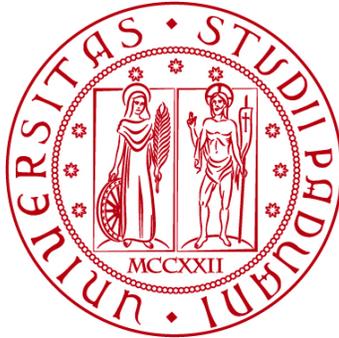


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea in Biologia Molecolare



ELABORATO DI LAUREA

Genotipizzazione ed espressione di *KDM1A* in un'ampia coorte di adenomi ipofisari somatotropi sporadici

Tutor: Prof. Gianluca Occhi
dipartimento di Biologia

Laureanda: Paola Mellia

ANNO ACCADEMICO 2023/2024

INDICE

ABSTRACT	1
1. STATO DELL'ARTE	3
1.1 <i>L'adenoma ipofisario secernente GH</i>	3
1.2 <i>L'aumento paradossoso di GH durante l'OGTT</i>	4
1.3 <i>Il recettore GIPR nel contesto dei tumori endocrini</i>	4
1.4 <i>Il gene KDM1A</i>	5
2. METODICHE	6
2.1 <i>Estrazione di DNA con tre diversi protocolli</i>	6
2.2 <i>Quantificazione tramite NanoDrop</i>	7
2.3 <i>Next Generation Targeted Exome Sequencing</i>	8
2.4 <i>Analisi dell'espressione genica di KDM1A e GIPR</i>	9
2.5 <i>Oligonucleotide-based array-comparative genomic hybridization</i>	9
2.6 <i>Multiplex PCR per l'analisi di STR</i>	10
2.7 <i>Corsa elettroforetica in gel d'agarosio</i>	11
2.8 <i>Elettroforesi capillare</i>	12
2.9 <i>Soluzioni per il miglioramento dell'elettroferogramma</i>	12
3. RISULTATI E DISCUSSIONE	14
3.1 <i>La risposta al test della tolleranza al glucosio orale</i>	14
3.2 <i>Varianti rilevate tramite Targeted exome sequencing e array-CGH</i>	15
3.3 <i>Analisi dei livelli di espressione di GIPR e KDM1A</i>	16
3.4 <i>Risultati della corsa elettroforetica in gel</i>	19
3.5 <i>Risultati della corsa elettroforetica capillare</i>	20
3.6 <i>Strategie testate per l'eliminazione dei picchi non allelici</i>	21
4. CONCLUSIONI	24
BIBLIOGRAFIA	26
APPENDICE	28

ABSTRACT

Il GIPR (Glucose-dependent insulintropic polypeptide receptor) è un recettore accoppiato a proteina G la cui espressione ectopica è stata rilevata in diversi tumori endocrini; in pazienti presentanti iperplasia surrenalica macro-nodulare bilaterale primaria GIP-dipendente la presenza di GIPR sembra essere direttamente correlata a varianti patogenetiche del gene *KDM1A*. In pazienti affetti da acromegalia dovuta all'adenoma ipofisario GH-secernente (GH-PAs), la causa dell'espressione ectopica del recettore non è ancora nota. Tuttavia questa espressione è stata osservata in circa 80% dei pazienti che mostrano un paradossale aumento di secrezione di GH dopo l'esecuzione del test orale di tolleranza al glucosio (OGTT).

Al fine di investigare le cause molecolari che determinano l'espressione alterata del recettore nei GH-PAs, sono stati analizzati i livelli di espressione di *KDM1A* e *GIPR* e sono state ricercate varianti patogenetiche nella sequenza del gene *KDM1A*; l'indagine è stata condotta in una larga coorte di pazienti affetti da acromegalia sporadica impiegando diverse tecniche di biologia molecolare.

In conclusione, non sono state identificate mutazioni germinali nel gene *KDM1A* ma è stata rilevata la sua aploinsufficienza nel 25% dei campioni analizzati, associata a una maggiore espressione di *GIPR* e, non sistematicamente, all'aumento paradossale di GH in seguito alla somministrazione del carico orale di glucosio.

1. STATO DELL'ARTE

1.1 L'adenoma ipofisario secernente GH

Gli adenomi ipofisari secernenti ormone della crescita (GH-PAs) si sviluppano da espansioni monoclonali di cellule somatotrope differenziate, sono di solito sporadici e non producono metastasi anche se in rari casi possono diventare invasivi; in contesto non patologico l'ormone della crescita (GH) è rilasciato dall'ipofisi anteriore, la cui attività è inibita o stimolata dall'ipotalamo, e media diversi effetti a livello cellulare, tissutale e metabolico, sia direttamente che indirettamente grazie al controllo che esercita sul rilascio del fattore di crescita insulino-simile (IGF-1), prodotto principalmente dagli epatociti [5]. L'eccessiva produzione di GH e IGF-1 dovuta agli adenomi somatotropi porta a gigantismo in età infantile e acromegalia in età adulta; questa rara e cronica malattia si sviluppa progressivamente in diversi anni ed è caratterizzata da cambiamenti dei tratti somatici, ingrandimento delle estremità e organomegalia, a cui si aggiungono diverse manifestazioni di comorbilità cardiovascolari, metaboliche, muscolo-scheletriche, endocrine, respiratorie e neoplastiche. La valutazione iniziale dei pazienti con sospetta acromegalia è basata su un'accurata anamnesi e un esame fisico, fondamentale poiché spesso rivela segni e sintomi inequivocabili legati agli effetti diretti e locali della massa tumorale o all'ipersecrezione del GH e IGF-1 [2,3].

In condizioni fisiologiche la produzione dell'ormone della crescita da parte delle cellule ipofisarie somatotrope è stimolata dal legame dell'ormone ipotalamico di rilascio della somatotropina (GHRH) al suo recettore accoppiato a proteina G, la cui subunità α attiva la via di segnale dell'adenilato ciclasi. In presenza di mutazioni attivanti nel gene *GNAS1*, l'attività GTPasica della subunità α , codificata da questo gene, viene inibita, portando all'attivazione costitutiva della via del cAMP/PKA; queste mutazioni sono responsabili dello sviluppo di circa il 30% dei casi di adenomi della ghiandola pituitaria e della conseguente ipersecrezione di GH. Fino ad ora non sono state identificate altre mutazioni *driver* e in meno del 5% dei casi la tumorigenesi ipofisaria è dovuta a mutazioni nella linea germinale di alcuni geni, tra cui *AIP*; nei rimanenti casi gli eventi molecolari che guidano la trasformazione neoplastica rimangono ancora elusivi [1,2]. I principali obiettivi della terapia contro l'acromegalia consistono nel normalizzare i livelli di IGF-1 e ridurre l'ipersecrezione di GH oltre che nell'alleviare i sintomi e gestire le complicazioni associate; l'approccio radioterapico è raramente utilizzato mentre la neurochirurgia tran-sfenoidale rappresenta la soluzione primaria insieme alla somministrazione di diverse classi di farmaci, tra cui gli analoghi della somatostatina e antagonisti del recettore del GH, soprattutto nei pazienti in cui l'intervento chirurgico sia non indicato o non risolutivo [4].

1.2 L'aumento paradossoso di GH durante l'OGTT

Circa il 65% dei somatotropinomi sono macroadenomi alla diagnosi, la quale arriva spesso in ritardo rispetto alle prime manifestazioni dei sintomi; la diagnosi si basa sulla quantificazione biochimica dei livelli sierici di IGF-1 e GH e, se necessario, sull'esecuzione del test di tolleranza al glucosio orale (OGTT). Quest'ultimo consiste nel somministrare al paziente 75g di glucosio per via orale dopo un digiuno notturno per poi rilevare i livelli plasmatici di GH ogni 30 minuti per un periodo di due ore, così da valutarne l'autonomia della secrezione [2]. Questo metodo si basa sull'effetto inibitorio che il glucosio notoriamente esercita sulla secrezione di GH: nei pazienti acromegalici, il glucosio orale non sopprime i livelli sierici di GH al di sotto del nadir, fissato a 0.4 µg/L, come invece accade in pazienti sani. Inoltre, in circa il 30% dei casi, in concomitanza con l'aumento dei livelli di glucosio, si osserva un aumento paradossoso dei livelli di GH; questo fenomeno è dovuto, in più dell'80% dei casi, all'espressione ectopica e conseguente attivazione inappropriata del recettore del peptide insulinotropico dipendente dal glucosio (GIPR) [6].

Di recente è stato condotto uno studio su un'ampia coorte di pazienti acromegalici, suddivisi in paradossi e non a seconda dei livelli di GH misurati dopo il test da carico orale di glucosio, con lo scopo di confrontare i loro profili endocrini, le manifestazioni cliniche della malattia e la risposta alla terapia. Dai risultati è emerso che i pazienti mostranti l'aumento paradossoso di GH presentano tumori più piccoli, meno invasivi e meno frequentemente co-secerenti prolattina, delineando dunque un profilo clinico migliore. Inoltre la risposta del GH allo stimolo iperglicemico in fase di diagnosi può essere utile per predire il successo della terapia medica primaria con SSA, in quanto nel gruppo dei pazienti con risposta paradossosa è stato osservato un tasso di remissione più elevato. Date queste evidenze, l'OGTT potrebbe diventare uno strumento dal potenziale prognostico importante e contribuire al delineamento di terapie mirate e personalizzate [7].

1.3 Il recettore GIPR nel contesto dei tumori endocrini

Il polipeptide insulinotropico dipendente dal glucosio (GIP) è un ormone incretinico secreto dalle cellule K del duodeno e del digiuno prossimale in risposta all'ingestione di nutrienti, in particolare zuccheri semplici e grassi. Partecipa al coordinamento tra assorbimento dei nutrienti e metabolismo, e influenza gli adattamenti sistemici che seguono la digestione metabolica; ha anche un effetto pleiotropico in quanto il suo recettore GIPR è espresso in diversi organi e tessuti quali il cervello, il cuore, lo stomaco e il pancreas, nelle cui cellule β stimola la secrezione di insulina.

Il GIPR è un recettore accoppiato a proteina G che attiva l'enzima adenilato ciclasi e produce come secondo messaggero l'adenosina monofosfato ciclico (cAMP); l'attivazione inappropriata dell'asse GIP/GIPR è ampiamente studiata in diversi disturbi metabolici, tra cui il diabete mellito di tipo 2, e di recente ha iniziato a essere investigata anche in diversi tumori endocrini umani sia per indagare un possibile coinvolgimento nella patogenesi di queste neoplasie, sia per ricercare nuovi approcci per lo sviluppo di metodi alternativi di diagnosi e terapie. Ad esempio in pazienti affetti da iperplasia

surrenalica primaria macronodulare bilaterale (PBMAH), una delle principali cause di ipercortisolemia endogena cronica associate alla sindrome di Cushing, l'espressione ectopica di *GIPR* è l'alterazione molecolare più comune rilevata e mima, in risposta al GIP, la trasduzione del segnale che in condizioni fisiologiche invece coinvolge l'attivazione del recettore MC2R da parte dell'ormone adrenocorticotropo (ACTH). Per quanto riguarda la ghiandola ipofisaria, nel tessuto sano il *GIPR* non è normalmente espresso, anche se livelli molto bassi del suo trascritto sono stati rilevati in alcuni campioni, più probabilmente dovuti alla presenza di cellule endoteliali che esprimono il recettore fisiologicamente.

Al contrario, due studi indipendenti hanno rilevato livelli anomali significativamente elevati di espressione di *GIPR* in circa il 30% degli adenomi ipofisari secernenti GH analizzati e nessuno di questi presentava mutazioni attivanti nel gene *GNAS1*; questa mutua esclusività lascia ipotizzare che i due eventi possano appartenere allo stesso meccanismo patogenetico coinvolgente il cAMP [7].

Evidenze recenti suggeriscono che la sovraespressione del gene *GIPR* possa avvenire precocemente nel processo di trasformazione neoplastica dei tumori endocrini però per quelli ipofisari i meccanismi molecolari responsabili devono ancora essere delucidati. È stato dimostrato che i somatotropinomi $GIPR^+$ sono caratterizzati da un profilo trascrizionale specifico e presentano una maggiore frequenza di riarrangiamenti genomici; per quanto riguarda l'aspetto epigenetico, dall'analisi di metilazione dell'intero genoma in campioni di adenomi somatotropi sono emersi profili diversi tra tumori $GIPR^+/GNAS^{wt}$ e $GIPR^-/GNAS^{mut}$: i primi mostrano un'ipermetilazione nel corpo del gene, non osservata invece nei secondi mentre entrambi non riportano differenze di metilazione nella regione promotoriale [8].

1.4 Il gene *KDM1A*

La Lysine-specific histone demethylase 1A (LSD1), nota anche come lysine (K)-specific demethylase 1A (KDM1A), è una proteina codificata nel genoma umano dal gene *KDM1A*, il cui locus genico si trova nella porzione distale del braccio corto del cromosoma 1 (1p36.12).

È un enzima regolatore dell'espressione genica che rimuove gruppi metili dalla lisina 4 dell'istone 3, contribuendo a creare un ambiente cromatinico non favorevole alla trascrizione; è stato comprovato il suo ruolo critico nella crescita degli ovociti, nell'embriogenesi, nell'ematopoiesi e nella differenziazione tissutale specifica. Al contrario la sua inattivazione risulta in un aumento della metilazione di istoni in regioni promotoriali e non, favorendo la sovraespressione di diversi geni.

Un'ulteriore dimostrazione che le alterazioni epigenetiche possano svolgere un ruolo chiave nell'espressione aberrante di *GIPR* nei tumori endocrini è emersa studiando le lesioni surrenali di pazienti affetti da PBMAH, nelle quali sono state trovate mutazioni germinali inattivanti proprio nel gene codificante *KDM1A*. Date queste nuove evidenze sperimentali, è coerente ipotizzare che le stesse mutazioni potrebbero essere coinvolte anche nell'espressione alterata di *GIPR* negli adenomi ipofisari sporadici GH-secernenti [9]: quest'ultimo aspetto è l'argomento centrale della ricerca che verrà di seguito presentata.

2. METODICHE

Lo studio ha coinvolto una coorte di 146 pazienti con età media 44 anni esaminati per acromegalia sporadica o gigantismo non sindromico tra il 2006 e il 2021 di cui 79 uomini e restanti donne; quest'ultima malattia è stata diagnosticata in 3 pazienti tramite la rilevazione di una velocità di crescita in età giovanile più elevata rispetto ai dati di riferimento o il raggiungimento di un'altezza finale significativamente superiore a quella della popolazione di confronto, insieme al rilevamento di concentrazioni di IGF-1 anormali rispetto agli standard per sesso e età. Per l'acromegalia invece la diagnosi si è basata sulla valutazione di caratteristiche cliniche quali elevate concentrazioni plasmatiche di IGF-1 e concentrazioni di GH non sopresse dopo la somministrazione dell'OGTT, a cui si è aggiunto lo staining del GH positivo eseguito su un campione di tessuto tumorale dopo l'operazione di rimozione chirurgica dell'adenoma. Il valore soglia di GH utile a differenziare i pazienti con una risposta paradossa all'OGTT è stato definito come un aumento maggiore del 20% rispetto al livello basale di GH entro 90 minuti dalla somministrazione del carico di glucosio.

I dati ottenuti dalla risonanza magnetica (MRI) dell'ipofisi hanno permesso di suddividere gli adenomi in microadenomi e macroadenomi e di valutarne l'invasività, definita come livello di invasione del seno cavernoso e delle strutture adiacenti. Successivamente sono state condotte analisi istochimiche e immunoistochimiche in campioni di tessuto tumorale per distinguere gli adenomi esclusivamente somatotropi da quelli somatotropi-lattotropi, positivi sia per il GH che per la prolattina, o da adenomi misti, secernenti altri ormoni.

2.1 Estrazione di DNA con tre diversi protocolli

L'estrazione del DNA dai campioni di adenomi somatotropi dei pazienti coinvolti nello studio è stata eseguita al fine di condurre analisi molecolari mirate alla rilevazione di mutazioni nella sequenza del gene *KDM1A* e di altri geni coinvolti in malattie dell'ipofisi quali *AIP*, *CABLES1*, *CDH23*, *CDKN1B*, *DICER1*, *GNAS*, *GPR101*, *MEN1*, *NF1*, *PI3KCA*, *PRKARIA*, *RET*, *USP8*, *VHL*, oltre che alcuni membri della famiglia delle demetilasi istoniche umane implicate, più in generale, nella tumorigenesi (*KDM3A*, *JMJD2A*, *JMJD2B*, *JMJD3*, *KDM6A*, *KDM5B*).

Il primo metodo sfrutta una miscela di fenolo, cloroformio e isoamil alcol (PCIA) aggiunti in rapporto 25:24:1 e richiede che il tessuto venga precedentemente omogenizzato e lisato, usando un buffer contenente generalmente EDTA, chelante di ioni bivalenti e SDS, detergente anionico che denatura le proteine e emulsiona i lipidi in modo da favorire la completa disgregazione della membrana cellulare. Il prodotto finale che si ottiene è una soluzione bifasica: la fase inferiore fenolica contiene le proteine e le altre macromolecole e può essere eliminata mentre la fase superiore acquosa, dove sono presenti gli acidi nucleici, va aspirata e conservata in una nuova microprovetta dove il DNA possa poi essere precipitato e ri-sospeso. L'estrazione tramite PCIA è laboriosa ma è economica e permette di ottenere un DNA di buona qualità, sebbene altri metodi più costosi garantiscano una resa migliore.

Come secondo metodo di estrazione è stato utilizzato il reagente TRIzol (Euromedex), miscela monofasica di fenolo e guanidina isotiocianato che è un agente caotropico in grado di solubilizzare le membrane, denaturare le proteine e inattivare le DNAsi e RNAsi: questa componente rende il TRIzol indicato per estrarre sia DNA che RNA direttamente dal campione tissutale. Per isolare i due acidi nucleici tra loro è necessario aggiungere cloroformio: l'omogenizzato così si separa in una fase organica inferiore contenente il DNA e in una acquosa superiore in cui si trova l'RNA; trasferendo le due fasi in diverse micropiastre è possibile far precipitare il relativo acido nucleico, rispettivamente aggiungendo etanolo al 70% e isopropanolo.

Infine è stato usato il kit di estrazione Quick-DNA/RNA (Zymo Research) che permette di ottenere sia un DNA genomico di alta qualità che l'RNA totale sfruttando colonnine di spin ricoperte da una matrice silicea a cui aderiscono gli acidi nucleici, mentre le altre macromolecole e impurità vengono lavate via con soluzioni di lavaggio.

2.2 Quantificazione tramite NanoDrop

Dopo l'estrazione degli acidi nucleici è importante valutarne la quantità sia per rilevare la presenza di contaminanti potenzialmente interferenti con la buona riuscita delle analisi successive sia perché ogni protocollo richiede l'utilizzo di una specifica quantità di acido nucleico e dunque conoscerne la concentrazione risulta imprescindibile.

La tecnologia NanoDrop è un'evoluzione dello spettrofotometro classico che permette di usare dei volumi minimi (1-2 μ L) per calcolare la concentrazione di soluzioni contenenti DNA a doppio o singolo filamento e anche RNA; il principio di funzionamento è basato sulla capacità delle basi puriniche e pirimidiniche di assorbire radiazioni UV, con un picco di assorbanza quando la lunghezza d'onda è 260nm: tramite la legge di Lambert-Beer è possibile correlare la quantità di luce assorbita dal DNA caricato con la sua concentrazione. Il software inoltre valuta la purezza dell'acido nucleico utilizzando dei rapporti tra assorbanze ottenute a diverse lunghezze d'onda: A260/A280 è l'indice di contaminazione da proteine mentre A260/A230 informa sulla presenza di carboidrati; se il valore calcolato rientra in un certo range ottimale, l'acido nucleico si può ritenere puro. Mentre NanoDrop è un'ottima tecnologia per ottenere una precisa e rapida misura della concentrazione di DNA, ha una sensibilità più limitata rispetto ai metodi basati sulla fluorescenza e non è in grado di informare della presenza di eventuale RNA nel DNA estratto.

2.3 Next Generation Targeted Exome Sequencing

Il next generation sequencing è una tecnologia sviluppata da quasi due decenni che ha rivoluzionato il modo di analizzare le sequenze di DNA, rendendo il processo precedentemente basato sul sequenziamento Sanger molto più rapido, efficiente e semplice. La principale caratteristica delle tecnologie di seconda generazione, tra cui il sistema Illumina Sequencing by Synthesis, è la miniaturizzazione e parallelizzazione del processo di sequenziamento: invece che dover analizzare una sequenza proveniente da un singolo campione per volta, il sistema permette di ottenere la sequenza di milioni di frammenti di DNA, anche appartenenti a campioni diversi, clonalmente amplificati e letti contemporaneamente in una singola corsa (run) di sequenziamento. Inoltre, grazie ad una maggiore sensibilità della tecnica è possibile rilevare varianti genetiche rare o nuove. Per questo lavoro è stato utilizzato il sequenziatore Miseq® Illumina™ sequencer analizzando i frammenti di DNA appartenenti a 23 pazienti per run.

Prima del sequenziamento i templati, ossia le molecole di DNA d'interesse, devono essere correttamente preparati affinché il sistema possa analizzarli: il DNA estratto infatti deve essere reso compatibile alla piattaforma in modo che possa ibridarsi. La costruzione della libreria dell'intero genoma è stata completata tramite l'utilizzo del NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina (NEB Inc.); il protocollo richiede un primo step di frammentazione del DNA estratto, che favorisce anche l'aumento dell'uniformità del coverage e un successivo step di riparazione delle estremità e ligazione degli adattatori, molecole complementari agli oligonucleotidi ancorati sulla piattaforma (flow cell). Gli adattatori contengono anche una breve sequenza, l'indice (index), che serve per associare ogni frammento di DNA correttamente e univocamente al campione di provenienza.

Il SureSelect^{XT} custom kit (Agilent) è stato utilizzato per l'arricchimento delle porzioni delle regioni esoniche dei geni d'interesse in quanto per questo lavoro era richiesto solo il sequenziamento dell'esoma e non dell'intero genoma; il kit è composto da biglie magnetiche rivestite di streptavidina a cui si legano brevi sequenze complementari (probes) alle porzioni esoniche da isolare.

Attraverso il paired-end sequencing si sono ottenute delle reads della lunghezza di 150pb che fanno entrambe riferimento allo stesso frammento di DNA ma a partire dalle due estremità opposte; i dati vengono sottoposti a una prima fase di elaborazione che consiste nel controllare la qualità di ogni read: soprattutto nel contesto di ricerca di possibili varianti patogeniche, è estremamente importante avere la certezza che ogni nucleotide sia stato letto correttamente, altrimenti si incorre nel rischio di annotare una base errata come una possibile variante.

L'analisi delle reads prosegue poi con il loro allineamento sul genoma scelto di riferimento, in questo caso genoma umano hg19, tramite software che sfruttano algoritmi veloci e complessi come Burrows-Wheeler Aligners; il file che restituisce è poi analizzato da un secondo software, in questo caso Annovar e SNPeff 4.0, che esegue il variant calling ossia compara la sequenza delle reads allineate con la sequenza di riferimento e restituisce tutte le differenze di singola base o riarrangiamenti più grandi che sono presenti.

2.4 Analisi dell'espressione genica di *KDM1A* e *GIPR*

Al fine di valutare i livelli di espressione dei geni *KDM1A*, *GIPR* e *HMBS* (idrossimetilbilano sintetasi) in campioni tumorali e di tessuto sano è stata condotta un'estrazione totale di RNA con i metodi già riportati; la resa è stata valutata tramite lo spettrofotometro NanoDrop mentre la qualità dell'RNA è stata determinata con l'Agilent Bioanalyzer 2100 che effettua un'elettroforesi capillare miniaturizzata all'interno di un chip e restituisce un elettroferogramma che informa su livello di degradazione dell'RNA e contaminazione da DNA genomico, eliminata successivamente tramite trattamento con Turbo DNA-free kit. Il gene *HMBS* è stato usato come controllo in quanto gene costitutivo (housekeeping) i cui livelli di trascrizione e traduzione rimangono stabili negli adenomi ipofisari e nel tessuto sano. L'RNA totale quindi è stato retrotrascritto con la trascrittasi inversa Maxima Reverse sfruttando random esameri come primers, in modo da avere una maggiore garanzia che l'intera lunghezza delle molecole fosse convertita in cDNA.

Per la quantificazione è stata usata la digital droplet PCR: si basa sul principio della PCR classica, ma introduce una novità, ovvero la separazione del campione in nano partizioni, le droplets, che permettono di ottenere la quantificazione assoluta, e non relativa, dei livelli di espressione dei geni target.

Prima di eseguire le analisi dei campioni è stato ottimizzato un sistema di analisi in parallelo (multiplexing) che permettesse, in una singola reazione di ddPCR, di rilevare i livelli di espressione dei tre geni d'interesse (*KDM1A*, *GIPR*, *HMBS*) contemporaneamente; sono state testate su un campione di cDNA di controllo diverse temperature, da 58°C a 62°C, e diversi rapporti di concentrazioni tra le sonde per l'amplificazione e la ddPCR™ Super Mix per le sonde.

Per poter essere suddiviso in droplets, il cDNA deve essere preparato con l'aggiunta di sonde TaqMan, a cui sono associati fluorofori reporter; questo permette al QX200 Droplet Reader, durante la fase di rilevamento dopo l'amplificazione, di misurare il segnale fluorescente all'interno di ogni singola droplet. Tutte quelle contenenti almeno una copia del DNA target sono considerate positive e il rapporto tra queste e quelle negative determina la concentrazione dei geni target nei campioni; tutti i campioni contenenti meno di 10000 droplets sono stati esclusi dalle analisi mentre per i restanti i dati sono riportati come rapporto tra segnale del gene target, *GIPR* o *KDM1A*, e segnale del gene costitutivo *HMBS*.

2.5 Oligonucleotide-based array-comparative genomic hybridization

L'ibridazione genomica comparativa basata su oligonucleotidi (array-CGH) permette l'identificazione di regioni genomiche duplicate e delete e il rilevamento di variazioni del numero di copia di cromosomi. La tecnica richiede l'utilizzo di una piattaforma (array) su cui sono fissati degli oligonucleotidi raggruppati in spot e complementari ai DNA genomici tra cui si vuole condurre il confronto, ovvero il DNA estratto dai campioni ipofisari (DNA test) e il DNA ottenuto da campioni di sangue derivanti dagli stessi pazienti (DNA reference). A seconda dell'applicazione, gli oligonucleotidi possono essere disegnati in modo da rappresentare l'intero genoma o solo determinate

regioni cromosomiali rilevanti e stabiliscono il livello di precisione con cui poi si identificheranno le regioni cromosomiali aberranti.

Per ogni paziente, i due DNA sono marcati con due fluorofori diversi e ibridati sull'array contemporaneamente, cosicché il sistema di rilevazione, ovvero l'Agilent Microarray Scanner, possa misurare il segnale emesso da ogni spot. In teoria, senza la normalizzazione dei dati e senza commettere errori di misura, per ogni alterazione genomica si dovrebbe sempre ottenere lo stesso risultante rapporto di intensità di fluorescenza tra i due DNA ibridati; nella realtà invece può essere che non tutte le cellule del campione presentino le stesse aberrazioni genomiche oppure che ci siano delle contaminazioni derivate da cellule di tessuto sano adiacente a quello tumorale, dunque sono necessari software capaci di interpretare le immagini, come il CytoGenomics software 4.0.3.12. Le alterazioni nel numero di copie che emergono per ogni cromosoma identificato possono poi essere ulteriormente filtrate secondo criteri arbitrari imposti dallo sperimentatore in modo da mantenere solo quelle d'interesse. Questa tecnica è laboriosa e ha delle limitazioni in quanto non può rilevare aberrazioni cromosomiche come traslocazioni bilanciate e inversioni ma rimane un metodo ad alto rendimento che, tramite l'utilizzo di arrays ad elevata risoluzione, è in grado di generare un rilevante numero di dati significativi.

2.6 Multiplex PCR per l'analisi di STR

Siccome la quantità di DNA estratto da 42 dei 146 pazienti è risultata insufficiente per condurre l'analisi di array-CGH, in questi campioni non si è potuto determinare se fosse presente la delezione di uno dei due alleli di *KDM1A*: i dati che si sarebbero potuti estrapolare sono purtroppo rimasti inaccessibili. Successivamente alla conclusione di questo lavoro, per poter analizzare anche i campioni di DNA poco concentrati, è stata messa a punto una diversa strategia, riportata in questa tesi, che richiedesse poco DNA per poter svolgere le analisi e fornisse dei buoni risultati, basata sull'analisi di specifici marcatori polimorfici tramite Polymerase Chain Reaction (PCR).

La multiplex PCR permette di esaminare molti loci genici simultaneamente amplificando in modo ripetitivo e estremamente selettivo diverse regioni di DNA delle quali si conoscono le sequenze terminali; in questo contesto, le sequenze da amplificare contengono tre short tandem repeats (STRs) d'interesse (AFM113XC7, AFMA300WB9 e AFMA129ZB5) adiacenti o interni al locus di *KDM1A*.

Il DNA estratto deve essere aggiunto a una soluzione, la master mix, nella quale si trovano tutte le componenti necessarie alla reazione di amplificazione: lo scopo della mix è mimare, in vitro, le condizioni che l'enzima DNA polimerasi termostabile trova in vivo in modo da favorire il suo corretto funzionamento. L'enzima scelto per questo protocollo è la hot-start Taq GOLD (ThermoFisher) che fornisce prodotti altamente specifici grazie all'aumentata sensibilità e alle modificazioni chimiche che la rendono inattiva a temperatura ambiente.

Al tampone di reazione (*buffer*), necessario al mantenimento costante del pH e della forza ionica, si aggiungono i desossiribonucleotidi (dNTPs) per la sintesi delle nuove eliche e il MgCl₂ come cofattore della polimerasi; seguono gli inneschi *forward* e *reverse* (primers), ovvero due brevi sequenze nucleotidiche a singola elica

complementari alle regioni esterne degli STR da amplificare e per ultima la Taq GOLD. Siccome questa PCR in specifico è seguita dalla corsa in elettroforesi capillare, i primer *forward* sono M13-*tailed* ovvero all'estremità 5' contengono una sequenza nota; utilizzandoli in combinazione con una terza tipologia di primer, gli M13 primer marcati con un fluoroforo all'estremità 5', permettono di rilevare il filamento *forward* durante la corsa in elettroforesi capillare.

	reagenti	volume per singolo campione μL
	buffer di reazione (10X)	1.25
	dNTPs (10mM)	0.25
	MgCl ₂ (25mM)	0.75
AFM113XC7	Primer Forward (10mM)	0.125
	Primer Reverse (10mM)	0.5
AFMA129ZB5	Primer Forward (10mM)	0.125
	Primer Reverse (10mM)	0.5
AFMA300WB9	Primer Forward (10mM)	0.125
	Primer Reverse (10mM)	0.5
	M13 marcato (10mM)	1
	Taq GOLD	0.08
	acqua DNAsi free	6.295
	TOTALE	11.5

La reazione avviene in un termociclatore e prevede il succedersi di 35 cicli di amplificazione, ciascuno composto da tre step di durata definita e con specifiche temperature: la denaturazione per 30 secondi a 94°C permette la separazione delle due eliche del DNA, l'annealing per 30 secondi a 60°C consente ai primers di appaiarsi alle regioni del template a cui sono complementari e l'estensione, durante la quale la TaqGOLD sintetizza, in direzione 5'-3', la nuova elica complementare di ogni frammento, per 30 secondi a 72°. Prima di iniziare con i cicli sono necessari 10 minuti a 94°C per attivare la polimerasi e dopo l'ultimo ciclo di amplificazione la fase di allungamento viene prolungata per 12 minuti.

2.7 Corsa elettroforetica in gel d'agarosio

La corsa elettroforetica in gel d'agarosio è stata impiegata per controllare la riuscita della precedente PCR; la tecnica prevede la preparazione di un gel al 2% a partire da agarosio in polvere disciolto nel tampone tris-acetato-EDTA (TAE) che funge anche da tampone di corsa. Viene aggiunto anche un colorante fluorescente per acidi nucleici, il GelRed (concentrato 1:50000), che permette la visualizzazione al transilluminatore delle bande di DNA ottenute. Una volta che il gel è solidificato, ai campioni può essere aggiunta un aliquota di Orange Loading dye, colorante che permette di seguire la migrazione dei campioni durante l'elettroforesi e possono essere caricati nei pozzetti. Viene anche caricato un marcatore di peso molecolare ovvero una miscela di frammenti di DNA a peso molecolare noto che si distribuiscono lungo il gel e servono come riferimento per valutare il posizionamento delle bande d'interesse ottenute dai prodotti amplificati.

Applicando un campo elettrico tramite l'inserimento, nella cassetta contenente il gel, di due elettrodi, i frammenti di DNA, carichi negativamente, migrano verso l'anodo attraverso la matrice del gel tanto più velocemente quanto minore è la loro lunghezza.

2.8 Elettroforesi capillare

L'elettroforesi capillare è sempre una tecnica di separazione che sfrutta un campo elettrico per distinguere frammenti di DNA con diverse lunghezze; rispetto all'elettroforesi su gel ha una risoluzione molto più elevata poiché permette di separare molecole che differiscono in lunghezza anche solo per un singolo nucleotide e dunque, in un normale gel d'agarosio, risulterebbero accorpati in un'unica banda.

Il sistema di rilevazione si compone di un laser che emette un fascio di luce a una certa lunghezza d'onda in grado di eccitare i fluorofori con cui gli ampliconi sono stati marcati e un detector all'estremità di uscita del capillare che assorbe la fluorescenza emessa; sfruttando fluorofori caratterizzati da diversi spettri di assorbimento e emissione è possibile marcare più campioni da far correre insieme. Il segnale viene inviato a un sistema di elaborazione dati che restituisce un elettroferogramma, ovvero un grafico costituito da diversi picchi, ciascuno dei quali rappresenta una popolazione di molecole di DNA della stessa lunghezza; per calcolare la lunghezza dei frammenti con precisione e valutare la qualità dei risultati è possibile usare il software Peak Scanner (ThermoFisher). Se per un singolo locus STR vengono visualizzati due picchi significa che quel campione è eterozigote mentre se il picco è singolo, entrambi gli alleli sono uguali e dunque è omozigote.

I prodotti di PCR ottenuti devono essere diluiti prima della corsa poiché a concentrazioni troppo elevate l'altezza e larghezza dei picchi aumenta, risultando nella perdita di risoluzione. I campioni possono poi essere aliquotati nella miscela di reazione, composta da LIZ, marcatore di posizione contenente frammenti marcati con uno specifico fluoroforo, acqua e formamide, necessaria per la denaturazione dei frammenti di DNA. Per favorire la completa separazione delle due eliche viene eseguita una breve denaturazione a 90° per 3 minuti al termociclatore.

2.9 Soluzioni per il miglioramento dell'elettroferogramma

L'interpretazione dell'elettroferogramma può essere compromessa a causa della presenza di picchi non informativi minori nell'intorno di quelli allelici, dovuti alla sintesi di ampliconi di lunghezza variabile tutti generati a partire dallo stesso template. Probabilmente la differenza nel numero di paia di basi è dovuta dall'aggiunta di una breve coda di adenosine di lunghezza non fissa all'estremità 3'-terminale dei prodotti di PCR catalizzata dalla DNA polimerasi stessa (*non-template addiction*).

Al fine di rimuovere questi picchi non allelici sono stati testati diversi metodi con lo scopo di uniformare la lunghezza di tutti gli ampliconi successivamente alla reazione di PCR; il DNA scelto per condurre le diverse prove era già stato amplificato con successo in precedenza.

La prima strategia è stata l'aggiunta di un'adenosina all'estremità 3' dei prodotti di PCR (*3'A-overhangs*), che devono però essere precedentemente purificati; dai

campioni appena amplificati quindi sono state eliminate tutte le componenti residue come dNTPs e primers tramite l'utilizzo del kit Wizard SV Gel and PCR Clean-Up System (Promega), composto da colonnine internamente ricoperte da una matrice silicea alla quale aderiscono gli ampliconi, mentre le altre componenti vengono lavate via con buffer salini. Successivamente si può preparare la miscela di reazione con il buffer da PCR già contenente gli ioni Mg^{2+} , la GOTaq DNA polimerasi (Promega), i prodotti di PCR purificati e i nucleotidi deossadenosina trifosfato (dATPs); la reazione avviene nel termociclatore o termoblocco per 30 minuti a 72°C.

reagenti	volume per singolo campione μL
prodotti di PCR purificati	12
dATPs (10mM)	1
Buffer con Mg^{2+} (10x)	10
Taq GOLD (5 U/ μ L)	0.2
ddH ₂ O	26.8
TOTALE	50

Come secondo metodo è stato eseguito un trattamento con ExoSAP: a ogni campione viene aggiunto un isovolume di fosfatasi alcalina ricombinante di gambero (SAP) per degradare i dNTPs rimanenti, e di esonucleasi I (Exo I), enzima idrolitico in grado di digerire in direzione 3'-5' primers e qualsiasi porzione di frammento a singolo filamento presente, rendendo piatte le estremità sporgenti degli ampliconi. La reazione viene condotta in termociclatore o termoblocco per 15 minuti a 37°C, durante la quale avviene la degradazione, e poi 15 minuti a 80°C per inattivare gli enzimi.

3. RISULTATI E DISCUSSIONE

3.1 La risposta al test della tolleranza al glucosio orale

Uno dei parametri valutati per la diagnosi dell'acromegalia è la secrezione autonoma di GH in risposta alla somministrazione del carico orale di glucosio; dei 146 pazienti coinvolti in questo studio affetti da gigantismo o acromegalia, 137 dispongono dei dati dell'OGTT. La loro concentrazione di GH è stata espressa come percentuale rispetto ai livelli basali misurati a digiuno; in 37 (27.4%) pazienti è stato rilevato un aumento di GH entro 90 minuti dalla somministrazione del carico orale pari o superiore al 20% rispetto al livello basale, ossia hanno mostrato l'aumento paradossale di GH (*figura 1*). Questo era stato già osservato in studi precedenti, nei quali però i criteri per differenziare i pazienti con una risposta paradossa da quelli con risposta normale erano diversi e di conseguenza variabile anche la numerosità dei due gruppi.

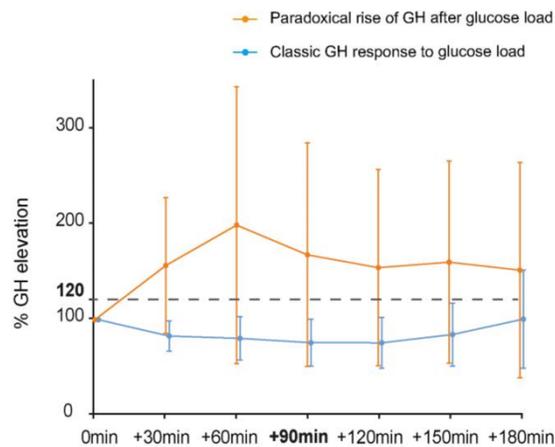


Figura 1: dati espressi come media \pm SD della risposta del GH dei pazienti paradossi (arancioni) e non (blu) durante l'OGTT

A partire da questi dati, è stato condotto un confronto tra i pazienti mostrandoti la risposta paradossa (GH-Par) e non (GH-NPar), considerando diversi parametri clinici, biochimici e tumorali; le differenze in età e sesso riscontrate nei due gruppi non risultano statisticamente significative, così come la risposta a trattamenti precedenti alla rimozione chirurgica e la presenza di tessuto tumorale residuo dopo l'operazione. Per quanto riguarda i parametri biologici misurati, i livelli basali di GH tra i due gruppi di pazienti non differiscono significativamente, suggerendo che questo valore non sia correlato alla risposta paradossa mentre i livelli di IGF-1 risultano più elevati nel gruppo GH-Par con un p-value molto significativo; la prolattina è risultata più concentrata nei pazienti con risposta del GH normale e, coerentemente, si è trovato che i tumori secernenti altri ormoni oltre al GH sono presenti in circa la metà dei pazienti GH-NPar, mentre per i pazienti GH-Par sono molto più comuni adenomi esclusivamente somatotropi. Dai risultati degli MRI la proporzione di macroadenomi rispetto ai microadenomi è simile in entrambi i gruppi ma i tumori dei pazienti GH-Par sono meno comunemente invasivi e tendono ad avere un diametro massimo inferiore.

Le differenze significative emerse per i parametri radiologici e biologici suggeriscono che esistano delle correlazioni con la risposta del GH al carico orale di glucosio, dunque viene riconfermato che l'OGTT ha un valore diagnostico importante, soprattutto in presenza di casi nei quali l'aumento dei valori di IGF-1 non sia così netto. Per questo motivo rimane una valida risorsa su cui continuare a investigare in quanto in futuro potrebbe diventare uno strumento per la caratterizzazione fenotipica precoce dei pazienti acromegalici e dei loro tumori.

3.2 Varianti rilevate tramite Targeted exome sequencing e array-CGH

Dai risultati del sequenziamento con tecnologia Illumina, mirato a valutare mutazioni in un pannello di geni d'interesse, non sono emerse varianti patogenetiche nel gene *KDM1A* già note e solo in un paziente ne è stata rilevata una di significato incerto che nel Genome Aggregation Database è associata a una frequenza molto bassa. Anche nei geni codificanti per le demetilasi istoniche non sono state rilevate varianti patogenetiche mentre 38 adenomi somatotropi sono risultati positivi alle mutazioni del gene *GNAS1* e 4 pazienti hanno mostrato una mutazione *gain-of-function* nel codone Arg201 di *GNAS*, rilevata anche in pazienti affetti dalla sindrome di McCune-Albright; in totale quindi sono risultati 42 i pazienti con mutazioni in questo gene e solo 5 di loro hanno anche mostrato l'aumento paradossale di GH dopo l'OGTT.

Per il gene *AIP* sono state identificate in 6 pazienti delle varianti patogenetiche, di cui una era una mutazione di frameshift presente nel 45% delle reads provenienti dal tessuto tumorale ma assente invece in quelle provenienti da DNA germinale, suggerendo quindi un possibile mosaicismo; solo in un campione sono risultate presenti sia la mutazione nel gene *AIP* che *GNAS*, entrambe assenti nel DNA germinale.

È stata riportata una mutazione di *GPR101* in un solo paziente e in tutti gli altri geni analizzati sono state rilevate solo varianti non maligne oppure non caratterizzate e di significato incerto.

Tramite l'array-CGH è stato analizzato il DNA di 104 adenomi ipofisari ed in 35 di questi è stata rilevata una delezione ricorrente nel braccio corto del cromosoma 1, comprendente anche il locus del gene *KDM1A* in 29 pazienti.

Nella *figura 2* sono riportati i risultati sia del Targeted exome sequencing che dell'array-CGH. Ogni campione è rappresentato da una colonna e per ogni riga si trovano i risultati degli esperimenti condotti. Nella prima riga tutti i 146 pazienti sono suddivisi in tre categorie a seconda del loro profilo di GH durante l'OGTT: in grigio ci sono i pazienti di cui questa informazione manca, a seguire in arancione è evidenziato il gruppo GH-Par e in blu quello GH-NPar. Le due righe successive sono occupate dai risultati dell'ibridazione genomica comparativa: rispettivamente rappresentano l'alterazione nel numero di copie del braccio corto del cromosoma 1 e le duplicazioni o delezioni comprendenti il locus *KDM1A*. A seguire le righe quattro, cinque e sei riportano i dati delle mutazioni somatiche o varianti patogenetiche rilevate nei geni *GNAS*, *AIP* e *KDM1A*. Questa rappresentazione grafica permette di notare che un terzo dei pazienti mostranti l'aumento paradossale di GH dopo l'OGTT presenta aploinsufficienza di *KDM1A*, rilevata invece solo nel 14.3% dei pazienti con risposta

normale all'OGTT; l'aumento paradossico di GH è quindi maggiormente osservato in pazienti con aploinsufficienza e i dati indicano che esista una correlazione significativa ma non esclusiva, inducendo a pensare che altre componenti non ancora note siano coinvolte. Al contrario, le mutazioni attivanti nel gene *GNAS* sono tre volte più prevalenti nel gruppo GH-NPar e solo in tre pazienti coesistono con la delezione di un allele di *KDM1A*. Inoltre, grazie alla profondità di sequenziamento impostata, emerge chiaramente l'architettura subclonale dei tumori ipofisari già descritta da altri studi precedentemente: i campioni analizzati in questo lavoro infatti presentano eterogeneità intra-tumorale, ossia coesistenza di molteplici sotto-popolazioni cellulari clonali caratterizzate da diversità a livello genomico.

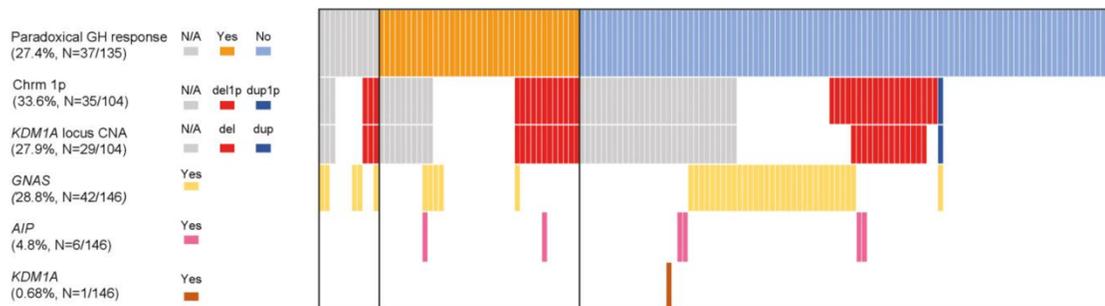


Figura 2: per ogni paziente (colonne) sono riportati nelle righe i risultati, se presenti, della risposta del GH all'OGTT, dell'array-CGH e del sequenziamento

3.3 Analisi dei livelli di espressione di *GIPR* e *KDM1A*

Per 28 campioni l'RNA estratto non è risultato sufficiente a condurre analisi di espressione genica mentre per i 118 campioni restanti, il livello di espressione di *GIPR* è risultato significativamente più elevato in un sottogruppo di pazienti, tutti mostrandoci la risposta di GH paradossa all'OGTT (figura 3).

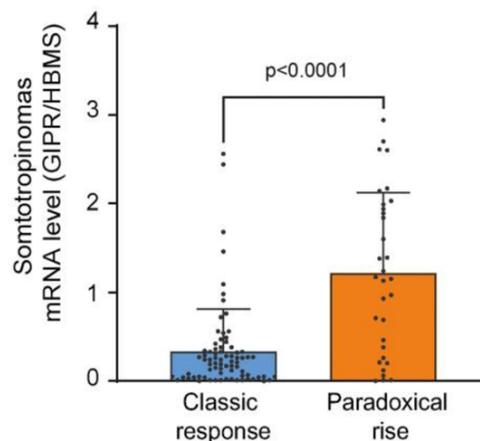


Figura 3: livelli di espressione relativo di *GIPR* nei campioni tumorali quantificato tramite ddPCR

Questi risultati supportano l'ipotesi già formulata in precedenza che l'espressione di *GIPR*, caratterizzante un determinato sottogruppo di adenomi ipofisari, non sia dovuta semplicemente a un'aumentata attivazione trascrizionale del gene ma piuttosto sia da considerare ectopica, in quanto nel tessuto ipofisario sano questo recettore risulta scarsamente presente, se non del tutto assente.

Invece, essendo che l'espressione massima assoluta di *GIPR* rilevata nei campioni analizzati (200 molecole/ μ g di mRNA) è risultata molto inferiore rispetto a quella rilevata nelle lesioni surrenali dei pazienti affetti da PBMAH dipendente da GIP (80000 molecole/ μ g di mRNA), è molto probabile che i meccanismi responsabili della sua espressione ectopica in questi due tumori endocrini differiscano; questo giustificherebbe anche la totale mancanza di varianti patogenetiche in *KDM1A*.

Mentre l'espressione ectopica aberrante di *GIPR* nella ghiandola surrenalica e la secrezione di ormone adrenocorticotropo sono correlate all'inattivazione biallelica di *KDM1A*, nei tessuti analizzati tumorali ipofisari non è mai stata rilevata la delezione di entrambi gli alleli del gene: è probabile che l'allele rimanente funzionale sia sufficiente a esercitare, almeno in parte, la sua attività di repressione dell'espressione genica e dunque potrebbe spiegare i livelli relativamente più bassi dell'espressione di *GIPR* in questi tumori.

Diversamente da quanto atteso, considerando tutti i 118 campioni i cui dati di espressione genica erano disponibili, la correlazione tra l'aumentata espressione di *GIPR* e diminuita espressione di *KDM1A* non è risultata significativa, come osservabile dalla *figura 4*, forse perché il legame tra questi due geni non è diretto, se non in specifici contesti. Al contrario, considerando di questi 118 solo gli 83 pazienti di cui anche i dati di array-CGH erano disponibili, è emersa una netta differenza nei livelli di espressione di *GIPR* tra gli adenomi ipofisari portanti la delezione di *KDM1A* e quelli con entrambi gli alleli (*figura 6*); inoltre, come previsto, nei campioni monoallelici, il livello di espressione di *KDM1A* è risultato inferiore rispetto ai campioni non deleti (*figura 7*).

I campioni sono stati ulteriormente suddivisi in due sottogruppi, $GIPR^+$ e $GIPR^-$, usando come criterio discriminante la quantità di trascritto del gene *GIPR* nel tessuto tumorale rispetto a quello rilevato nel tessuto ipofisario sano; in nessuno dei campioni $GIPR^+$ sono state rilevate mutazioni nel gene *GNAS* (*figura 8*), e considerando tutti e soli i campioni $GIPR^+$ e $GIPR^-/GNAS^{wt}$ risulta un livello di correlazione inversa più significativo tra l'espressione del recettore e l'aploinsufficienza di *KDM1A* (*figura 5*). L'assenza totale di mutazioni attivanti nel gene *GNAS* negli adenomi ipofisari $GIPR^+$ rafforza la preesistente ipotesi che questi due geni siano implicati in due alternativi, e potenzialmente mutualmente esclusivi, meccanismi di trasformazione neoplastica degli adenomi ipofisari.

Inoltre, il 65,4% dei campioni con aploinsufficienza di *KDM1A* sono risultati $GIPR^+$ mentre dei campioni con entrambi gli alleli di *KDM1A* solo il 15,8% è stato classificato $GIPR^+$ (*figura 9*). Questa differenza è statisticamente significativa e supporta l'ipotesi che il numero di alleli di *KDM1A* influenzi l'espressione del gene *GIPR*.

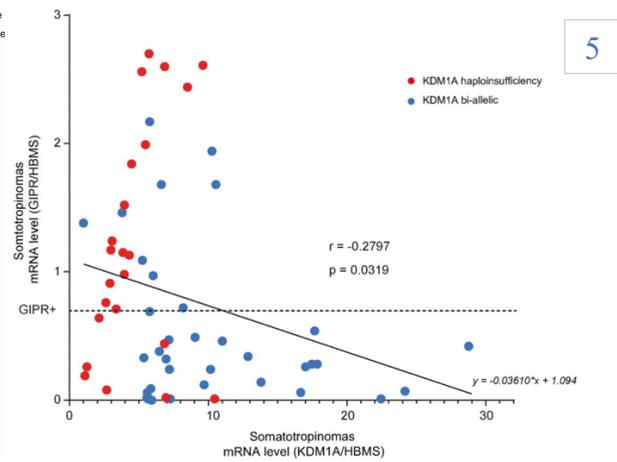
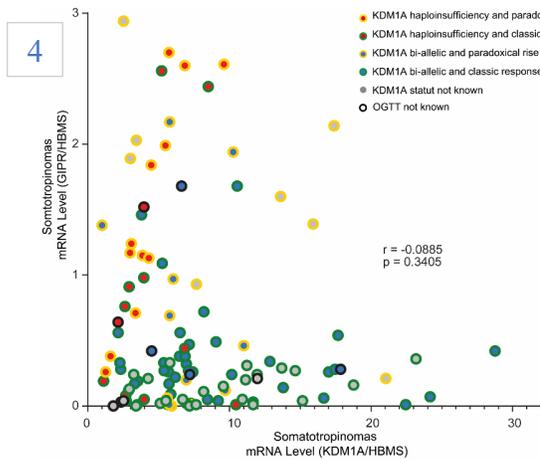


figura 4 e 5: correlazione dei livelli di espressione relativa di GIPR e KDM1A in tutti i campioni di adenoma ipofisario somatotropo con i dati disponibili (a sinistra) e solo nei campioni senza mutazioni attivanti in GNAS (a destra)

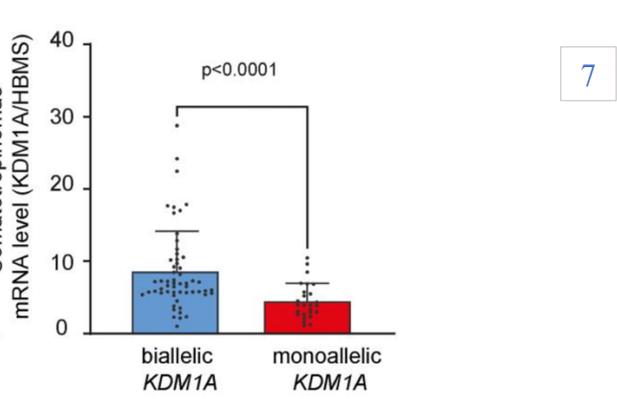
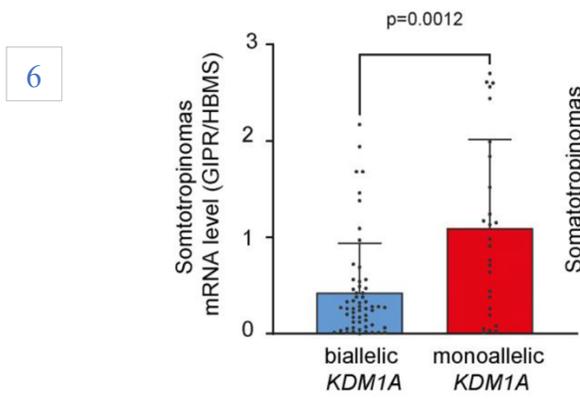


Figura 6 e 7: livelli di espressione di GIPR e KDM1A valutati in campioni con e senza la delezione di un allele di KDM1A. I campioni deleti mostrano una maggiore espressione di GIPR e minore di KDM1A

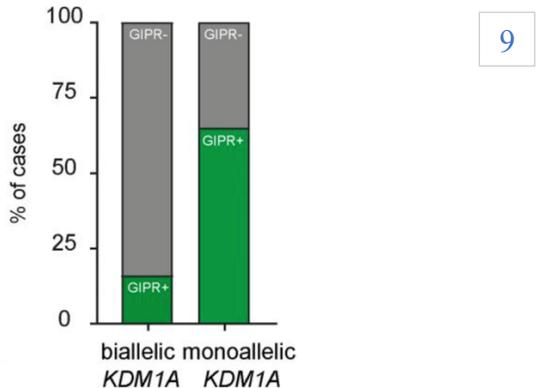
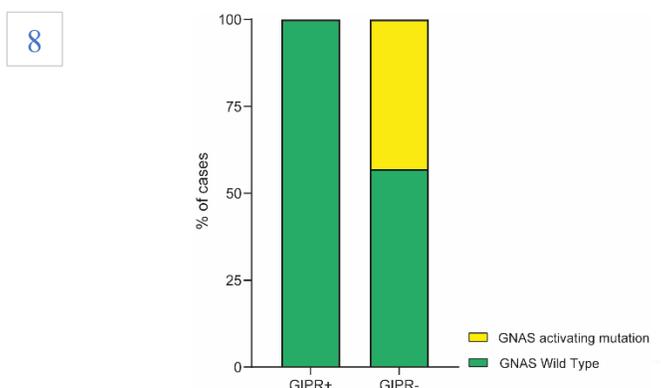


Figura 8 e 9: percentuale di pazienti portanti la mutazione attivante nel gene GNAS suddivisi nei gruppi GIPR+ e GIPR- (a sinistra) e percentuale di pazienti classificati come GIPR+ e GIPR- suddivisi sulla base del numero di alleli di KDM1A rilevati

3.4 Risultati della corsa elettroforetica in gel

Per poter rilevare l'eventuale delezione di uno dei due alleli di *KDM1A* anche nei campioni con quantità insufficienti di DNA è stato messo a punto protocollo basato sulla rilevazione di particolari marcatori molecolari, gli STR, che mappano nelle regioni genomiche adiacenti o interne a *KDM1A*, tramite la corsa in elettroforesi capillare.

Inizialmente sul DNA estratto si esegue una multiplex PCR, con lo scopo di isolare e amplificare solo le brevi sequenze contenenti gli STR d'interesse. Vengono anche sempre amplificati un controllo positivo (CP) e un controllo negativo (CN): il primo è un DNA della stessa specie che si è certi venga amplificato e il secondo è solitamente solo acqua, dunque non dovrebbe restituire nessun prodotto di amplificazione.

Poi, per confermare la riuscita della reazione e controllare di aver ottenuto tutte le popolazioni di ampliconi volute, i prodotti di PCR vengono fatti correre elettroforeticamente su un gel d'agarosio, come quello in *figura 10*. Se nel gel non viene visualizzata nessuna banda relativa al CP, significa che si è verificato qualche errore nella preparazione della miscela di reazione o durante la reazione stessa; le possibilità sono varie, dal non funzionamento corretto della DNA polimerasi alla presenza di contaminanti inibitori o al malfunzionamento del termociclatore. Al contrario, nel controllo negativo non dovrebbe essere visualizzata nessuna banda in quanto, durante la reazione di amplificazione, la polimerasi non avrebbe dovuto trovare alcun template; la presenza di una banda è indice di una contaminazione, potenzialmente derivante da reagenti, strumentazione o spazio di lavoro.

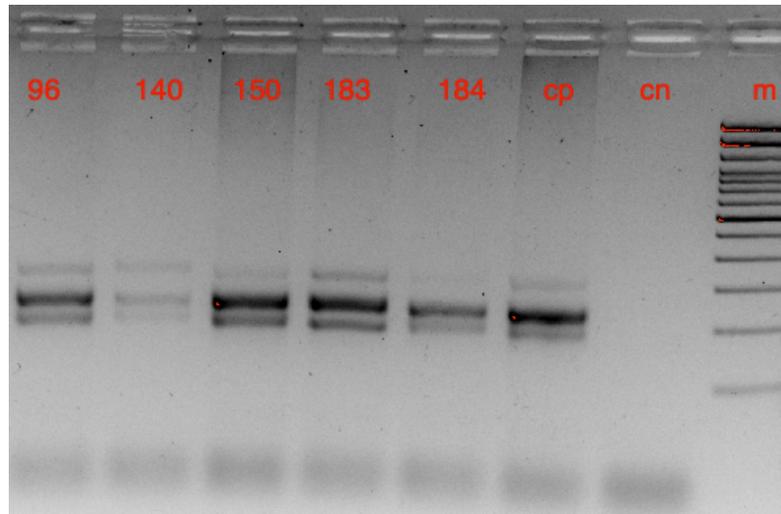


Figura 10: corsa elettroforetica di ampliconi ottenuti da campioni di DNA estratto da tessuto ipofisario tumorale. In tutti i campioni sono visibili le tre bande attese, le cui dimensioni dall'alto al basso sono 293pb, 207pb, 162pb. Il CP e CN confermano il successo della PCR.

3.5 Risultati della corsa elettroforetica capillare

L'elettroferogramma risultante dalla corsa elettroforetica in capillare fornisce informazioni sul genotipo dei campioni analizzati mostrando, per ogni allele rilevato nei tre loci STR d'interesse, un picco. Affinché i risultati del DNA estratto dai campioni di adenoma ipofisario GH-secernente siano correttamente interpretabili è necessario eseguire le stesse analisi anche su un campione di DNA germinale, proveniente ad esempio da sangue, dello stesso paziente così da poter confrontare gli aplotipi ottenuti. Quando nel campione germinale uno o più loci presentano due alleli, ovvero sono eterozigoti, mentre nel DNA tumorale gli stessi risultano omozigoti, è probabile che in quest'ultimo la regione genica interessata abbia subito una delezione; al contrario, se il numero degli alleli presenti coincide per tutti i loci in entrambi i profili, non è avvenuta nessuna delezione. Nella *figura 11* si osserva il profilo di un paziente GIPR⁻ privo di delezione mentre nella *figura 12* è evidenziata la delezione di un allele in uno dei loci esaminati, appartenente a un paziente GIPR⁺. Le frecce rosse evidenziano gli alleli dei tre STR considerati e i picchi arancioni derivano dal LIZ, marcatore di peso molecolare utilizzato che permette al software di calcolare la dimensione (S) dei frammenti. L'altezza dei picchi (H) è data dall'intensità di fluorescenza rilevata dunque è proporzionale alla quantità di frammenti di DNA presenti; analizzare campioni in cui la concentrazione di DNA è molto limitata oppure eseguire una diluizione eccessiva dopo la PCR possono portare a ottenere elettroferogrammi con picchi molto bassi o del tutto assenti. La mancanza completa di alleli in un locus STR, visibili invece negli altri loci analizzati, è probabilmente dovuta a un malfunzionamento dei primers specifici per l'amplificazione del locus assente.

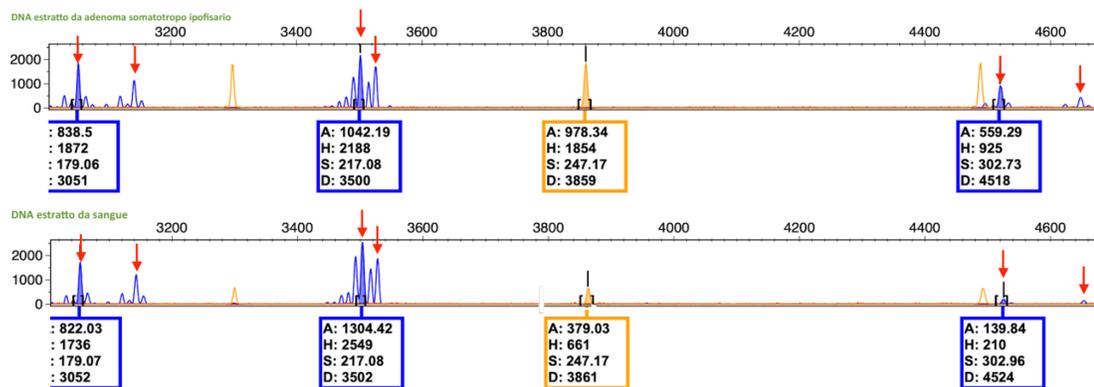


Figura 11: elettroferogramma di un paziente non deleto

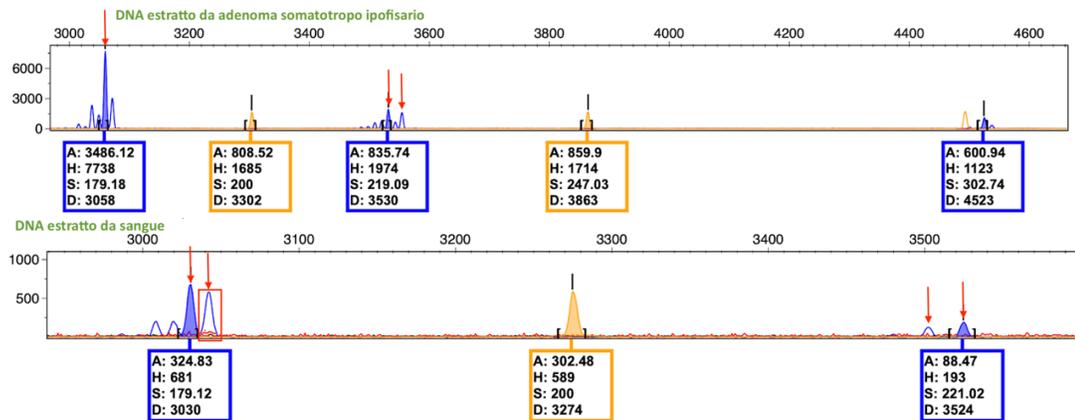


Figura 12: elettroferogramma di un paziente delecto

3.6 Strategie testate per l'eliminazione dei picchi non allelici

L'interpretazione dell'elettroferogramma può essere resa più complessa quando, invece che essere presente per ogni allele un singolo picco ben definito, se ne trova uno principale circondato da picchi di altezza minore; questi sono probabilmente causati dall'aggiunta di una breve coda di adenosine di lunghezza non fissa all'estremità 3'-terminale dei prodotti di PCR catalizzata dalla DNA polimerasi (*non-template addiction*). Così, a partire da un singolo template di data lunghezza, non viene generata una popolazione omologa di ampliconi, come invece ci si aspetterebbe; questo può rischiare di mascherare loci omozigoti come eterozigoti, risultando problematico soprattutto in studi in cui la dimensione degli ampliconi è di fondamentale rilevanza, come nel caso della genotipizzazione tramite STR. Al fine di trovare un protocollo funzionale per eliminare i picchi non allelici, dunque non informativi, dall'elettroferogramma, sono stati valutati diversi metodi potenzialmente efficaci per rendere tutti gli ampliconi della stessa dimensione.

Dopo l'amplificazione, un'aliquota di amplificato non trattata (NT) è stata tenuta per essere corsa direttamente in elettroforesi capillare mentre il restante prodotto è stato suddiviso ulteriormente per essere purificato tramite kit e poi adenilato oppure per essere sottoposto al trattamento enzimatico ExoSAP.

L'elettroferogramma del campione trattato con ExoSAP non differisce da quello del campione NT in numero di picchi, però si osserva una diminuzione nella loro altezza; questo risultato ha portato a supporre che il trattamento potrebbe avere maggior riuscita aumentando la durata dello step di digestione. Così si è ripetuto l'esperimento lasciando agire i due enzimi idrolitici per 30 minuti invece che 15; effettivamente si è osservato un ulteriore abbassamento di alcuni picchi e il cambiamento della loro disposizione in uno dei due loci, come evidenziato dal riquadro rosso della *figura 13*. In un terzo tentativo si è lasciata proseguire la reazione per un'ora ma, al contrario di quanto atteso, la qualità dei picchi non è migliorata.

I due campioni su cui è stata eseguita l'adenilazione successiva alla purificazione tramite kit o con ExoSAP hanno prodotto elettroferogrammi con picchi molto bassi che non permettono di fare nessuna valutazione; probabilmente questo è dovuto ai passaggi intermedi di diluizione che questi protocolli richiedono, oltre alla perdita di DNA che si può verificare durante le purificazioni. Per confermare questa ipotesi si potrebbero

ripetere gli esperimenti partendo però da campioni di DNA maggiormente concentrati oppure eseguire delle quantificazioni tramite NanoDrop dopo i passaggi di purificazione dei templati in modo da rilevare eventuali perdite di DNA e, in base a queste, valutare se sia opportuno proseguire.

Per rendere piatte le estremità sporgenti degli ampliconi si sarebbe potuta utilizzare anche la T4 DNA polimerasi che ha una potente attività esonucleasica in direzione 3'-5' su porzioni a singolo filamento; questo approccio presenta un limite in quanto in presenza di un eccesso di enzima gli ampliconi vengono degradati e inoltre l'attività esonucleasica non si ferma esattamente quando termina la sporgenza, ma rimuove anche una breve porzione del filamento a doppia elica ed è quindi necessario aggiungere al buffer di reazione sufficienti dNTPs in modo che l'enzima stesso possa riparare la sequenza mancante.

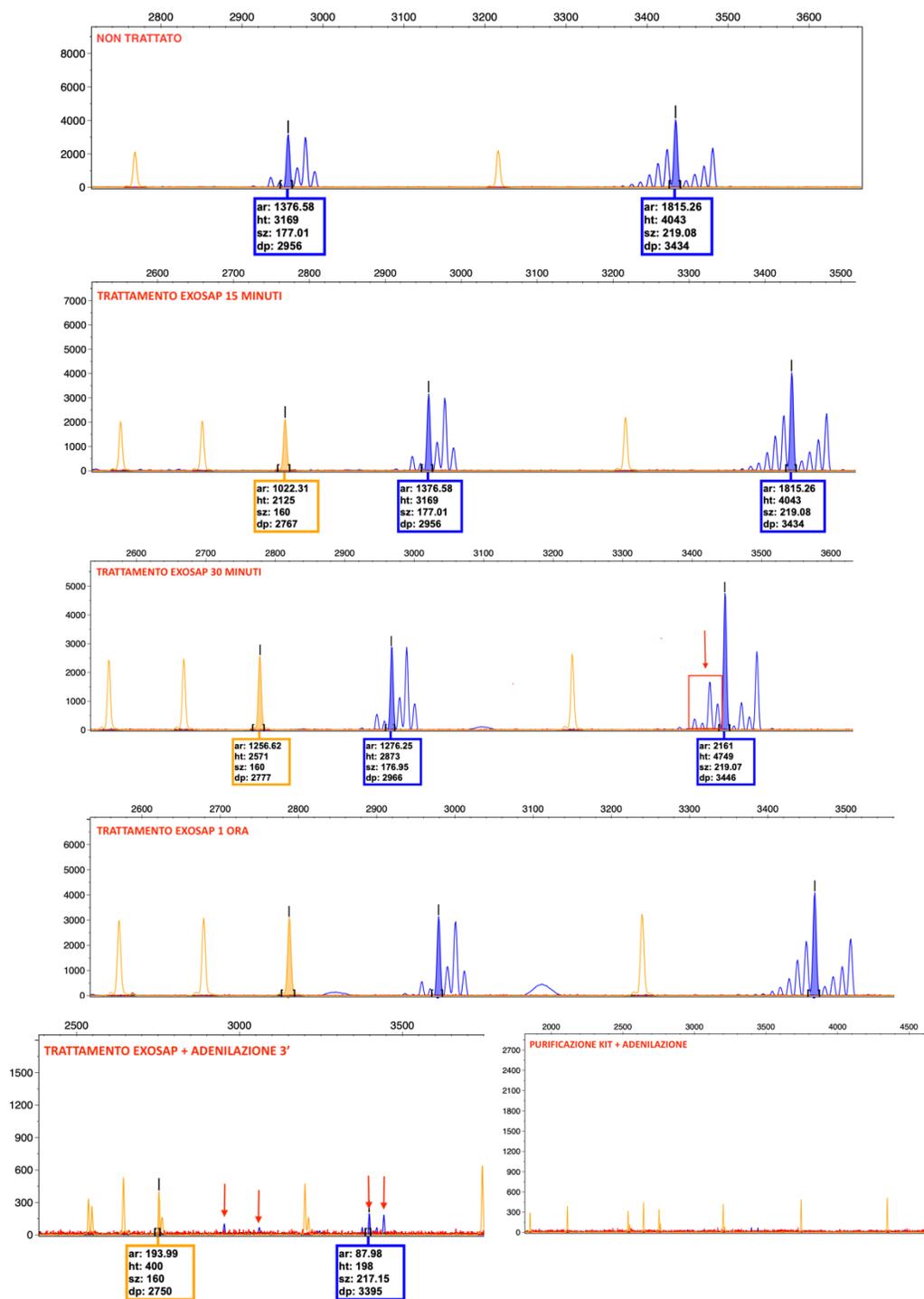


Figura 13 : elettroferogrammi di tutti i protocolli utilizzati al fine di ridurre o eliminare del tutto la presenza dei picchi non allelici

4. CONCLUSIONI

Lo scopo principale di questo lavoro era indagare uno dei possibili meccanismi epigenetici che potrebbe spiegare l'espressione marcata di *GIPR* nei GH-PAs; a tal ragione, le analisi si sono focalizzate sulla rilevazione di alterazioni genetiche nel gene *KDM1A* in quanto di recente è stato dimostrato che sue varianti patogenetiche sono determinanti per l'espressione di *GIPR* nell'iperplasia surrenalica bilaterale causante sindrome di Cushing alimentare. I risultati emersi hanno portato a escludere che i meccanismi di sviluppo di questi due tumori endocrini siano condivisi, ma è stato messo in luce un nuovo aspetto caratterizzante un sottogruppo di campioni tumorali ipofisari analizzati: la delezione più o meno estesa del braccio corto del cromosoma 1, comprendente anche il locus genico *KDM1A*.

A partire da questa nuova rivelazione, l'analisi dei dati si è concentrata sul ricercare un nesso con altre caratteristiche osservate nei pazienti acromegalici, ovvero la risposta di GH paradossa durante l'OGTT e la sovraespressione marcata di *GIPR*.

L'aploinsufficienza di *KDM1A* è stata rilevata maggiormente in adenomi ipofisari derivanti da pazienti GH-Par e inoltre più della metà sono risultati appartenenti anche al gruppo $GIPR^+$; è chiaro che in qualche modo queste tre variabili siano connesse, ma l'esistenza di campioni tumorali aploinsufficienti appartenenti a pazienti del gruppo GH-NPar e $GIPR^-$ suggerisce il coinvolgimento anche di altri meccanismi, forse epigenetici.

In uno studio recente è stata condotta un'analisi trascrittomico tramite sequenziamento al fine di rilevare variazioni nell'espressione genica nei tumori ipofisari secernenti GH, suddivisi sulla base della densità di granulazione [10]; potrebbe rivelarsi interessante eseguire lo stesso esperimento a partire da campioni tumorali con e senza aploinsufficienza di *KDM1A*, così da poter individuare altri geni significativamente sovraespressi o repressi, potenzialmente appartenenti agli stessi pathways in cui è coinvolto *GIPR*.

Per quanto riguarda la presenza di picchi non allelici negli elettroferogrammi dei campioni analizzati, con i trattamenti utilizzati si sarebbe dovuto osservare uno spostamento dei picchi da destra verso sinistra a formarne uno unico, rappresentante una singola popolazione di ampliconi tutti delle stesse dimensioni; il trattamento con ExoSAP per 30 minuti ha prodotto i risultati migliori, pur non ancora pienamente soddisfacenti.

Da un'analisi condotta a posteriori, è stato ipotizzato che la differenza nelle dimensioni dei prodotti di PCR non sia dovuta esclusivamente all'attività di adenilazione all'estremità 3' della DNA polimerasi ma sia il risultato di un appaiamento errato tra i due filamenti di DNA errato durante la fase di sintesi. Questi picchi aspecifici, definiti stutter, derivano da ampliconi più corti di una o più unità ripetute rispetto all'allele reale e sono originati molto più frequentemente durante l'amplificazione dei microsatelliti dinucleotidici. Il modello di scivolamento (*slippage model*) spiega questo fenomeno: durante la sintesi del nuovo filamento, l'enzima DNA polimerasi può temporaneamente dissociarsi dal prodotto di PCR in formazione lasciando così che i

filamenti di DNA si separino parzialmente per poi riappaiarsi, potenzialmente generando un disallineamento a causa della presenza delle brevi sequenze ripetute. Nell'elettroferogramma, soprattutto nei loci omozigoti, si genera un pattern identificativo formato da quattro o più picchi di altezza crescente da sinistra verso destra. Per ridurre la formazione dei picchi stutter diversi studi suggeriscono di modificare le condizioni di amplificazione, abbassando la temperatura di estensione e variando la concentrazione di ione Mg^{2+} : potrebbe rivelarsi vantaggioso e utile condurre alcuni tentativi [11, 12]

BIBLIOGRAFIA

- [1] Chasseloup F, Regazzo D, Tosca L, Proust A, Kuhn E, Hage M, Jublanc C, Mokhtari K, Dalle Nogare M, Avallone S, Ceccato F, Tachdjian G, Salenave S, Young J, Gaillard S, Parker F, Boch AL, Chanson P, Bouligand J, Occhi G, Kamenický P. KDM1A genotyping and expression in 146 sporadic somatotroph pituitary adenomas. *Eur J Endocrinol.* 2024 Feb 1;190(2):173-181. doi: 10.1093/ajendo/lvae013. PMID: 38330165.
- [2] Fleseriu M, Langlois F, Lim DST, Varlamov EV, Melmed S. Acromegaly: pathogenesis, diagnosis, and management. *Lancet Diabetes Endocrinol.* 2022;10(11):804-82. [https://doi.org/10.1016/S2213-8587\(22\)00244-3](https://doi.org/10.1016/S2213-8587(22)00244-3) link diretto <https://www.sciencedirect.com/science/article/pii/S2213858722002443>
- [3] Regazzo D, Losa M, Albiger NM, Terreni MR, Vazza G, Ceccato F, Emanuelli E, Denaro L, Scaroni C, Occhi G. The GIP/GIPR axis is functionally linked to GH-secretion increase in a significant proportion of gsp- somatotropinomas. *Eur J Endocrinol.* 2017 May;176(5):543-553. doi: 10.1530/EJE-16-0831. Epub 2017 Feb 8. PMID: 28179449.
- [4] Ershadinia N, Tritos NA. Diagnosis and Treatment of Acromegaly: An Update. *Mayo Clin Proc.* 2022 Feb;97(2):333-346. doi: 10.1016/j.jmayo.2021.11.007. PMID: 35120696.
- [5] Álvarez-Nava F, Lanes R. GH/IGF-1 Signaling and Current Knowledge of Epigenetics; a Review and Considerations on Possible Therapeutic Options. *Int J Mol Sci.* 2017 Oct 5;18(10):1624. doi: 10.3390/ijms18101624. PMID: 28981462; PMCID: PMC5666699.
- [6] D Regazzo, M Losa, N M Albiger, M R Terreni, G Vazza, F Ceccato, E Emanuelli, L Denaro, C Scaroni, G Occhi, The GIP/GIPR axis is functionally linked to GH-secretion increase in a significant proportion of gsp- somatotropinomas, *European Journal of Endocrinology*, Volume 176, Issue 5, May 2017, Pages 543–553, <https://doi.org/10.1530/EJE-16-0831>
- [7] Regazzo D, Barbot M, Scaroni C, Albiger N, Occhi G. The pathogenic role of the GIP/GIPR axis in human endocrine tumors: emerging clinical mechanisms beyond diabetes. *Rev Endocr Metab Disord.* 2020 Mar;21(1):165-183. doi: 10.1007/s11154-019-09536-6. PMID: 31933128.
- [8] Hage M, Chaligné R, Viengchareun S, Villa C, Salenave S, Bouligand J, Letouzé E, Tosca L, Rouquette A, Tachdjian G, Parker F, Lombès M, Lacroix A, Gaillard S, Chanson P, Kamenický P. Hypermethylator Phenotype and Ectopic GIP Receptor in GNAS Mutation-Negative Somatotropinomas. *J Clin Endocrinol Metab.* 2019 May 1;104(5):1777-1787. doi: 10.1210/jc.2018-01504. PMID: 30376114.
- [9] Dalle Nogare M, D'Annunzio S, Vazza G, Regazzo D, Picello L, Denaro L, Voltan G, Scaroni C, Ceccato F, Occhi G. The Methylation Analysis of the Glucose-Dependent Insulinotropic Polypeptide Receptor (GIPR) Locus in GH-Secreting Pituitary Adenomas. *Int J Mol Sci.* 2023 May 25;24(11):9264. doi: 10.3390/ijms24119264. PMID: 37298217; PMCID: PMC10252571.
- [10] Kim K, Kim Y, Kim SH, Moon JH, Kim EH, Lee EJ, Oh CM, Ku CR. Differential gene expression and pathway analysis in growth hormone-secreting pituitary tumors according to granulation pattern.

Front Oncol. 2024 Jul 30;14:1423606. doi: 10.3389/fonc.2024.1423606. PMID: 39139281; PMCID: PMC11319302.

[11] Seo SB, Ge J, King JL, Budowle B. Reduction of stutter ratios in short tandem repeat loci typing of low copy number DNA samples. *Forensic Sci Int Genet.* 2014 Jan;8(1):213-8. doi: 10.1016/j.fsigen.2013.10.004. Epub 2013 Oct 17. PMID: 24315611.

[12] Mulero JJ, Chang CW, Hennessy LK. Characterization of the N+3 stutter product in the trinucleotide repeat locus DYS392. *J Forensic Sci.* 2006 Sep;51(5):1069-73. doi: 10.1111/j.1556-4029.2006.00227.x. PMID: 17018083.

KDM1A genotyping and expression in 146 sporadic somatotroph pituitary adenomas

Fanny Chasseloup,¹ Daniela Regazzo,² Lucie Tosca,³ Alexis Proust,⁴ Emmanuelle Kuhn,^{5,6} Mirella Hage,¹ Christel Jublanc,⁶ Karima Mokhtari,⁷ Mattia Dalle Nogare,⁸ Serena Avallone,² Filippo Ceccato,^{2,9} Gerard Tachdjian,³ Sylvie Salenave,¹ Jacques Young,¹ Stephan Gaillard,⁵ Fabrice Parker,¹⁰ Anne-Laure Boch,⁵ Philippe Chanson,¹ Jerome Bouligand,^{3,4,11} Gianluca Occhi,^{8,*†} and Peter Kamenický^{1,*†}

¹Inserm, Université Paris-Saclay, Physiologie et Physiopathologie Endocrinienne, AP-HP, Hôpital Bicêtre, Service d'Endocrinologie et des Maladies de la Reproduction, Centre de Référence des Maladies Rares de l'Hypophyse, 94270 Le Kremlin-Bicêtre, France

²Department of Medicine, Endocrinology Unit, University of Padova, 35128 Padova, Italy

³AP-HP, Hôpital Antoine Bécère, Université Paris-Saclay, Service d'Histologie, Embryologie et Cytogénomique, 92140 Clamart, France

⁴AP-HP, Hôpital Bicêtre, Service de Génétique Moléculaire et d'Hormonologie, 94270 Le Kremlin Bicêtre, France

⁵AP-HP, Hôpital Pitié-Salpêtrière, Service de Neurochirurgie, 75013 Paris, France

⁶AP-HP, Hôpital Pitié-Salpêtrière, Unité Hypophyse, 75013 Paris, France

⁷AP-HP, Institut du Cerveau—Paris Brain Institute—ICM, Sorbonne Université, Inserm, CNRS, Hôpitaux Universitaires La Pitié Salpêtrière—Charles Foix, Service de Neuropathologie, Onconeurothèque, 75013 Paris, France

⁸Department of Biology, University of Padova, 35128 Padova, Italy

⁹Endocrine Disease Unit, University-Hospital of Padova, 35128 Padova, Italy

¹⁰AP-HP, Hôpital Bicêtre, Service de Neurochirurgie, 94270 Le Kremlin-Bicêtre, France

¹¹Inserm, Physiologie et Physiopathologie Endocrinienne, Université Paris-Saclay, Le Kremlin-Bicêtre 94270, France

*Corresponding authors: Department of Biology, University of Padova, Via Ugo Bassi 58/B, Padova 35128, Italy. Email: gianluca.occhi@unipd.it (G.O.); Faculté de Médecine Paris-Saclay, Inserm U1185, 63 rue Gabriel Péri, Le Kremlin Bicêtre Cedex 94276, France. Email: peter.kamenicky@universite-paris-saclay.fr (P.K.).

Abstract

Importance: A paradoxical increase of growth hormone (GH) following oral glucose load has been described in ~30% of patients with acromegaly and has been related to the ectopic expression of the glucose-dependent insulinotropic polypeptide (GIP) receptor (GIPR) in somatotropinomas. Recently, we identified germline pathogenic variants and somatic loss of heterozygosity of lysine demethylase 1A (*KDM1A*) in patients with GIP-dependent primary bilateral macronodular adrenal hyperplasia with Cushing's syndrome. The ectopic expression of *GIPR* in both adrenal and pituitary lesions suggests a common molecular mechanism.

Objective: We aimed to analyze *KDM1A* gene sequence and *KDM1A* and *GIPR* expressions in somatotroph pituitary adenomas.

Settings: We conducted a cohort study at university hospitals in France and in Italy. We collected pituitary adenoma specimens from acromegalic patients who had undergone pituitary surgery. We performed targeted exome sequencing (gene panel analysis) and array-comparative genomic hybridization on somatic DNA derived from adenomas and performed droplet digital PCR on adenoma samples to quantify *KDM1A* and *GIPR* expressions.

Results: One hundred and forty-six patients with sporadic acromegaly were studied; 72.6% presented unsuppressed classical GH response, whereas 27.4% displayed a paradoxical rise in GH after oral glucose load. We did not identify any pathogenic variant in the *KDM1A* gene in the adenomas of these patients. However, we identified a recurrent 1p deletion encompassing the *KDM1A* locus in 29 adenomas and observed a higher prevalence of paradoxical GH rise ($P = .0166$), lower *KDM1A* expression (4.47 ± 2.49 vs 8.56 ± 5.62 , $P < .0001$), and higher *GIPR* expression (1.09 ± 0.92 vs 0.43 ± 0.51 , $P = .0012$) in adenomas from patients with *KDM1A* haploinsufficiency compared with those with 2 *KDM1A* copies.

Conclusions and relevance: Unlike in GIP-dependent primary bilateral macronodular adrenal hyperplasia, *KDM1A* genetic variations are not the cause of *GIPR* expression in somatotroph pituitary adenomas. Recurrent *KDM1A* haploinsufficiency, more frequently observed in *GIPR*-expressing adenomas, could be responsible for decreased *KDM1A* function resulting in transcriptional derepression on the *GIPR* locus.

Keywords: pituitary, neuroendocrinology, genetics, acromegaly

† G.O and P.K contributed equally with shared senior authorship.

Significance

A paradoxical increase of growth hormone (GH) following oral glucose load has been described in ~30% of patients with acromegaly and has been related to the ectopic expression of the glucose-dependent insulinotropic polypeptide (GIP) receptor (GIPR) in somatotropinomas. The ectopic expression of GIPR also reported in GIP-dependent primary bilateral macronodular adrenal hyperplasia is related to the genetic inactivation of the lysine demethylase 1A (KDM1A). Therefore, we investigated the role of KDM1A in the ectopic GIPR expression in somatotropinomas. We did not identify KDM1A pathogenic variants in the adenomas from 146 sporadic acromegaly patients. However, we identified recurrent KDM1A haploinsufficiency, which was associated with lower KDM1A expression and higher GIPR expression, and more frequently observed in adenomas from patients displaying a paradoxical GH rise.

Importance

Growth hormone (GH)-secreting pituitary adenomas are responsible for chronic GH excess leading to gigantism during childhood and acromegaly in adulthood. In addition to GH hypersecretion, they can be revealed by local compressive mass effect and associated deficiencies in pituitary functions.¹ Despite recent advancements in genetics, their molecular pathogenesis remains incompletely understood. Activating mutations in the *GNAS* gene, encoding the G protein-activating α subunit, are responsible for promoting the development of adenomas and GH hypersecretion in ~30% of somatotropinomas.² Further, around 5% of them result from germline genetic variations predisposing to pituitary tumorigenesis,^{3,4} in the remaining 2/3 of cases, the molecular driver events are not known.

A subset of patients with acromegaly presents a specific biochemical profile with a paradoxical increase in serum GH concentrations following an oral glucose load.⁵ Similar to cases of glucose-dependent insulinotropic polypeptide (GIP)-dependent Cushing's syndrome,⁶ this paradoxical increase in GH appears to be associated with the ectopic expression of the GIP receptor (GIPR) in pituitary adenomas.⁷⁻⁹ We recently reported that GIP-dependent primary bilateral hyperplasia with Cushing's syndrome is a genetic disease caused by germline-inactivating pathogenic variants of the lysine demethylase 1A (*KDM1A*, also known as *LSD1*) with loss of heterozygosity of the second *KDM1A* locus in the adrenal lesions.¹⁰

The aim of the present study was to investigate whether the ectopic expression of GIPR in sporadic somatotroph pituitary adenomas resulted from a similar mechanism of *KDM1A* genetic inactivation.^{10,11} We conducted targeted exome sequencing (gene panel analysis) in somatotroph pituitary adenoma samples to analyze genes involved in pituitary tumorigenesis and *KDM1A*, performed copy number alteration (CNA) analysis, and assessed the pituitary expressions of *KDM1A* and *GIPR*.

Participants and design

Study design

We conducted a multicenter retrospective cohort study at tertiary referral centers for pituitary diseases in France and Italy. All patients gave their written informed consent for genetic analyses. The study was approved by all local Ethics Committees within guidelines on the ethical conduct of human research consistent with the Declaration of Helsinki. Tissue sample collection was approved by the "Comité de Protection des Personnes Ile de France V" (approval no. 2020-A02899-30) and by the "Azienda Ospedaliera di Padova" ethical committees (approval no. AOP1782).

Patients

Patients included in this study were examined for sporadic acromegaly or gigantism between 2006 and 2021. Acromegaly was diagnosed based on clinical features, elevated Insulin-Like Growth Factor 1 (IGF-1) plasma concentrations, and un-suppressed GH concentrations after an oral glucose tolerance test (OGTT)¹ and by a positive GH staining in somatotropinoma tissue after surgery. Gigantism was defined by excessively rapid growth velocity for age in children or adolescents (>97th percentile) or a final height greater than a +2 SD score above normal for the reference population, together with abnormal IGF-1 values for sex and age. The paradoxical response of GH to the OGTT was defined as an increase of >20% from baseline within 90 min after the glucose load.^{5,12} Based on pituitary magnetic resonance imaging, the adenomas were divided into microadenomas (maximum diameter <10 mm) and macroadenomas (maximum diameter >10 mm). Histological and immunohistochemical confirmation of GH staining was obtained in tumor samples. Mixed somatotroph-lactotroph adenomas displayed immunopositivity for both GH and prolactin. Plurihormonal adenomas displayed immunopositivity for one or more hormones along with GH, besides prolactin. Data regarding clinical parameters, such as gender, age at diagnosis, date of surgery, GH and IGF-1 concentrations, and tumor size, were collected. Invasiveness was defined by the invasion of the cavernous sinus or other adjacent structures identified in magnetic resonance imaging (MRI) studies. Residual tumor was defined as residual adenoma tissue identified on follow-up MRI studies.

Procedures

DNA and RNA extraction methods from somatotropinomas are available in the [Supplementary material](#). Targeted next-generation sequencing (gene panel analysis) was used to sequence high-resolution somatic DNA from all patients (more details are provided in the [Supplementary material](#)) with a panel of genes involved in pituitary diseases, including *AIP*, *CABLES1*, *CDH23*, *CDKN1B*, *DICER1*, *GNAS*, *GPR101*, *MEN1*, *NF1*, *PI3KCA*, *PRKAR1A*, *RET*, *USP8*, and *VHL*. The panel also included the members of the lysine demethylase family involved in human tumorigenesis *KDM1A*, *KDM3A*, *JMJD2A*, *JMJD2B*, *JMJD3*, *KDM6A*, and *KDM5B*.⁷

Oligonucleotide-based array-comparative genomic hybridization (array CGH) analysis was used to analyze genomic imbalances using 60K and 180K oligonucleotide arrays (Agilent Technologies). Copy number alterations were considered of importance if they could be defined by 3 or more oligonucleotides spanning at least 82 and 26 kb (for 60K and 180K arrays, respectively) and were not identified in the Database of

Genomic Variants. Genomic disruption was quantified and defined by 3 groups, affecting under <1% of the genome, between 1% and 12% of the genome, and over 12% of the genome (Supplementary material).

Gene expression analysis was performed on total RNA extracted from tissues. The *GIPR* and *KDM1A* absolute quantification in somatotropinomas was determined by droplet digital polymerase chain reaction using the QX200 AutoDG Droplet Digital PCR System (Bio-Rad) as reported elsewhere.¹³ Probes are detailed in the Supplementary material. For each sample, data are reported in arbitrary unit (ie, the ratio of *GIPR* to *HMBS* [*hydroxymethylbilane synthase*] expression).

Statistical analysis

No sample size calculations were done, and all acromegaly patients with available somatic DNA (extracted from adenoma samples) of sufficient quality for genotyping were included. Data are presented as mean (\pm SD). Fisher's exact tests or χ^2 tests were used to compare categorical variables. Correlation

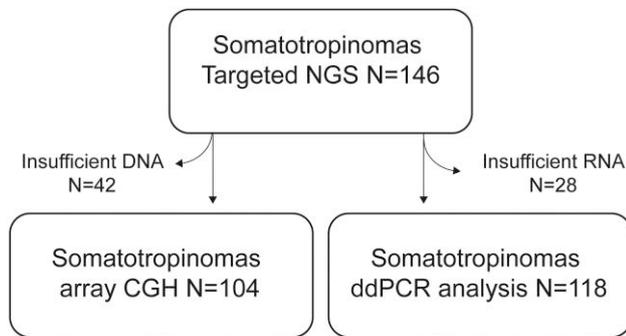


Figure 1. A flowchart of somatotropinomas included in the molecular study.

analysis was made using a 2-tailed Spearman nonparametric correlation analysis and a linear regression model. A nonparametric Mann–Whitney *U* test or a parametric unpaired *t*-test was used to compare 2 groups using the Shapiro–Wilk test to assess normal distribution. Data were analyzed with Prism (version 9.0.0) GraphPad software. *P*-values of <.05 were considered statistically significant (95% CI).

Results

Clinical presentation

The study included 146 acromegaly patients (Figure 1). Among the 146 study patients, 79 (54%) were men and 67 (46%) were women. The mean age at diagnosis was 43.9 ± 13.0 years. Three patients had nonsyndromic gigantism, and the remaining had apparently sporadic acromegaly. Age-adjusted IGF-1 concentration at diagnosis was $226 \pm 183.0\%$ of the upper limit of normal (% ULN). Growth hormone concentrations during the OGTT were expressed as a percentage relative to the baseline levels. Growth hormone concentrations during the OGTT were available for 135 patients (92.4%): 98 patients (72.6%) had a classic pathological GH response profile during the OGTT, and 37 patients (27.4%) displayed a paradoxical rise in their GH concentrations following the oral glucose load. The mean GH increase in the paradoxical response group was $52.8 \pm 27.8\%$ of the baseline values (Figure 2A). A comparison of clinical, biochemical, and tumoral characteristics between patients with and without paradoxical GH rise is provided in Table 1. Age at diagnosis and sex ratio were similar between patients with and without paradoxical responses. The IGF-1 elevation expressed as a percentage of the ULN for age was higher in the paradoxical response group than in the group with a classic GH response. The proportions of macroadenoma to microadenoma were similar in both groups; however, tumors were smaller and less invasive in the paradoxical response group.

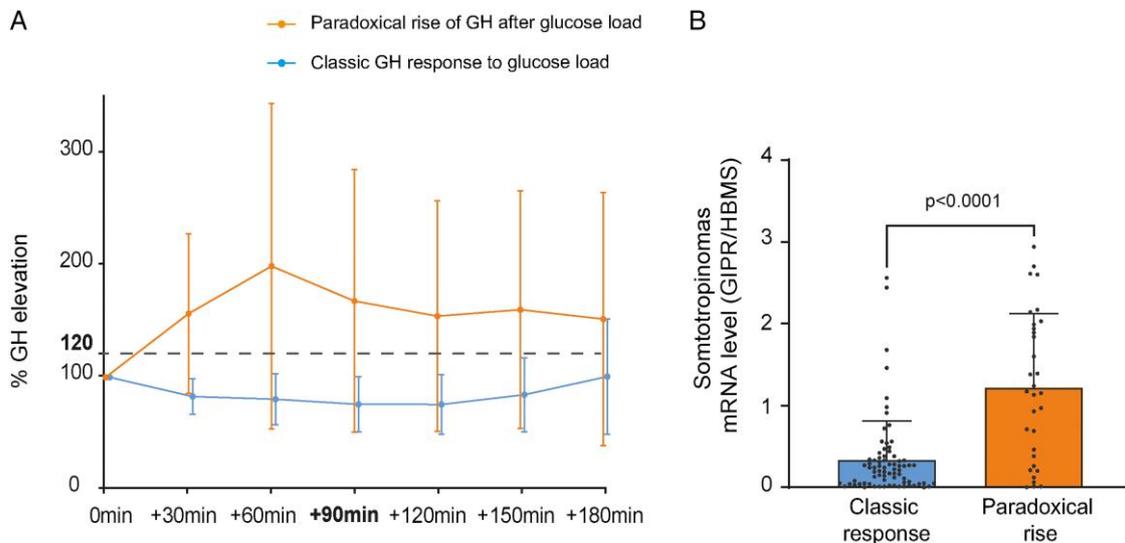


Figure 2. Growth hormone (GH) response during an oral glucose tolerance test (OGTT). (A) GH profiles during the OGTT in patients with (upper line, in orange) and without (lower line, in blue) paradoxical response. GH levels after glucose load are expressed as a percentage of basal GH levels. Data are expressed as mean \pm SD. GH response was classified as paradoxical if the threshold of 120% of baseline GH levels (gray dashed line) was reached within the first 90 min. (B) Relative *GIPR* expression in somatotropinomas quantified by droplet digital PCR (ddPCR). A quantification of *GIPR* expression was performed in 76 patients with a classical response (left blue histogram) and 32 patients with a paradoxical response (right orange histogram). The histograms represent mean and SD, and individual quantifications are shown as points. A comparison was performed by using an unpaired *t*-test. Normal distribution was assessed using a Shapiro–Wilk test.

Table 1. Characteristics of patients.

	Available data OGTT (<i>n</i> = 135)	Paradoxical response of GH after glucose load (<i>n</i> = 37) (27.4%)	Normal response of GH after glucose load (<i>n</i> = 98) (72.6%)	P-value
Clinical parameters				
Age	131	47.5 (36.0; 57.0)	44 (33.0; 52.0)	.19
Male	135	21 (56.8%)	50 (51.1%)	.57
Preoperative medical treatment	126	10 (29.4%)	30 (32.6%)	.83
Residual tumor after surgery	85	11 (57.9%)	43 (65.1%)	.60
Gigantism	135	0	3	>.99
Biological parameters				
IGF-1 (% ULN)	126	374.0 ± 116.7%	310.1 ± 109.9%	.0056
Basal GH (mUI/L)	91	57.3 ± 68.7	68.4 ± 87.8	.60
IGT or DM	93	8 (34.8%)	15 (21.4%)	.26
Prolactin, µg/L	128	15.7 ± 17.0	34.2 ± 106	.0144
Radiological parameters				
Invasive tumor	123	12 (34.3%)	49 (55.7%)	.0452
Maximum tumor diameter at diagnosis (mm)	124	14.8 ± 5.8	17.7 ± 8.3	.0430
Macroadenoma	127	30 (85.7%)	82 (89.1%)	.56
Histological parameters				
Mixed or plurihormonal tumor	126	10 (27.8%)	46 (51.1%)	.0186
Ki-67	108	2.5%	3%	.11

Data are presented as medians (percentile 25%-75%) for continuous variables or mean ± SD and *n* (%) for categorical variables. IGF-1 is expressed as a percentage of the upper limit of normal. Abbreviations: DM, diabetes mellitus; GH, growth hormone; IGT, impaired glucose tolerance test; OGTT, oral glucose tolerance test; ULN, upper limit of normal.

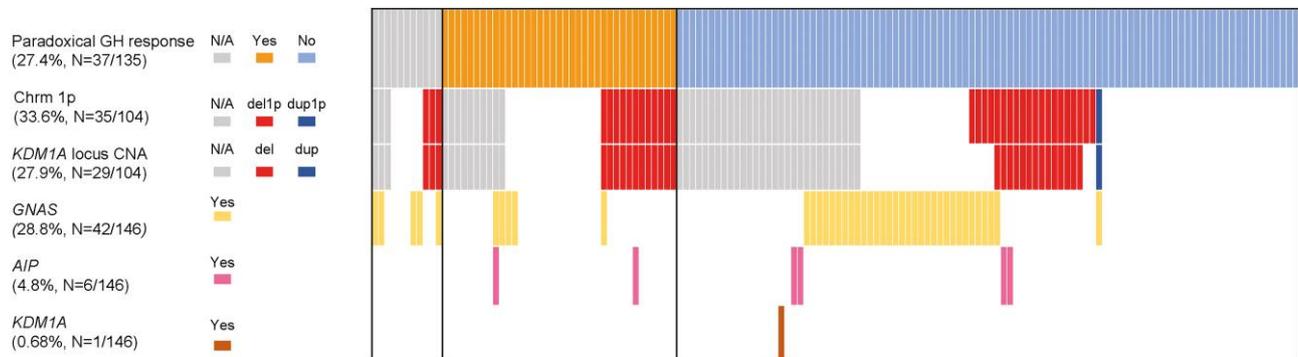


Figure 3. Genetic and cytogenetic abnormalities in 146 somatotropinomas. All 146 somatotropinomas analyzed by targeted exome sequencing (gene panel analysis) and 104 by using an array CGH are shown. Each sample is represented by one column. The first line displays GH profiles after oral glucose load dividing patients in three sections: the orange boxes (middle section) represent somatotropinomas from patients presenting a paradoxical rise and the blue boxes (right section) represent patients with a classical GH response. Left section represents patients without OGTT information. The second and third lines represent array-CGH analyses results, among which the upper line represents the presence or absence of CNA of the short arm of chromosome 1, and the lower line displays an analysis of the *KDM1A* locus. Among patients with a paradoxical rise of GH concentrations (in orange), 32.4% displayed *KDM1A* haploinsufficiency. In patients with a classical response (shown in blue), 14.3% displayed *KDM1A* haploinsufficiency. In the next line, somatic-activating mutations in *GNAS* are represented in yellow. *AIP* pathogenic variants are represented in pink in the following line. In the last line, the VUS of the *KDM1A* gene is represented in brown. Several VUSs, likely benign or benign variants, were found in the *SDH* gene family as well as *NF1*, *CDH23*, *CDKN1B*, *DICER1*, *CABLES1*, *PRKACB*, *RET*, *MEN1*, and *PRKAR1A* (these data are not shown). Missing data are shown in gray.

No differences in histological parameters were observed between the 2 groups.

Genetic analyses

DNA extracted from 146 somatotropinomas was examined by performing a gene panel analysis (Figure 1). We did not identify any *KDM1A* somatic pathogenic variants, independently of their GH profile to the OGTT (Figure 3). One patient presented with a variant of uncertain significance (VUS) of *KDM1A* (NM_015013.4:c.1397C > G [p.Ser466Cys]) that has been reported with a very low frequency in the Genome Aggregation

Database (gnomAD; *f* = 0.00000803). Similarly, we did not identify any pathogenic variants affecting other analyzed genes involved in histone methylation. *GNAS* pathogenic variants (NM_000516.7:c.601C > T [p.Arg201Cys] and NM_000516.7:c.680A > T [p.Gln227Leu]) were found in 38 somatotropinomas. Additionally, 4 patients displayed variations in the *GNAS* gene affecting Arg201 predicted to be pathogenic and previously described in McCune–Albright syndrome: *GNAS* NM_000516.7:c.602G > A (p.Arg201His), *GNAS* NM_000516.7:c.601C > A (p.Arg201Ser), and *GNAS* NM_000516.7:c.601C > G (p.Arg201Gly) raising the overall prevalence of *GNAS* pathogenic variant to 28.8% of

somatotropinomas. Five patients displayed both a paradoxical rise of GH after the OGTT and a *GNAS*-activating variant (Figure 3). *AIP* pathogenic variants were found in 6 patients (4.8%). One of them displayed a somatic pathogenic frameshift variant (NM_001302959:c.467_468del: [p.K156fs]) identified in 45% of overall reads in the tumor DNA, but that was not detected in germline DNA by Sanger sequencing and NGS, suggesting mosaicism (Figure 3). One patient carried both a somatic pathogenic variant of *GNAS* (NM_000516.7:c.601C>G [p.Arg201Gly]) and a somatic variant of *AIP* (NM_001302959:c.508C>T: p.Q170X). One previously reported patient displayed a *GPR101* variant (NM_054021:c.924G>C [p.E308D]).¹⁴ Several VUSs were found in the *SDH* gene family. Finally, only benign or likely benign variants were identified in *NF1*, *CDH23*, *CDKN1B*, *DICER1*, *CABLES1*, *PRKACB*, *RET*, *MEN1*, and *PRKAR1A*.

Copy number alterations

One hundred and four somatotropinomas were analyzed using array CGH, and GH response to the OGTT was available for 96 of them. Overall, we identified a recurrent deletion affecting the short arm of chromosome 1 in 35 patients analyzed (33.6%), encompassing the *KDM1A* locus in 29 (27.9%) patients (Figure 3). Among the 70 somatotropinomas that exhibit a biallelic *KDM1A* gene locus and for which OGTT profiles were available, 15 (21%) and 55 (79%) displayed a paradoxical rise and a classic response of GH after glucose load, respectively. Among patients with only one copy of *KDM1A*, 12 (46%) displayed a paradoxical rise of GH after the OGTT, and 14 (54%) displayed a classical response of GH. The paradoxical rise of GH was therefore more prevalent among patients displaying *KDM1A* haploinsufficiency compared with those with two copies of *KDM1A* ($P = .0166$). Overall array-CGH profiles identified an association with the occurrence of a paradoxical rise and a higher rate of genomic disruption of over 12% [$P = .0292$; odds ratio (OR) = 3.26; 95% CI: 1.13-9.46].

GIPR and *KDM1A* expression analyses

The expression analysis of *GIPR* and *KDM1A* was carried out on 118 of the 146 samples. Of these, 108 samples belonged to patients with available OGTT profiles (29.6% with and 70.4% without a paradoxical response). The relative expression of *GIPR* was 0.33 ± 0.48 in the classical response group and 1.22 ± 0.91 in the paradoxical response group ($P < .0001$; Figure 2B).

Exploring the potential association between a decrease in *KDM1A* expression and an increased expression of *GIPR*, we examined the correlation between these 2 parameters. When expression data of all 118 samples were analyzed, we did not observe a statistically significant correlation ($r = -0.09$, $P = .34$; Figure S1). Focusing solely on 83 cases analyzed by using the array CGH revealed that relative *GIPR* expression was higher in somatotropinomas with *KDM1A* haploinsufficiency ($n = 26$, 1.09 ± 0.92) than in those with conserved *KDM1A* alleles ($n = 57$, 0.43 ± 0.51 , $P = .0012$; Figure 4A). Conversely, and in line with expectations, somatotropinomas exhibiting *KDM1A* haploinsufficiency displayed lower *KDM1A* expression compared with cases without *KDM1A* CNA (4.47 ± 2.49 and 8.56 ± 5.62 , respectively, $P < .0001$; Figure 4B).

Further, we classified samples into *GIPR*⁺ and *GIPR*⁻ groups, according to their *GIPR* expression relative to the normal pituitary.⁷ Among patients with a paradoxical GH response, 68.4% were classified as *GIPR*⁺. Further, only 15.8% of samples biallelic in the *KDM1A* locus were *GIPR*⁺, whereas 65.4% of samples displaying *KDM1A* haploinsufficiency were *GIPR*⁺ ($P < .0001$; Figure 4C). We did not identify *GNAS* mutations in *GIPR*⁺ adenomas (Figure S2). When excluding samples with *GNAS* mutations (in these samples, the principal driver event of tumorigenesis is identified), *GIPR* expression was inversely correlated with the expression of *KDM1A* ($r = -0.2797$, $P = .0319$; Figure 4D).

Conclusions and relevance

This joint large French and Italian cohort of acromegaly patients permitted us to confirm that the paradoxical GH response biologically reveals a subgroup of patients with specific pituitary tumor characteristics. The proportion of patients displaying a paradoxical GH rise after glucose loading diverges between 20% and 37% depending on the criteria used.^{8,9,15-17} Based on our experiences, we have recently redefined the definition and classified a GH response as paradoxical if the GH increased beyond the threshold of 120% of the baseline level within the first 90 min after glucose ingestion.^{5,12} According to this definition, 37 patients of our cohort (27.4%) presented with a paradoxical response pattern. These patients presented with a higher increase of IGF-1, smaller and less invasive tumors that are predominantly pure somatotroph adenomas, as previously reported by us^{9,15} and others.¹⁶⁻¹⁸ Granulation pattern has not been analyzed in this study, but we and others have previously published that somatotropinomas from patients with paradoxical GH response are predominantly densely granulated adenomas.^{9,15,17,19} The OGTT was the gold standard for acromegaly diagnostic confirmation²⁰ as its real diagnostic yield in cases with undoubtedly increased IGF-1 was questionable.²¹ The added value of this test could rely, in the future, on fine phenotyping of patients and their tumors.

Within this subset of pituitary tumors, recent findings from our groups have revealed that a differential *GIPR* expression characterizes around 1/3 of somatotropinomas.⁷⁻⁹ We consider this expression of *GIPR* in certain somatotroph adenomas to be ectopic (rather than an overexpression of a receptor physiologically expressed at lower levels), since the relative expression of *GIPR* in the healthy pituitary is very low, at the limit of significance.⁷ The reasons behind the differential *GIPR* expression at the pituitary level are not yet fully elucidated. However, similar to other endocrine tumors,²² its connection with epigenetic changes has been hypothesized and, to a certain extent, also verified.^{9,23}

This study represents a further step forward in clarifying the mechanism underlying this phenomenon. It is the first attempt to explore whether, as observed in bilateral adrenal hyperplasia associated with GIP-dependent Cushing's syndrome,¹⁰ genetic alterations involving *KDM1A* could lead to changes in *GIPR* expression in GH-secreting pituitary adenomas. Using gene panel analysis, we did not identify any pathogenic variants of *KDM1A* in somatotroph adenomas of patients with acromegaly, regardless of whether they presented a paradoxical GH elevation and/or ectopic *GIPR* expression. Conversely, we observed recurring deletions on the short

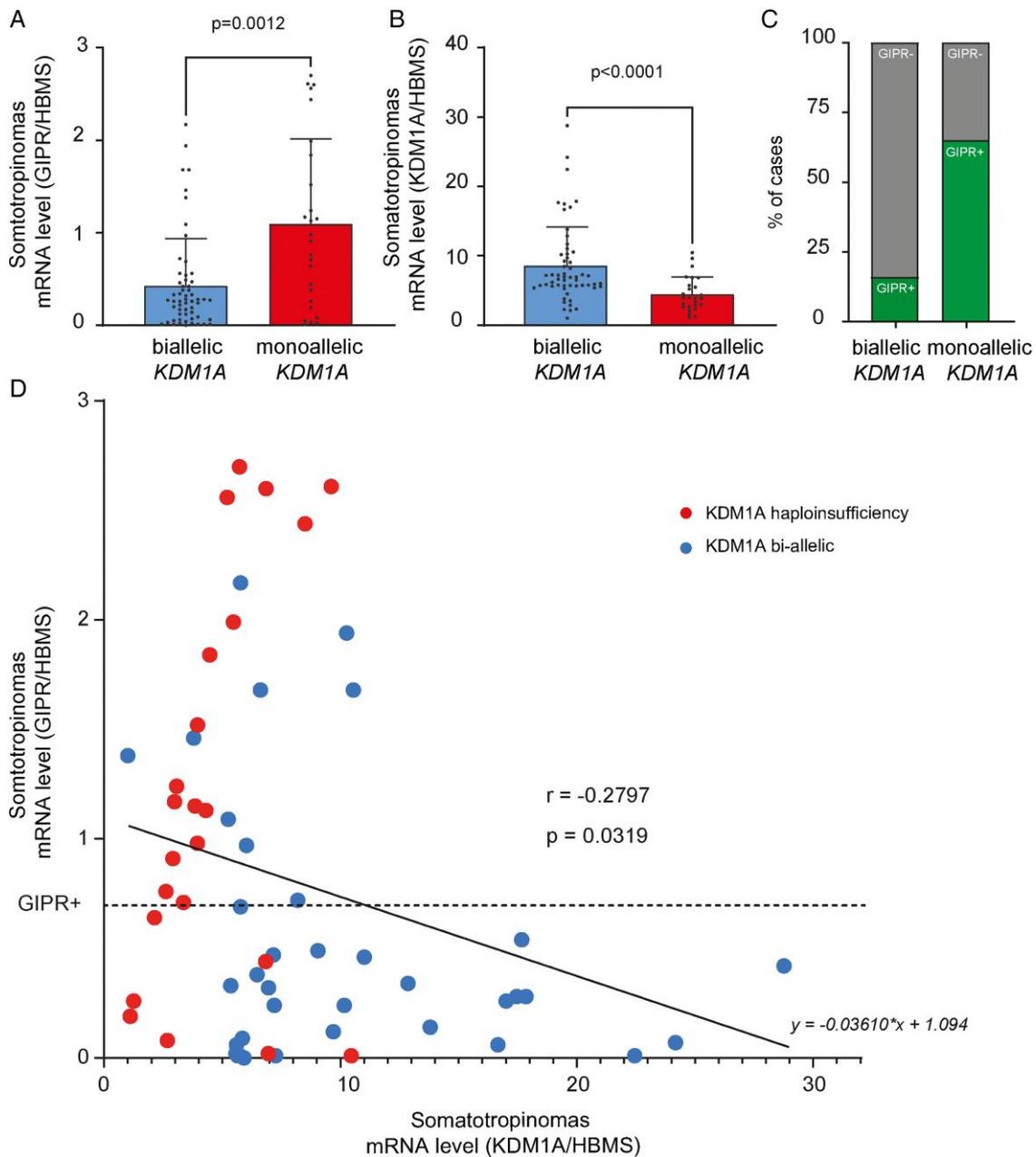


Figure 4. GIPR and KDM1A expressions in somatotroph pituitary adenoma samples. Expression analysis was available in 83 samples with array-CGH data and 57 somatotropinomas with both *KDM1A* alleles and 26 somatotropinomas with *KDM1A* haploinsufficiency. (A) *GIPR* expression was higher (1.09 ± 0.92) in samples displaying *KDM1A* haploinsufficiency than those with both *KDM1A* alleles ($n = 57$, 0.43 ± 0.51 , $P = .0012$). (B) *KDM1A* expression was lower in somatotropinomas from patients with *KDM1A* haploinsufficiency (4.47 ± 2.49 and 8.56 ± 5.62 , respectively, $P < .0001$) than in samples without *KDM1A* copy number alteration. (C) Somatotropinomas were classified according to *GIPR* expression status. Among the samples biallelic in the *KDM1A* locus, 15.8% were classified as GIPR+, while 65.4% were classified as GIPR+ in the samples displaying haploinsufficiency for *KDM1A* ($P < .0001$). (D) Correlation in all samples, excluding samples displaying a *GNAS*-activating mutation: *GIPR* expression was inversely correlated with the expression of *KDM1A* ($r = -0.2797$, $P = .0319$). The dotted line represents a cutoff for somatotropinomas-positive *GIPR* expression (GIPR+). For each sample, data are reported in arbitrary units (ie, the ratio of *GIPR/HMBS* and *KDM1A/HMBS* mRNAs). The histograms represent mean \pm SD. Comparisons were performed by using a Mann–Whitney *U* test. Normal distribution was assessed using a Shapiro–Wilk test. Correlation analysis was made using a Spearman nonparametric correlation analysis and a linear regression model.

arm of chromosome 1, encompassing the *KDM1A* locus, within a specific subgroup of somatotroph adenomas, more frequently associated with the paradoxical GH profile and *GIPR* overexpression. Furthermore, *KDM1A* haploinsufficiency led to a decrease in *KDM1A* expression, in line with a biallelic expression of this gene in normal pituitary tissue. Despite this clear *KDM1A* haploinsufficiency–*GIPR* expression link, no significant correlation was found between

KDM1A and *GIPR* transcript steady-state levels across the complete cohort. On the other hand, a negative correlation emerged when we focused solely on cases lacking *GNAS* pathogenic variants, whose transforming role has been extensively demonstrated. Interestingly, only 3 patients displaying both activation mutation of *GNAS* and *KDM1A* haploinsufficiency were identified further suggesting that *KDM1A/GIPR* contribute to an alternative and maybe mutually exclusive neoplastic

25. Wang J, Hevi S, Kurash JK, *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet.* 2009;41(1):125-129. <https://doi.org/10.1038/ng.268>
26. Hage M, Viengchareun S, Brunet E, *et al.* Genomic alterations and complex subclonal architecture in sporadic GH-secreting pituitary adenomas. *J Clin Endocrinol Metab.* 2018;103(5):1929-1939. <https://doi.org/10.1210/jc.2017-02287>
27. Välimäki N, Demir H, Pitkanen E, *et al.* Whole-genome sequencing of growth hormone (GH)-secreting pituitary adenomas. *J Clin Endocrinol Metab.* 2015;100(10):3918-3927. <https://doi.org/10.1210/jc.2015-3129>

Supplemental Data

Table of Content

Supplementary Data	2
Supplementary methods.....	2
DNA extraction.....	2
Next Generation Targeted Exome Sequencing (Gene Panel Analysis)	2
Oligonucleotide based array-comparative genomic hybridization analysis (array-CGH) .	2
RNA extraction and gene expression analysis	2
Supplementary figures and table	3
Supplementary Table: clinical and genetic characteristics of individual patients.....	3
Supplementary figure S1: GIPR and KDM1A expression in somatotroph pituitary adenoma samples.....	4
Supplementary figure S2: percentage of patient displaying GNAS activating mutation according to GIPR expression status.	10

Supplementary Data

Supplementary methods

DNA extraction

DNA was extracted from somatotropinomas samples using the phenol-chloroform-isoamyl alcohol extraction method, TRIzol reagent (Euromedex), or Quick-DNA/RNA (Zymo Research) according to the manufacturer's instructions.

Next Generation Targeted Exome Sequencing (Gene Panel Analysis)

Whole genome libraries were prepared using NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina (NEB Inc.) robotized on a Biomek Span 8 workstation (Beckman). Enrichment was processed using SureSelect XT custom kit (Agilent) robotized on a Biomek 4000 workstation (Beckman). Paired-end 2×150 bp sequencing was performed by batch of 23 patients on a Miseq® Illumina™ sequencer. Read alignment (aligned to Human genome hg19), variant calling, and annotation were done with Annovar and SNPeff 4.0. Standard QC analysis was performed on the data using FastQC. PCR duplicates were marked with Picardmetrics. The mean coverage was 235 X with 97 % of the bases being covered at 20X.

Oligonucleotide based array-comparative genomic hybridization analysis (array-CGH)

Genomic imbalances were analyzed using 60K and 180K oligonucleotide arrays (Agilent Technologies). DNA from pituitary samples was compared to sex-matched blood donor DNA. Hybridization was performed according to the manufacturer's protocol. The slides were scanned on an Agilent Microarray Scanner. Images processing and data analysis were performed with CytoGenomics software 4.0.3.12 (Agilent Technologies). ADM2 algorithm was used for statistical analysis. Copy number alterations were considered of importance if they could be defined by 3 or more oligonucleotides spanning at least 82Kb and 26Kb (for 60K and 180K arrays, respectively), and were not identified in the Database of Genomic Variants (<http://projects.tcag.ca/cgi-bin/variation/gbrowse/hg19>). The Genome Browser used to analyze genes content was hg19, Build37 (<http://genome.ucsc.edu/>). Genomic disruption was quantified and defined by three groups, affecting under <1% of the genome, between 1 and 12% of the genome and over 12% of the genome.

RNA extraction and gene expression analysis

Total RNA was extracted from tissues using TRIzol reagent (Euromedex) or the Quick-DNA/RNA (Zymo Research). RNA integrity and yields were evaluated by an Agilent 2100 Bioanalyzer (Agilent Technologies) and a Nanodrop spectrophotometer (NanoDrop Technologies), respectively. Before reverse transcription, RNA samples were treated with the Turbo DNA-free kit (Ambion) to remove DNA contamination. RNA was reverse transcribed with Maxima Reverse Transcriptase (Thermofisher) with random hexamers following the manufacturer's instructions. The *GIPR* and *KDM1A* absolute quantification in somatotropinomas was determined by droplet digital PCR (ddPCR) using the QX200 AutoDG Droplet Digital PCR System (Bio-Rad). Probes were obtained from Thermofisher (*GIPR*, Hs00609210_m1; *KDM1A*, Hs01002741_m1; *HMBS*, Hs00609296_g1) while all other chemicals were purchased from Bio-Rad. Before performing samples analyses, a multiplex set-up was optimized. A range of temperatures (from 58 to 62°C, with 2°C intervals), and different relative ratios of targets amplification primer/probe mix (*GIPR*:*HMBS*, 1:1, 0.75:1.25), were tested on a control cDNA and the amplification conditions established. ddPCR was performed as reported elsewhere¹³. Briefly, after dividing the cDNA mixture (i.e., ddPCR™ Super Mix for probes, targets (*GIPR*, 0.75 µl; *KDM1A*, 1.25 µl) and reference (*HMBS*) amplification primer/probe mix, and 10 ng of cDNA template) into 20,000 droplets through a QX200™ Droplet Generator, samples were amplified (10 min at 95°C, 40 cycles of 30 sec at 94°C, and 1 min at 58°C, and a final step at 98°C for 10 min). All cycling steps were performed with a ramp rate of 2°C/sec. After a 4°C O/N incubation to promote droplets stabilization, fluorescence was read on the QX200 Droplet Reader using a two-channel setting to detect FAM (*GIPR* and *KDM1A*) and VIC (*HMBS*) and analyzed by the QX Manager 1.2 Standard Edition. Samples with <10,000 droplets were excluded from the analysis. For each sample, data are reported in arbitrary units (i.e., the ratio of *GIPR*/*HMBS* and *KDM1A*/*HMBS* mRNAs).

Supplementary figures and table

Supplementary Table: clinical and genetic characteristics of individual patients

Patient ID	Age at diagnostic	Gender	Paradoxical response of GH after glucose load	Preoperative treatment	Residual tumor after surgery	IGF-1 (%UIN)	Prolactin concentration (µg/L)	IGT or DM	Invasive	Maximum tumor diameter	Immunohistochemistry	Ki-67	ArrayCGH	ddPCR expression analysis	GNAS NM_000516.7 (VAF*)	AIP NM_001302960 & NM_001302959 (VAF)	KDM1A NM_015013 (VAF)
1	48	M	NO	NO	NO	338	9	YES	YES	10	GH, PRL	3	Available	NA**	c.601C>T p.(Arg201Cys) (35%)		
2	33	M	NO	NO	YES	416	3.4	NO	YES	17	GH	3	Available	Available			
3	52	F	NO	NO	NO	494	10	NO	NO	15	GH	2	Available	Available	c.601C>T p.(Arg201Cys) (43.7%)		
4	48	F	NO	NO	YES	251	14	IGT	YES	17	GH,PRL	3	Available	Available			
5	52	M	NO	NO	YES	465	9	NO	YES	15	GH	6	Available	Available			
6	37	F	NO	NO	YES	375	9.1	NO	YES	11	GH	3	Available	NA			
7	47	F	NO	YES	YES	429	8.32	NO	YES	16	GH	2	NA	Available	c.601C>T p.(Arg201Cys) (43.4%)		
8	27	M	NO	YES	YES	340	185	NO	YES	17	GH,PRL	6	Available	Available			
9	35	F	NO	YES	YES	343	16.5	NO	YES	10	GH,PRL	2	Available	Available			
10	51	M	NO	NO	YES	194	35.5	YES	YES	17	GH,PRL	3	Available	Available	c.680A>T p.(Gln227Leu) (35.1%)		
11	47	M	NO	NO	YES	446	18	NO	YES	29	GH,PRL,TSH	2	Available	Available			
12	33	M	NO	NO	YES	524	14.5	YES	YES	22	GH	3	Available	Available	c.601C>T p.(Arg201Cys) (28%)		
13	39	M	NO	YES	NO	308	6.1	NO	NO	18	GH	1	Available	NA			
14	57	M	NO	NO	NO	364	6.6	IGT	NO	11	GH,PRL	1.5	Available	Available	c.601C>T p.(Arg201Cys) (45%)		
15	27	M	NO	NO	YES	271	28	NO	YES	27	NA	NA	NA	Available			
16	22	M	NO	NO	NO	319	16.4	NO	NO	27	GH,PRL	7	Available	Available			
17	28	F	NO	YES	YES	228	52	NO	YES	23	GH	2.5	Available	Available			
18	38	M	NO	NO	NO	383	67.5	NO	NO	15	GH,PRL	5	Available	Available	c.680A>T p.(Gln227Leu) (53.8%)		
19	67	F	YES	NO	YES	486	5.4	YES	YES	12	GH	3	Available	NA			
20	56	F	NO	NO	NO	279	8.7	NO	NO	9	GH,αSU	1	Available	Available			
21	23	M	NO	NO	NO	284	16.5	NO	NO	12	GH,PRL	1	Available	Available	c.601C>T p.(Arg201Cys) (49.2%)		
22	17	M	NO	NO	YES	198	24.5	NO	YES	32	GH	5	Available	Available			
23	26	M	NO	YES	YES	330	17	IGT	YES	59	GH	6	Available	Available			
24	73	F	NO	NO	YES	511	43	NO	YES	34	GH, PRL, LH	5	Available	Available			

25	42	F	NO	NO	YES	174	41	NO	NO	YES	24	GH, PRL	NA	Available	Available	NM_001302960: c.2T>C;p.M1T (85%)
26	37	M	NO	NO	YES	463	25.5	YES	NO	NO	18	GH, PRL	1	Available	Available	c.680A>T p.(Gln227Leu) (39.7%)
27	21	M	NO	YES	YES	319	17	NO	NO	YES	25.5	GH, PRL	3	Available	Available	c.601C>T p.(Arg201Cys) (50.5%)
28	48	F	NO	YES	YES	327	5.6	YES	YES	YES	23	GH	0	Available	NA	
29	37	F	NO	YES	NO	153	11.8	NO	NO	NO	8	GH	1	Available	NA	
30	57	F	NO	NO	YES	153	11.5	NO	YES	YES	20	GH	2	Available	Available	
31	47	M	YES	NO	NO	581	10.5	NO	NO	NO	12	GH	4	Available	Available	
32	48	M	YES	NO	YES	544	6.1	IGT	YES	YES	14	GH	4	Available	Available	
33	30	M	YES	YES	YES	429	32	NO	NO	YES	26	GH	1	Available	Available	
34	28	M	YES	NO	YES	329	7	NO	NO	YES	10	GH	4	Available	Available	
35	39	M	YES	NO	NO	492	5.7	IGT	NO	NO	9	GH	3	Available	Available	NM_001302959: c.467_468del;p.K156f s (Germine AIP WT) (31.5%)
36	35	M	YES	NO	NO	293	3.4	YES	NO	NO	12	GH	1	Available	Available	
37	32	M	YES	YES	YES	470	26.1	YES	YES	YES	20	GH	4.5	Available	Available	
38	40	M	YES	NO	YES	241	8.6	NO	NO	YES	21	GH	3	Available	Available	
39	49	M	YES	NO	NO	502	9.6	NO	NO	NO	12	GH	2	Available	Available	
40	57	F	YES	NO	NO	409	63.3	IGT	NO	NO	10	GH	3	Available	Available	
41	48	M	NO	NO	NO	141	5.9	YES	NO	NO	11	GH, FSH, LH	1.5	Available	NA	
42	60	M	NO	NO	YES	450	14.4	NO	NO	YES	19	GH, PRL	3	Available	Available	
43	46	M	NO	YES	NO	443	6.2	NO	NO	YES	12	GH	1	Available	NA	
44	63	F	NO	NO	YES	250	36	NO	NO	NA	NA	NA	NA	Available	Available	
45	33	M	YES	NO	NA	200	5.7	NO	NO	NO	9.8	NA	NA	Available	Available	
46	47	M	YES	NO	NO	430	0	NO	NO	NO	11	GH, LH	3	Available	Available	
47	52	M	YES	NO	NO	300	5.9	NO	NO	YES	17	GH	2	Available	Available	
48	29	F	NO	YES	YES	170	24	NO	NO	YES	23	GH	3	Available	NA	
49	48	F	NO	NO	NO	170	8.4	NO	NO	NA	10	GH	3	Available	NA	
50	50	M	NO	NO	NO	300	23.4	NO	NO	NO	17	GH, PRL	2	Available	Available	
51	56	M	NA	NO	YES	NA	NA	NO	NO	YES	NA	GH	2	Available	Available	
52	73	M	NO	NO	YES	188	7.6	YES	NO	NO	25	GH	5	Available	NA	

53	60	F	YES	NO	NO	YES	68.7	YES	NO	25.5	GH, PRL	2	Available	Available	c.601C>T p.(Arg201Cys) (46.1%)
54	40	F	NO	NO	NO	YES	19.7	YES	NO	13	GH, PRL	3	Available	Available	c.601C>T p.(Arg201Cys) (44.8%)
55	36	M	NA	NO	NO	YES	5	NO	YES	NA	GH	3	Available	Available	c.601C>T p.(Arg201Cys) (45.2%)
56	54	M	NO	NO	NO	YES	11.9	NO	YES	40	GH	2	Available	NA	
57	55	M	NO	YES	NO	NO	7.4	NO	NO	9	GH, PRL	1	Available	Available	
58	28	F	NO	NA	NO	NO	33	NO	NA	9	GH	3	Available	NA	
59	NA	M	NA	NA	NA	NA	NA	NA	NA	NA	GH	NA	Available	Available	
60	29	M	NO	NO	NO	NO	NA	NO	YES	26	GH, PRL	4	Available	NA	
61	57	F	YES	YES	YES	YES	7.6	YES	NO	17	GH, PRL, ACTH, αSU	1	Available	NA	c.680A>T p.(Gln227Leu) (44.1%)
62	45	M	YES	NO	NO	NO	6.8	NO	NO	12	GH, PRL, ACTH	1	Available	Available	
63	44	M	NO	NO	YES	NO	13	NO	NO	12	GH, PRL	2	Available	Available	c.601C>T p.(Arg201Cys) (48.9%)
64	36	M	NO	NO	YES	NO	28	NO	NO	23	GH, PRL, FSH, LH, TSH	3	Available	Available	
65	34	M	NO	NO	YES	NO	20.6	NO	YES	4.5	GH, LH and TSH	NA	Available	Available	
66	66	F	NO	NO	NO	NO	9.7	NO	YES	14	GH, PRL, FSH, αSU, TSH	3	Available	Available	c.601C>T p.(Arg201Cys) (41.1%)
67	53	M	NO	NO	YES	NO	9	NO	YES	17	GH, PRL	1	Available	Available	
68	44	M	NO	NO	YES	NO	27	NO	NO	NA	GH, PRL	4	Available	NA	c.601C>T p.(Arg201Cys) (64.7%)
69	54	M	YES	NO	YES	NO	9.2	NO	NO	12	GH	1	Available	Available	c.602G>A (p.Arg201His) (47.5%)
70	28	F	NO	NO	YES	NO	10.4	NO	YES	15	GH, PRL	3	Available	NA	
71	55	M	NO	NO	YES	NO	26	NO	NO	21	GH, αSU	2	Available	Available	c.601C>T p.(Arg201Cys) (64.2%)
72	27	M	NO	NO	NO	NO	8.8	NO	YES	25	GH, PRL	6	Available	Available	c.601C>T p.(Arg201Cys) (53.8%)
73	20	F	NO	NO	YES	NO	22.6	NO	YES	28	GH	5	Available	Available	
74	54	M	NA	NO	NA	YES	12	YES	NO	6	GH	2	Available	NA	c.601C>T p.(Arg201Cys) (53.8%)
75	37	F	NO	NO	YES	NO	36	NO	YES	24	GH, PRL	1	Available	Available	
76	57	F	YES	NO	NO	NO	34	NO	YES	25	GH, FSH, LH	1	Available	Available	
77	49	F	YES	NO	NA	NO	9.4	NO	YES	10	GH, PRL	1	Available	NA	

78	38	M	NO	NO	NO	YES	285	135	YES	YES	28	GH	4	Available	Available	c.680A>T p.(Gln227Leu) (42.5%)
79	35	M	NO	NO	NO	YES	253	7.8	YES	NO	19	GH, PRL	4.5	Available	Available	
80	46	M	NO	NO	NO	NO	290	50	NO	YES	9.6	GH, PRL	4	NA	Available	c.680A>T p.(Gln227Leu) (54.7%)
81	65	F	YES	NO	NO	YES	390	10	NO	NO	15	GH	3	Available	Available	
82	33	F	NO	NO	NO	YES	140	27	NO	YES	18	GH, PRL	2.5	Available	Available	
83	56	M	NA	NO	NO	YES	420	NA	YES	YES	12	GH	2.5	Available	Available	c.601C>T p.(Arg201Cys) (44%)
84	59	F	NA	NO	NA	YES	264	10.2	YES	NO	10	NA	NA	Available	Available	
85	49	M	NO	NO	YES	310	7.8	NO	NO	YES	13	GH	3	Available	Available	
86	38	F	NO	NO	NO	NO	190	10	NO	NO	11	GH, PRL	2	Available	Available	c.601C>T p.(Arg201Cys) (41.8%)
87	30	M	NO	NO	NO	NO	442	4.5	NO	YES	18	GH, αSU	2	Available	Available	
88	30	M	NO	NO	NO	NO	540	17	NO	YES	18	GH	4	NA	Available	
89	60	F	NO	NO	YES	234	11.4	NO	NO	YES	16	GH, PRL	5.5	NA	NA	
90	53	F	NO	YES	YES	581	37	NO	NO	YES	28	GH	4	Available	NA	
91	59	M	NO	NO	NO	NO	429	6.6	YES	NO	18	GH	2.5	Available	Available	c.680A>T p.(Gln227Leu) (49.5%)
92	34	F	NO	NO	YES	232	53	NO	NO	NO	13	GH, PRL	2.5	Available	Available	c.601C>A (p.Arg201Ser) (41.5%)
93	45	M	NO	NO	YES	321	3.6	NO	NO	YES	18	NA	NA	Available	Available	
94	NA	F	NO	NO	NA	NA	NA	4.7	NA	NO	11	NA	NA	Available	Available	c.601C>T p.(Arg201Cys) (47.4%)
95	50	F	NO	NO	NA	154%	15.1	15.1	NA	NO	10	GH	NA	NA	Available	
96	21	M	YES	NO	NA	169%	5.3	5.3	NA	NO	15	GH, PRL	NA	NA	Available	c.601C>T p.(Arg201Cys) (45.4%)
97	51	F	NO	NO	NA	384%	9.5	9.5	NA	NA	9	GH	3	Available	Available	c.680A>T p.(Gln227Leu) (63.6%)
98	NA	F	NA	NA	NA	NA	NA	NA	NA	NA	NA	GH, PRL	3	NA	Available	c.601C>T p.(Arg201Cys) (57.8%)
99	41	F	NO	NO	NA	289%	14.7	14.7	NA	YES	NA	GH	3	NA	Available	
100	32	M	NO	YES	NA	326%	>1000	>1000	NA	NO	10	GH, PRL	NA	Available	Available	NM_001302959: c.733C>T;p.R245X (85%)
101	51	M	YES	NO	NA	421%	4.2	4.2	NA	NO	10	GH	NA	Available	Available	
102	NA	M	NA	NA	NA	NA	NA	NA	NA	NA	NA	GH, FSH	NA	NA	Available	c.680A>T p.(Gln227Leu)

(46.4%)

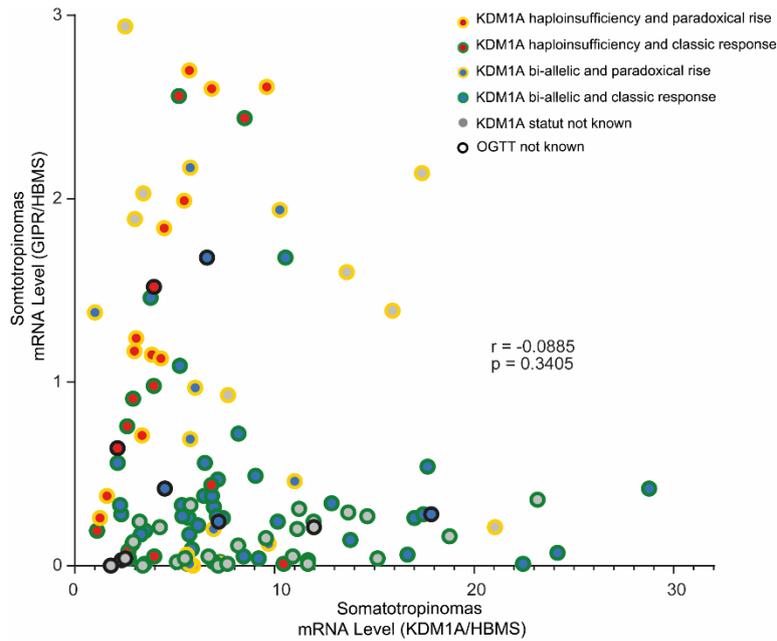
103	57	M	NO	YES	NA	552%	10.4	NA	NO	14	NA	NA	NA	Available
104	37	F	NO	YES	NA	309%	21	NA	NO	16	GH, LH	3	NA	Available
105	51	M	YES	YES	NA	296%	27.1	NA	YES	23	GH, PRL, FSH	NA	NA	Available
106	36	M	YES	NO	NA	391%	10	NA	NO	9	GH	3-5	NA	Available
107	NA	F	NO	NA	NA	NA	NA	NA	NA	NA	GH	NA	NA	NA NM_001302959 c.496C>T;p.Q166X (50%)
108	60	F	YES	YES	NA	404%	55.5	NA	NO	10	GH	NA	NA	Available
109	40	F	NO	YES	NA	243%	23.4	NA	YES	22	GH	NA	NA	NA
110	52	F	NO	YES	NA	179%	6.1	NA	YES	12	GH	3	NA	Available
111	55	M	NO	YES	NA	203%	22.2	NA	NO	20	GH	1	NA	Available
112	36	F	YES	NO	NA	NA	25	NA	YES	24	GH	NA	Available	Available
113	13	F	YES	YES	NA	NA	8.3	NA	NO	28	GH	NA	NA	NA c.601C>G (p.Arg201Gly) (14.5%) NM_001302959 : c.508C>T;p.Q170X (100%)
114	21	F	NO	YES	NA	302%	6.5	NA	YES	27	GH	NA	NA	Available
115	47	F	YES	NO	NA	NA	6.6	NA	YES	10	GH	NA	NA	NA
116	NA	F	YES	NA	NA	NA	NA	NA	NA	NA	GH	NA	NA	Available
117	22	F	NO	NA	NA	NA	NA	NA	NA	NA	GH	3	NA	Available c.680A>T p.(Gln277Leu) (49%)
118	53	M	NO	YES	NA	309%	92.3	NA	NO	13	GH	3.5	NA	Available c.602G>A p.(Arg201His) (39.1%)
119	62	F	NO	NO	NA	179%	13.2	NA	NO	4	GH	3	NA	Available
120	51	M	NA	YES	NA	296%	27.1	NA	YES	NA	NA	NA	Available	Available
121	40	F	NO	YES	NA	114%	6	NA	NO	11	NA	NA	NA	Available
122	53	F	YES	NO	NA	303%	4	NA	NO	9	GH	3	NA	Available
123	44	F	NO	YES	NA	269%	18	NA	YES	20	GH	3	NA	Available
124	60	M	YES	YES	NA	409%	10.6	NA	NO	17	GH	3	NA	Available
125	NA	M	NA	NA	NA	NA	NA	NA	NA	NA	GH	NA	NA	Available
126	67	F	NO	YES	NA	369%	22.7	NA	YES	11	GH	3	Available	Available
127	21	M	NO	NO	NA	449%	4.3	NA	NO	NA	GH	3	Available	Available
128	26	M	NO	YES	NA	455%	28.7	NA	NO	16	GH	1	NA	Available c.1397C>G;p.S466C (45.5%)
129	47	M	YES	NA	NA	431%	9.7	NA	NO	13	GH	3	Available	Available

130	49	F	NO	YES	NA	272%	34	NA	NO	5	GH, PRL	NA	NA	Available
131	69	F	YES	YES	NA	614%	NA	NA	NO	14	GH, PRL	NA	Available	Available
132	NA	F	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	Available	Available
133	NA	F	NO	NA	NA	NA	NA	NA	NA	NA	NA	NA	Available	Available
134	71	F	NO	YES	NA	NA	NA	NA	NA	NA	GH	NA	Available	Available
135	45	F	NO	YES	NA	303%	44.1	NO	NO	13	GH	3	Available	c.601C>T p.(Arg201Cys) (47.3%) Available
136	38	F	NO	YES	NA	361%	5.3	NA	NO	24	GH, PRL	3.5	Available	NA c.601C>T p.(Arg201Cys) (31.8%) Available
137	46	F	YES	YES	NA	245%	26	NO	NO	16	GH	3	Available	Available
138	59	F	YES	YES	NA	186%	19	NO	NO	7	GH	3	NA	Available c.601C>T p.(Arg201Cys) (17.8%) Available
139	35	M	NO	YES	NA	186%	196	NO	YES	30	GH, PRL	3	NA	Available NM_001302959: c.634C>T R212W
140	45	F	NO	NO	YES	254%	16	NO	NO	15	GH, FSH, LH, TSH	3	NA	Available c.601C>T p.(Arg201Cys) (38.6%) Available
141	52	M	NO	NO	NA	294%	6.7	YES	NO	11	GH	3	NA	NA c.680A>T p.(Gln227Leu) (32.5%) Available
142	54	M	NO	NO	NA	352%	10.7	NA	NO	10	GH, PRL	3	NA	NA c.680A>T p.(Gln227Leu) (37%) Available
143	20	M	YES	NA	NA	314%	2.5	NA	NA	NA	GH, PRL	1	NA	Available
144	40	F	NO	NA	NA	407%	15.8	NA	NA	NA	GH	3.5	NA	Available
145	38	F	NO	NA	NA	250%	17.1	NA	NA	16	GH, PRL	3	NA	Available c.601C>T p.(Arg201Cys) (40%) Available
146	51	M	NO	NO	NA	444%	3.3	NA	YES	18	NA	NA	Available	Available

* VAF: Variant Allele Frequency. **NA: Not Available.

Supplementary figure S1: GIPR and KDM1A expression in somatotroph pituitary adenoma samples.

Expression analysis was available in 118 patients. Samples with <10,000 droplets were excluded from the analysis. For each sample, data are reported in arbitrary units (i.e., the ratio of *GIPR/HMBS* and *KDM1A/HMBS* mRNAs). Correlation analysis was made using a logistic regression model.



Supplementary figure S2: percentage of patient displaying GNAS activating mutation according to GIPR expression status.

