

No	質問	回答
1	マウスサンプルを使った疾患研究をしているのですが、このエンハンサー領域の解析はヒトサンプルでの解析が適切ということでしょうか。	ご質問有難うございます。マウスの5' scRNA-seqのサンプルでも適用可能です。実際にマウスでも解析を別プロジェクトで行っています。
2	5'でscRNAseqを外注にて行う場合、何か注意が必要でしょうか。	ご質問有難うございます。1) Read 1,2共に長め(100bp PEや150bp PEなど)シーケンスすること、2) 5' scRNA-seqのライブラリ以外の構造のライブラリもしくはPhiXが10-20%混ざっていること(固定配列の部分にcomplexityを出すため)、この2つが注意点になります。
3	双方向に転写がみられるbtEIはH3K27acやBRD4とかとも相関は見られているのでしょうか？	ご質問有難うございます。H3K27acやH3K4me1のシグナルと相関は確認しています。BRD4は調べておりませんでした。MED1は相関していました。
4	確からしいエンハンサー領域を解析可能な量読み取るためには、リード数は通常のscRNA-seqよりも多く(深く)読む必要があるのでしょうか？理解が甘く、基礎的な質問で申し訳ございません。	ご質問有難うございます。遺伝子発現解析の時と同等程度、25000-50000 reads/cellで問題ありません。エンハンサー領域をしっかり捉えるにはリード数よりも細胞数(クラスターの中に少なくとも数千細胞)が大事と考えています。
5	UMAPにおいて、通常の遺伝子発現、エンハンサーRNA発現、エンハンサーRNA発現+通常の遺伝子発現を用いた時に細胞の分離は変化しますか？特異的な細胞はみられますか？	ご質問有難うございます。明確に分かれてくるものがあるわけではないが、多少は良くなる印象です。面白いのは転写開始点で見ているのでアイソフォーム毎の発現を見ることも可能です。「ここはロングアイソフォームが、ここはショートアイソフォームが発現している」という違いが見られます。
6	今後はどのようにこの研究が発展されるとお考えですか？	ご質問有難うございます。他の疾患(精神疾患や生活習慣病など)の解明のために、脳などの他の臓器でも同等の解析も可能と思います。また今回見つけたエンハンサー領域でかつマウスでも保存されている領域をゲノム編集で潰して機能を見るなども考えています。
7	エンハンサーRNAは1細胞あたり何種類くらい検出できますか？	ご質問有難うございます。分子数はリードのカウントから直接的に推測するのは難しいですが、deduplicate後のリードの0.5%前後がeRNAに対応していると考えます。
8	エンハンサーRNAの発現量はそのままエンハンサーの活性化度合いと解釈できるものなのでしょうか。	ご質問有難うございます。おおむね発現量が高ければ活性が高い傾向にあります。
9	BulkのCAGEなどと比較した時にoligo dTによるinternal primingでは何%程度eRNAをキャプチャーできているのでしょうか。	ご質問有難うございます。8割程度は両者の方法でオーバーラップしているように思います。
10	Cap由来のGを判定する方法ですが、転写開始点付近のreference配列がpolyGの場合も特定できるのでしょうか？	ご質問有難うございます。リファレンスにGがあるとするとCap由来のGだとしてもReapTECで除くこととなります。PolyGが長いとエンハンサーRNAが取れていないかもしれません。しかしエンハンサーの領域はACGTの出現率が均等なのでGに偏るのは珍しく、大体は網羅できていると考えています。
11	今回の論文でT細胞に着目したのはどのような理由でしょうか	ご質問有難うございます。同定したエンハンサーとGWASのstudyとの統合解析をおこないたいと考えました。1) 自己免疫疾患やアレルギー疾患ではGWAS研究が数多く取り組まれているのでデータも豊富であること。2) CD4+のT細胞は免疫の管制塔として重要で、自己免疫疾患、アレルギー疾患にも関わる面白い細胞であること。3) また、採血からT細胞を採取でき、研究しやすいこと、から今回CD4+T細胞ではじめました。
12	細胞種特異的なエンハンサーが生じるメカニズムはなんでしょうか？(エンハンサー配列の修飾？ 転写因子の発現？)	ご質問有難うございます。エンハンサー配列はACGTの出現率が均等であるパターンで配列を作りやすいこと、また細胞種ごとに転写開始点のbindingが異なることが理由と考えています。
13	同じく理研のHon Chung Chau博士が開発したSCAFEも同様に5' scRNA-seqを用いてTSSを解析する方法ですが比較するとどのようなメリット、デメリットがあるのでしょうか	ご質問有難うございます。メリット①としては、我々の手法では低発現のエンハンサーも検出しやすいことが挙げられます。というのも、我々は最初にunencoded Gの情報で、リードをfilteringしているのに対して、彼は、まずリードのクラスターを作って、そのクラスターの中にどのくらいunencoded Gを持つリードが存在しているかでfilteringしており、いつfilteringを行っているかが違うからです。メリット②としては、我々はあえて両方向性に転写が起きている領域のみに注目していることで、より機能的なエンハンサー領域に絞られていることかと思えます。デメリットとしては、リファレンスにGがあるとどうしても過度にリードをfilteringすることになり、promoter領域の発現値は低く見積もっている可能性がある、ことと思えます。
14	バルクRNA-seqで転写開始点解析を実施してそこまで良い思い出はないのですが、シングルセルになることで精度や情報量に違いはありそうですか？	ご質問有難うございます。Bulk RNA-seq (CAGE?) で必要とされていたRNA量よりもはるかに少ない量でエンハンサー解析ができ、シングルセルとして細胞ごと、クラスターごとの解析ができるだけでなく、bulk扱いでの解析もできますので、情報量ははるかに多いと思います。精度という意味では、同じくらい(CAGE法はそもそもcap trapper法を使っている)、精度が良いが、ReapTEC法を利用するとさらに精度が増す)かと思えます。