーション前のデータを用いて、構造化や解析手法の整合性を検討した。

一方で、この集団に対して、1000Genomesの全人種をレファレンスゲノムとして用い、これらを元に全ゲノムインピュテーションを実施した。関連解析は、Plink1.9を用いて相互作用の検討を行った。

（倫理面への配慮）

「ゲノム倫理指針」に従って、試料提供者のプライバシーを保護する。

C．研究結果

本研究に用いた集団の構造化などの検討をQQ plotを用いて行い、問題がないことを確認した。また施設間での既取得ゲノム情報を用いた全ゲノムインピュテーションが完了した。今後は、喫煙情報を元に層別化解析を行い、個々の関連解析を実施する予定である。

D．考察&結論

今回の解析結果は、アジア人集団においては過去最大の症例数を用いた関連解析となる。来年度以降は、候補感受性遺伝子座に対して、検証研究を用いて、新規感受性遺伝子座の同定を目指す。今後これらのデータは、リスクモデリングを構築するために必要なものであり、今後肺がんに対する高危険度群捕捉手法の確立が期待される。

F．研究発表

論文発表・学会発表
なし

G．知的財産権の出願・登録状況

なし

研究報告書補足資料

課題名：喫煙量依存的な肺がんリスクを規定する遺伝素因の同定

共同研究者：

国立がん研究センター：河野隆志、白石航也、島津太一、片野田耕太、アドリアン・シャルヴァ

理化学研究所：久保充明

京都大学：松田文彦

愛知県がんセンター：松尾恵太郎、伊藤秀美

東京大学：醍醐 弥太郎

方法・結果：

検出研究を実施するため、理化学研究所/東京大学、国立がん研究センター/バイオバンクジャパン、京都大学/愛知県がんセンターにて肺腺がんと診断された症例群と3つの健常者コホート群において、既を取得されているSNP chip dataを元に、症例と健常群内の集団遺伝構造の有無の検討を行った(図1)。その結果、QQ plotによる構造化を検討したところ、 $\lambda_{GC}=1.10$ 以下であったため、構造化は認められなかった。

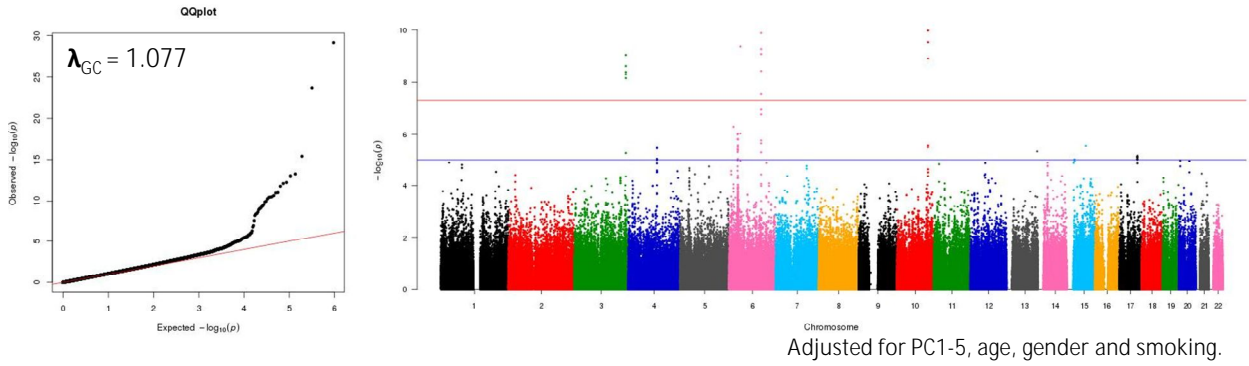
次に喫煙との交互作用を検討するため、喫煙情報が得られた症例の選別並びに診療情報の収集並びに確認を行った。喫煙量に関しては多くの症例で確認できたが、バイオバンクジャパンの症例については、現在喫煙者のみ喫煙本数と喫煙期間の情報が得られるため、症例数を減らしての解析となった。共変量は年齢(49歳以下、50-59歳、60-69歳、70歳以上)、性別、喫煙量(非喫煙者、 $0 < \text{Pack-year (PY)} \leq 20$ 、 $20 < \text{PY} \leq 40$ 、 $\text{PY} > 40$)とし、SNPと喫煙量との交互作用の検討を行った。現状では、各GWASとも少なくともGWASレベルに到達している感受性遺伝子座は認められていない。

最後に各施設で取得されたゲノム情報を用いて、1000Genomesの全人種レファレンスパネルを基にした全ゲノムインピュテーションをImpute2というソフトウェアを用いて実施した。

考察：

SNP chip dataだけでは、現状では全ゲノムレベル($P = 5.0 \times 10^{-8}$)を下回るSNPはまだ得られていない。今後はさらにゲノムデータが取得されている症例を追加して、さらに検出研究を大規模化させる。また公開データベースの活用も検討する。一方で、喫煙量依存的な検討を行う場合、喫煙していない非喫煙者を本解析に含めることが適切であるかも検討する必要がある。今後は非喫煙者を除外して同様の解析を行い、喫煙量依存的な遺伝要因が同定できるかどうかさらに検討する予定である。今後は候補となる感受性遺伝子座に関して検証研究を実施し、喫煙量依存的な肺がん感受性遺伝子があるかどうかを検証する。

(1) NCCHとBBJ肺がん症例



(2) 京都大学と愛知県がんセンター肺がん症例

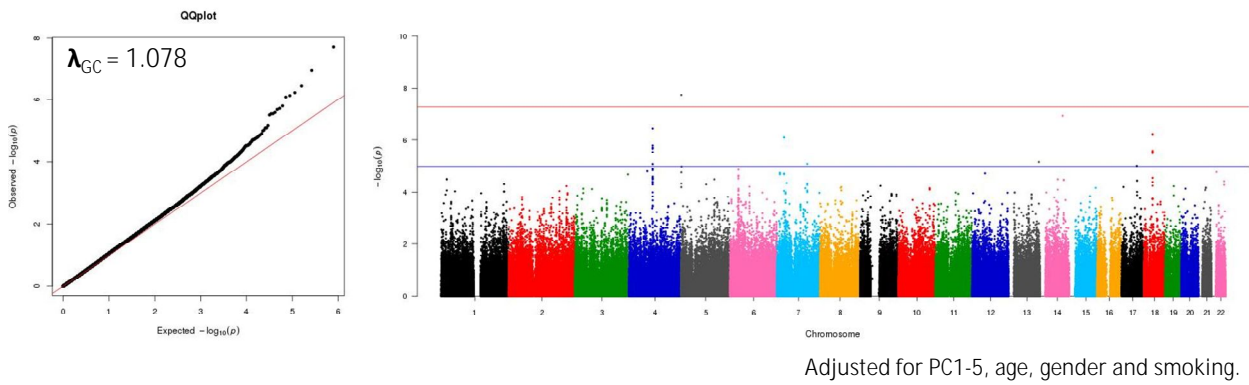
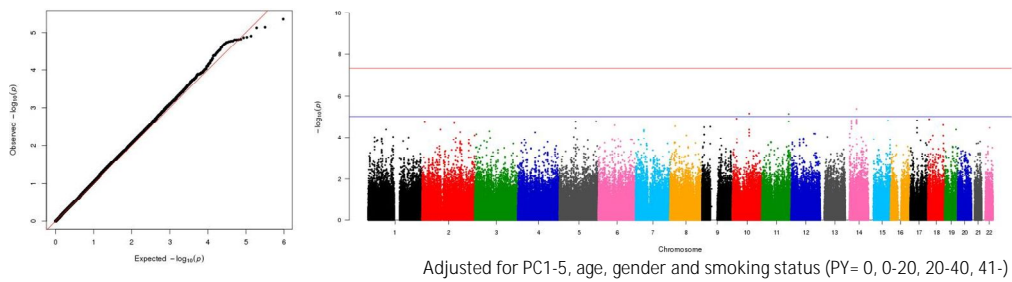


図1. 各施設別での肺腺がんリスクに対する全ゲノム関連解析の結果
 QQ plotによる構造化の有無を検討したところ、 $\lambda_{GC}=1.10$ 以下のため構造化は認められなかった。
 マンハッタンプロットは、横軸が染色体別を示し、縦軸は関連の強さを示す。

(1) NCCHとBBJ肺がん症例



(2) 京都大学と愛知県がんセンター肺がん症例

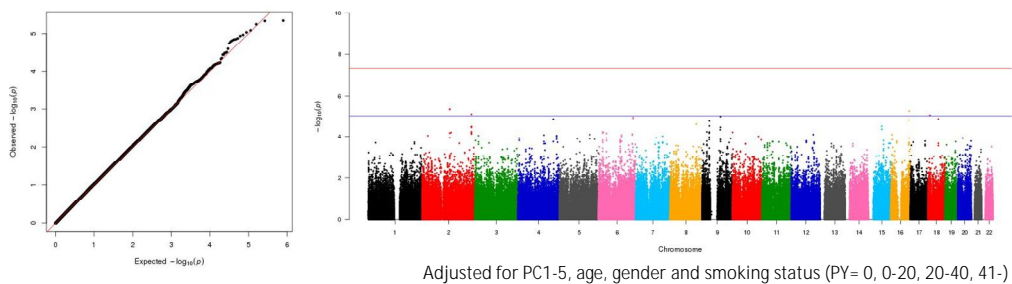


図2. 喫煙量との交互作用の検討